UNIVERSITÉ DE STRASBOURG

cnrs dépasser les frontières

Habilitation à Diriger des Recherches

présentée par

Stéphane ARNOUTS

le 5 septembre 2018

Chronique d'une mort annoncée

Le destin des galaxies dans un contexte cosmologique

devant le jury composé de :

Rapporteurs	Hervé Dole Ariane Lançon Simona Mei
Examinateurs	Stéphane Colombi Benoit Famaey
Garant	Laurent Cambresy

Observatoire astronomique
de Strasbourg

Table des matières

1

Contexte

Des résultats d'Edwin Hubble (Hubble 1926) à la future mission Euclid (Laureijs et al. 2011), la cosmologie observationnelle repose en grande partie sur les traceurs privilégiés que sont les galaxies. Ces phares cosmiques mettent l'univers en lumière, depuis ses origines jusqu'à sa remarquable structuration locale à grande échelle (Colless et al. 2003). Au delà de l'intérêt cosmologique de ces traceurs (Perlmutter et al. 1997; Eisenstein et al. 2005), leur origine et leur évolution sont parmi les plus intrigants chapitres de la formation des structures cosmiques. Comment est apparue la séquence de Hubble ? Comment les galaxies acquièrent leur gaz et leur moment angulaire ? Comment est régulée la formation d'étoiles et quel est le rôle des processus rétro-actifs ? L'évolution est-elle prédéfinie en partie par les conditions initiales (nature), ou est elle influencée par l'environnement (*nurture* : fusions ou interactions, accrétions, ...) ? Pour appréhender ces questions, une vision multi-échelles est nécessaire, en s'intéressant à la fois aux propriétés internes des galaxies ainsi qu'à leur environnement immédiat (milieu circum-galactique) ou à grande échelle.

Grâce à l'avalanche de données accumulées au cours de ces vingt dernieres années (Madau & Dickinson 2014) et les développements des simulations numériques et hydrodynamiques (e.g. Springel 2010; Dubois et al. 2012; Kim et al. 2014, pour des comparaisons), une vision cohérente des processus de formation et d'évolution des galaxies a émergé. Ces processus mettent en jeu plusieurs acteurs : la croissance hiérarchique des structures qui se place dans un paradigme cosmologique dont les conditions aux limites sont désormais bien établies (Springel et al. 2006; Planck Collaboration et al. 2016) et le milieu intergalactique qui contient ~80% des baryons (McQuinn 2016). Dans le cadre de la formation hiérarchique des structures, les fluctuations de densité primordiales croissent de manière anisotrope sous l'effet de la gravité (Zel'dovich 1970). Au cours de ce processus, la matière s'écarte des régions sous-denses, s'écoule le long de feuillets, ou murs, qui s'enroulent pour former des filaments, le long desquels la matière s'écoule vers des noeuds, correspondant aux pics de densité, donnant naissance à la toile cosmique (ou *cosmic web*, Bond et al. 1996). Bien que contre-intuitive pour un modèle cosmologique dominé par la matière noire froide, cette formation hiérarchique reproduit remarquablement bien la distribution des galaxies à grande échelle (Frenk & White 2012). Au dessus d'une densité critique, la matière noire s'effondre pour former des halos dans lesquels les baryons accrétés vont pouvoir former les galaxies. La distribution de masses, dominée initialement par les petites masses, évolue au cours du temps via une accrétion continue ou par fusions successives. La physique des baryons est plus complexe. A grande échelle, le gaz suit les gradients de potentiel gravitationnel imposés par la matière noire formant des filaments d'hydrogène ionisé, le milieu inter-galactique (IGM), dans lequel se forment les galaxies. Le gaz accrété au sein des halos refroidit et s'effondre sur un disque (White & Rees 1978), dont le spin est dicté par celui acquis lors de l'effondrement des halos de matière noire (White 1984; Hahn et al. 2009). La formation d'étoiles se déclenche lorsque le gaz froid excède une certaine densité. La quantité de gaz disponible dans les halos est directement régulée par la croissance des halos (Bouché et al. 2010).

Ce cadre théorique a été confronté à la multitude d'observations accumulée ces dernières années.
• Le déclin de l'activité stellaire : La mesure de l'activité stellaire cosmique a montré que l'univers ne cesse de former moins d'étoiles avec le temps et que l'activité stellaire était 10 fois plus intense il y a 10 milliards d'années (e.g. Schiminovich et al. 2005; Madau & Dickinson 2014). Cette faible activité récente, implique que l'essentiel de la masse des galaxies était déjà assemblé (plus de 50%) à cette époque, y compris les galaxies massives (Arnouts et al. 2007; Moutard et al. 2016b), laissant une moindre place aux processus de fusions récentes. L'augmentation du taux de formation d'étoiles cosmique est en fait accompagné d'une augmentation d'activité pour l'ensemble des galaxies actives, quelle que soit leur masse (Noeske et al. 2007b). Ces observations sont corroborées par la plus forte fraction de gaz observée (Daddi et al. 2010), la nature plus perturbée des disques de galaxies lointaines (Kassin et al. 2012). Une telle évolution est aussi consistante avec les prédictions des simulations numériques qui trouvent une accrétion sur les halos plus efficace à grand z. La corrélation, entre l'activité stellaire et la masse des galaxies, nous informe sur le mécanisme dominant qui gouverne la formation d'étoiles dans les galaxies. La faible dispersion dans cette relation favorise une évolution séculaire, basée sur un mode d'accrétion continue gouvernant la formation stellaire, plutôt que des phases de fusions successives, entrainant des épisodes stochastiques d'activité (Noeske et al. 2007a). En revanche, à plus forte masse stellaire, la dispersion augmente et une fraction importante des galaxies dévie de cette relation universelle (Ilbert et al. 2015), avec une formation stellaire affaiblie, signe d'une cessation progressive de leur activité stellaire et d'une migration vers la population de galaxies passives ou elliptiques.

• Le phénomène de *downsizing* : Cette découverte majeure révèle que les galaxies les plus massives ont eu leur pic d'activité stellaire plus tôt dans le passé que celles de plus faible masse (Cowie et al. 1996; Juneau et al. 2005). Cette évolution, apparemment anti-hiérachique, a aussi son origine dans le modèle d'effondrement des halos. La distribution de masse des halos est modifiée par la modulation du champ de densité à grande échelle, induite par les grandes structures. Elle permet aux halos de passer le seuil d'effondrement plus tôt et donc génère plus de halos massifs dans les régions sur-denses du champ de densité (Bond et al. 1991). Ce processus de formation biaisée des galaxies permet d'expliquer la forte corrélation des amas de galaxies (Kaiser 1984), des galaxies lumineuses/massives observées à grand redshift (Adelberger et al. 1998; Arnouts et al. 1999), et explique la bimodalité en couleur de la distribution de galaxies dans l'univers local (Postman & Geller 1984). Dans un tel scénario, les galaxies elliptiques/massives locales se sont formées très tôt lors d'épisodes intenses d'activités stellaires dans les noeuds de la toile cosmique. Les simulations hydrodynamiques montrent que des courants froids peuvent pénétrer au coeur des halos de matière noire et alimenter en gaz de telles galaxies pour soutenir leur activité stellaire, leur permettant d'acquérir très vite, une grande partie de leur masse stellaire (Kereš et al. 2005; Dekel & Birnboim 2006).

• Les mécanismes stoppant l'activité stellaire dans les galaxies (*quenching*) : Par des arguments liés uniquement à la physique non collisionnelle de la matière noire et la formation biaisée des halos, il est possible d'expliquer qualitativement le déclin de la formation d'étoiles durant les 10 derniers milliards d'années, et l'effet de *downsizing*. Toutefois la fonction de masse des halos diffère de celle des galaxies, avec un excès de halos de matière noire aux deux extrémités de masse. Ces différences sont en partie liées à la physique plus complexe des baryons. Plusieurs mécanismes peuvent empécher le gaz de former des étoiles. Les effets rétroactifs liés aux vents stellaires, aux supernovae, probablement dominants pour les galaxies de faible masse, ou les processus violents associés aux noyaux actifs, principalement dominants à forte masse (voir discussion Silk & Mamon 2012). Dans les halos au delà d'une certaine masse critique, le gaz accrété subit un chauffage gravitationnel avec un temps de refroidissement plus long que le temps dynamique (Dekel & Birnboim 2006). Le réservoir de gaz peut aussi être arraché lorsque la galaxie entre dans un halo plus massif ou lors de fusions de spirales (voir Peng et al. 2010, pour le rôle séparé de l'environnement et de la masse). Ces mécanismes mal compris sont gérés par les modèles semi-analytiques ou les simulations hydrodynamiques et les observations peuvent aider à contraindre les recettes physiques.

• Au delà des variations à grande échelle du contraste de densité, l'environnement lié aux structures anisotropes de la toile cosmique impacte-t-il les propriétés des galaxies ? Les simulations prédisent que l'histoire d'assemblage des halos est sensible aux effets de marée. Un halo à proximité d'une structure peut voir son accrétion ralentie ou stoppée ; l'orientation du spin des halos tend à s'aligner avec la structure filamentaire voisine (Pichon et al. 2011; Codis et al. 2012). Les simulations hydrodynamiques montrent que cet effet persiste pour les galaxies (Dubois et al. 2014) et cette orientation privilégiée du spin a récemment été confirmé observationnellement (Trujillo et al. 2006; Tempel et al. 2013). Mesurer le rôle spécifique de ces effets de marée, dans l'assemblage de la masse et du taux de formation stellaire, est moins facile, car il faut pouvoir démêler la contribution spécifique de la géométrie de l'environnement, de celle de la densité locale (Chen et al. 2017; Malavasi et al. 2017; Kraljic et al. 2018). De telles investigations ouvrent toutefois une nouvelle voie d'exploration pour les futurs grands relevés spectroscopiques.

Le scénario de formation et d'évolution discuté ci-dessus est l'accomplissement de multiples observations, accumulées sur l'ensemble du spectre électromagnétique. L'approche multi-longueur d'onde fût essentielle pour passer des grandeurs observationnelles aux grandeurs physiques, permettant des comparaisons directes aux simulations et modèles théoriques. Les grands relevés spectroscopiques, ainsi que les redshifts photométriques, ont permis d'accéder à la dimension temporelle indispensable pour reconstruire cette histoire des galaxies. La taille, de plus en plus significative de ces relevés, a permis de s'affranchir des effets de variance cosmique, dont souffraient les premiers sondages (Somerville et al. 2004) et de dessiner les grandes structures à des redshifts de plus en plus élevés (Garilli et al. 2014), permettant d'analyser le rôle de l'environnement sur une large plage temporelle. La figure 1.1 montre cette évolution, du premier champ profond (Hubble Deep Field, HDF) obtenu avec le HST au dernier grand relevé spectroscopique VIPERS sondant l'univers lointain ($< z > \sim 0.8$).

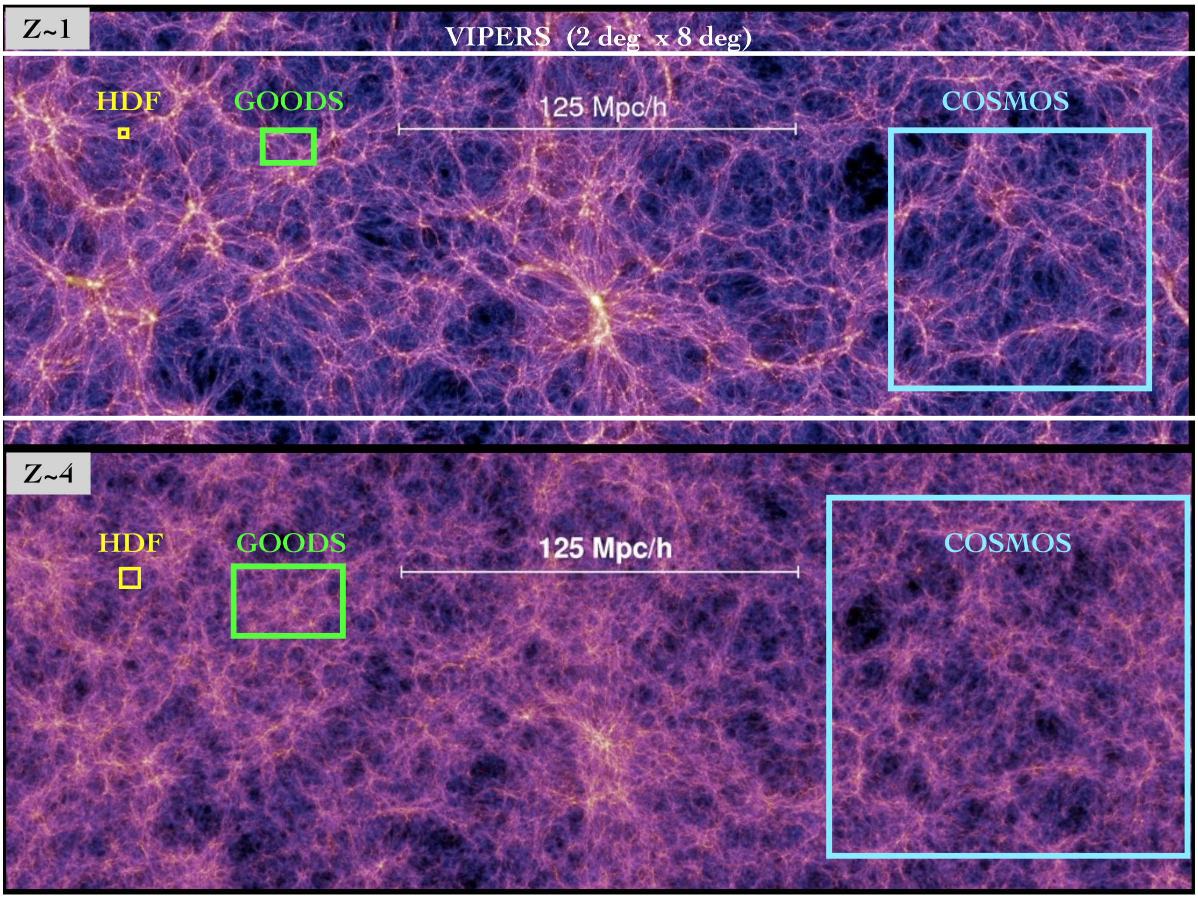


Figure 1.1: Tailles relatives de différents sondages extragalactiques : Hubble Deep Field (2.5×2.5 arcmin^2 ; HDF), Great Observatories Origins Deep Survey (10×16 arcmin^2 ; GOODS), Cosmic Evolution Survey (1.4×1.4 deg^2 ; COSMOS) et VIMOS Public Extragalactic Redshift Survey (8×2 deg^2 ; VIPERS) aux redshifts z∼ 1 (haut) et z∼ 4 (bas). Les images en fond montre l'évolution du champ de densité de la matière noire sur une épaisseur d'∼15 Mpc de la simulation Millenium, quand l'univers était âgé de 4.7 Gyr (haut) et 1 Gyr (bas). Ces images illustrent comment les sondages HDF et GOODS de tailles modestes sont affectés par la variance cosmique, bien qu'ils permettent d'observer des galaxies faibles à grand redshift. En revanche les sondages VIPERS et COSMOS explorent une large gamme d'environnements, mais sont limités à bas redshift (z<1.2, VIPERS) ou biaisés vers des galaxies lumineuses à grand z (COSMOS).

2

Activité de recherche et son évolution

2.1 Emission des galaxies : du pixel aux propriétés physiques

Au cours de ces vingt dernières années, les développements instrumentaux, en terme de sensibilité et de résolution, ont ouvert une nouvelle ère en révélant l'émission des galaxies individuelles sur l'ensemble du spectre électromagnétique. L'émission stellaire est désormais accessible de l'ultraviolet lointain (observé par le satellite GALEX) à l'infra-rouge proche et moyen (avec le satellite Spitzer/IRAC); l'absorption du rayonnement UV par les poussières du milieu interstellaire est ré-émise dans l'infra-rouge thermique (far-IR, avec Spitzer et Herschel); l'émission non thermique témoigne des phénomènes violents (supernovae [SN], noyaux actifs [AGN] en radio, mid-infrared and rayon-X). Combiné avec les mesures de redshifts obtenues soit avec les relevés spectroscopiques soit avec la technique des redshifts photométriques (section 2.1.2), les observations aux différentes longueurs d'onde ont permis des avancées majeures sur notre compréhension des processus d'évolution des galaxies en suivant deux approches distinctes :

— L'approche statistique, basée sur la mesure des fonctions de luminosité (LFs) à différentes époques : les LFs UV et Far-IR et leurs grandeurs intégrées (la densité de luminosité) ont permis de déterminer l'évolution globale de l'histoire de la formation d'étoile cosmique (SFRD) au cours des 10 derniers milliards d'années (section 2.2), indépendamment des histoires complexes subies par les galaxies individuellement. les LFs dans l'infra-rouge proche et moyen ont permis de mesurer les phases majeures d'assemblage de la masse stellaire au sein des galaxies (SMD) (section 2.3). Cette approche repose sur les hypothèses suivantes : les émissions UV et Far-IR (dans ce dernier cas, re-processée) sont dominées par les étoiles massives de courte durée de vie c'est à dire qu'elles sont un traceur de la formation d'étoile instantannée; les émissions Near-IR et Mid-IR sont dominées par les étoiles évoluées de masses intermédiaires qui constituent l'essentiel de la masse stellaire d'une galaxie; et finalement il est nécessaire de faire l'hypothèse d'une fonction de masse initiale universelle. Cette approche permet de répondre à deux questions fondamentales sur la formation des galaxies : existe-t-il une époque caractéristique pour la formation d'étoiles? quelle quantité de baryons, observée aujourd'hui, est déjà bloquée au sein des galaxies dans le passé?

— L'approche individuelle, basée sur l'analyse de la distribution d'énergie spectrale des galaxies : Les observations dans les différents domaines de longueurs d'onde ont conduit à l'émergence d'une multitude de populations spécifiques : classification en galaxies rouges et bleues, LBGs (Lyman Break galaxies), EROs (extremely red galaxies), UVLGs (UV luminuous galaxies), ULIRGs (ultra-luminuous Infra-red galaxies), SMGs (sub-millimeter galaxies), LAEs (Lyman-alpha emitters), ... Il est difficile de placer cet ensemble hétérogène dans un scénario unifié d'évolution des galaxies, en particulier à cause de la difficulté de suivre la transformation continue de ces diverses sous-populations au cours du temps et d'un cadre théorique, encore partiel, pour interpréter les observables. Toute classification à partir de grandeurs observées ne reflète pas nécessairement des différences fondamentales entre les galaxies en terme de propriétés physiques telles que l'activité stellaire (SFR), la masse stellaire ($M_\star$), le contenu en poussière. C'est cette continuité des propriétés physiques qui permet de mesurer et comprendre l'évolution des galaxies et de relier les différentes populations basées sur différents critères de sélection. Pour accéder à ces quantités physiques, il est nécesaire d'exploiter des modèles de populations stellaires synthétiques qui prédisent, pour une histoire de formation d'étoiles donnée, l'émission intégrée des populations stellaires à différents âges. Ces prédictions sont ensuite confrontées aux observations multi-longueurs d'onde (section 2.1.3).

Avant de pouvoir réaliser de telles analyses, plusieurs étapes sont nécessaires pour convertir les photons reçus aux diverses longueurs d'onde, avec des instruments aux caractéristiques très différentes, en une série de mesures exploitables pour reconstruire la distribution d'énergie spectrale de chaque galaxie, estimer son redshift et dériver les propriétés physiques. Ces trois étapes sont discutées ci-dessous.

2.1.1 Photométrie UV en champ encombré

La photométrie, à diverses longueurs d'onde, donne une vision très variable du ciel, du fait de l'origine différente de l'émission des sources ainsi qu'aux caractéristiques (résolution et sensibilté) des différents instruments. En particulier, le faible pouvoir résolvant des satellites UV ou Far-IR induit une confusion des sources qui impose de traiter ce problème avec des logiciels dédiés, comme dans le cas du satellite GALEX. Lancé en 2005, GALEX (Martin et al. 2005) a exploré le ciel ultraviolet dans les bandes Far-UV (135−175 nm) et Near-UV (170−275 nm) avec trois grands relevés d'imagerie : All-Sky (AIS), Medium (MIS) et Deep (DIS) Imaging Surveys. Sa faible résolution angulaire (PSF avec une FWHM~5arcsec) affecte l'extraction et la photométrie des sources UV dès la magnitude NUV~22-23 (Xu et al. 2005), alors que les champs profonds, avec un temps d'intégration typique de $T_{exp} \sim 30{,}0000$ sec/pixel, permettent de détecter des sources jusqu'à NUV~25-25.5 avec un signal sur bruit $S/N \sim 5$. Une approche alternative aux logiciels classiques d'extraction de sources, comme SExtractor (Bertin & Arnouts 1996), fût de développer une photométrie avec priors qui optimise l'attribution du flux UV à un ensemble de priors optiques via une modélisation/reconstruction de l'image GALEX. Dans ce cas spécifique, considérer que toutes les émissions UV ont une contrepartie optique est une approximation raisonnable, de part la pronfondeur des images optiques utilisées (cad les données du CFHTLS) et la forme typique des SEDs dans le pire des cas correspondant à un spectre plat ($f_\nu \propto$ ctt). Les priors optiques sont convolués à la résolution de GALEX, et l'attribution du flux de chaque prior est obtenue en maximisant la vraisemblance entre l'image simulée et l'image observée à l'aide d'un algorithme itératif d'Expectation-Maximisation (EM Guillaume et al. 2005). Le principe de la méthode (*EMphot*) décrit dans Llebaria et al. (2008); Conseil et al. (2011) a été validé à l'aide d'images simulées. L'impact sur les comptages GALEX dans les champs profonds est illustré dans la Figure 2.1. Cette nouvelle extraction réduit le nombre de sources brillantes (excés de flux lié à la fusion de sources) et détecte plus de sources faibles, jusqu'à la limite théorique attendue. Cette nouvelle photométrie permet d'exploiter de manière optimale les relevés profonds GALEX dans les champs COSMOS et VIPERS-MLS (Moutard et al. 2016b). Elle est mise à disposition de la communauté (http ://cesam.lam.fr/galex-emphot).

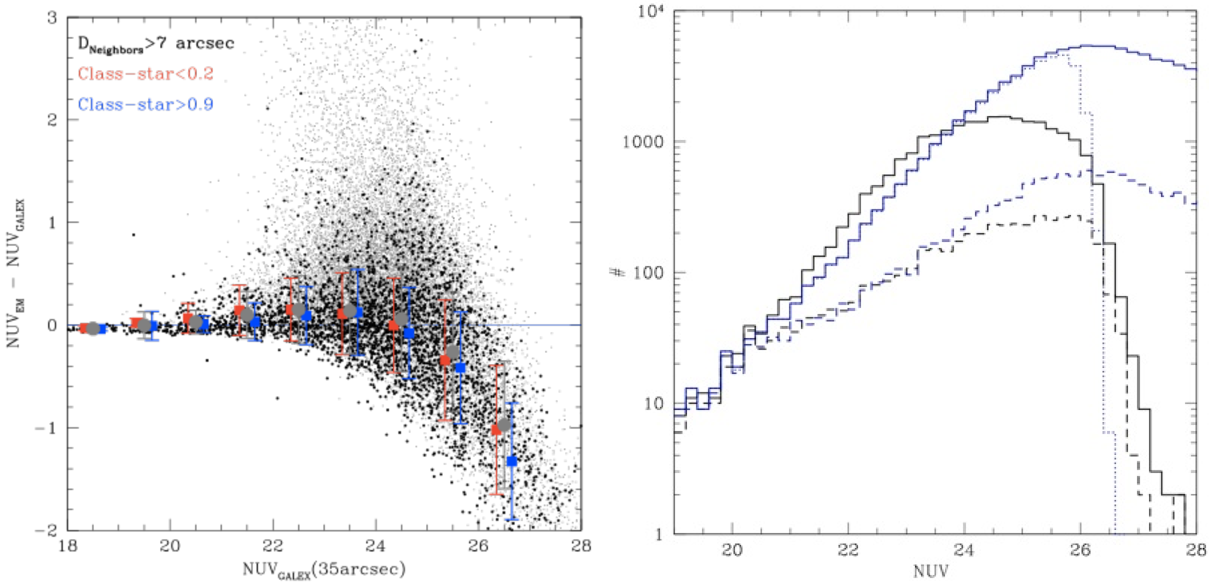


FIGURE 2.1: Gauche : Comparaison des flux UV mesurés avec le pipeline GALEX (basée sur le logiciel SExtractor, Bertin & Arnouts 1996) et le nouvel outil EMPhot, qui montre la sur-estimation systématique du flux des sources (points gris) avec la photométrie classique. Pour les sources isolées (points noirs) en revanche les deux méthodes sont comparables pour les sources compactes (en bleu) et étendues (en rouge, grâce à l'utilisation des vignettes optiques). Droite : Comptage de sources dans un champ profond GALEX avec le pipeline (lignes noires), la nouvelle photométrie avec priors (lignes bleues) et pour les sources isolées (lignes tirets). EMPhot tend à diminuer le flux des sources dû aux effets de proximité et à augmenter la détection de sources faibles grâce aux priors optiques. Le comptage des sources avec un bon signal sur bruit (S/N≥ 5) est montré avec la ligne pointillée (Moutard et al. 2016b).

2.1.2 Redshifts photométriques

Les redshifts photométriques sont devenus essentiels dans l'exploitation des sondages photométriques multi-couleurs. L'estimation des redshifts à partir de la photométrie seule est une vieille idée (Baum 1962; Puschell et al. 1982). Malgré le faible échantillonage de la distribution d'énergie spectrale des sources (relié à la largeur des filtres, $R \sim \lambda/\Delta\lambda$ ~5-10), cette méthode permet de contraindre la forme du continuum et d'identifier les structures dominantes (cassures de Lyman et à 4000Å) afin d'estimer le décalage spectral. Les redshifts de sources faibles au delà des capacités spectroscopiques des télescopes actuels et pour de grands échantillons peuvent ainsi être obtenus. Ces deux aspects ont contribué à rendre cette technique extrêment attractive pour les sondages dédiés à l'évolution des galaxies où certaines analyses statistiques (fonction de luminosité, clustering, ...) ne nécessitent pas forcément une grande précision sur le redshift. Elle est aussi devenue une composante essentielle des grands relevés cosmologiques pour contraindre l'énergie noire avec les analyses de lentilles gravitationnelles faibles et la fonction de masse des amas. Dans de telles analyses, l'incertitude des redshifts photométriques par rapport aux redshifts spectroscopiques peut-être en partie contrôlée avec l'utilisation de la fonction de densité de probabilité (PDF) attribuée à chaque source plutôt que la valeur discrète du redshift (voir applications dans Arnouts et al. 2002, 2007).

Deux types d'approches sont traditionnellement utilisés pour estimer les redshifts photométriques :

— méthode d'ajustement de spectres (*SED fitting*) : l'ensemble des codes (Arnouts et al. 1999; Benítez 2000; Bolzonella et al. 2000; Assef et al. 2010, ...) exploite une librairie de spectres prédéfinis (observés ou théoriques). Les flux prédits à travers les divers filtres, pour chaque spectre placé à différents redshifts, sont comparés aux couleurs observées et le meilleur ajustement est obtenu par minimisation du χ^2. Les PDFs sont extraites en marginalisant sur l'ensemble des paramètres (SED, atténuations, âges,...). Cette technique, motivée physiquement, ne nécessite pas un grand échantillon spectroscopique et quand une bonne librairie de spectres a été trouvée, elle peut être appliquée à d'autres sondages. Toutefois, elles sont souvent intensives en temps de calcul avec l'exploration de l'ensemble de la grille des flux prédits par les modèles. La pauvre connaissance de l'atténuation des spectres par la poussière et l'impact de l'inclinaison des galaxies sur les couleurs intégrées introduisent une dégénérescence dans l'espace redshift−couleurs difficile à corriger.

Le code *Le Phare*, que nous avons développé avec Olivier Ilbert (Arnouts et al. 1999; Ilbert et al. 2006), appartient à cette catégorie. La figure 2.1.2 (gauche) montre la précision obtenue dans les champs profonds Hubble Deep Field Nord et Sud. Plusieurs développements ont permis d'améliorer les performances du code (Ilbert et al. 2009) : l'ajustement des points-zeros des différentes bandes photométriques à l'aide d'un échantillon spectroscopique, l'inclusion des raies d'émission à l'aide de lois empiriques (e.g. Kennicutt 1998), l'utilisation de plusieurs lois d'atténuation (Arnouts et al. 2013). *Le Phare* a obtenu de bonnes performances dans les divers tests à l'aveugle réalisés sur plusieurs échantillons (e.g. Hildebrandt et al. 2010; Dahlen et al. 2013). Il est couramment utilisé par la communauté.

— méthode empirique (*Machine Learning*) : avec l'augmentation constante du nombre de redshifts spectroscopiques, il est devenu possible d'utiliser des algorithmes de Machine Learning supervisés qui optimisent la fonction permettant d'associer les propriétés photométriques (flux, couleurs) aux redshifts provenant d'un échantillon spectroscopique. Ces méthodes incluent les réseaux de neurones artificiels (Collister & Lahav 2004), les fits polynomiaux locaux (Csabai et al. 2007), les *Random Forests* (Carliles et al. 2010). Elles nécessitent de grands échantillons d'entrainement et ne peuvent être appliquées en dehors du domaine de redshifts et des conditions imposés par les échantillons spectroscopiques d'entrainement.

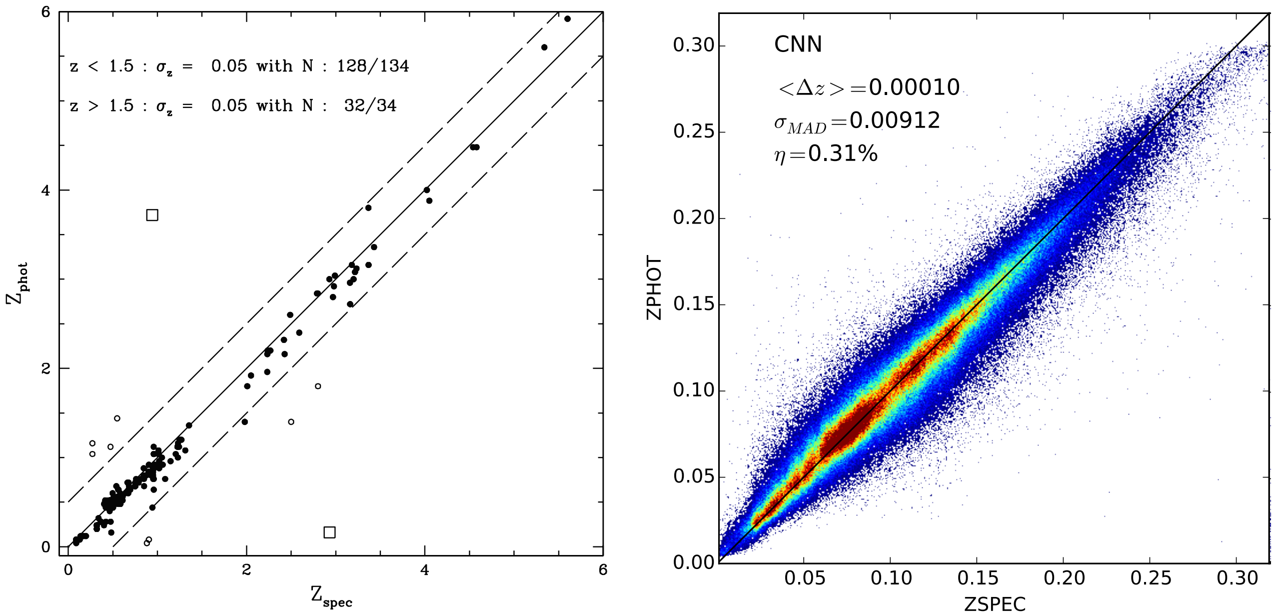


Figure 2.2: Gauche : Précision des redshifts photométriques avec le code *Le Phare* dans les champs Hubble Deep Fields, avec ces 150 redshifts spectroscopiques disponibles (Arnouts et al. 2002). Droite : Précision des redshifts photometriques dans le SDSS avec un algorithme de "Deep Learning" basé sur l'exploitation des vignettes multi-bandes ugriz (Pasquet et al. 2018)

La méthode de *SED fitting* a plusieurs avantages sur les techniques empiriques. Un échantillon spectroscopique sur toute la gamme de redshift n'est pas nécessaire et les propriétés physiques (par ex. le type de galaxies) sont naturellement estimées. Toutefois la méthode est sensible aux effets systématiques dans la photométrie alors qu'ils sont implicitement pris en compte dans les méthodes empiriques. Ces dernières méthodes commencent à donner de meilleurs résultats que les méthodes de *SED fitting*, mais l'absence de couverture spectroscopique dans certaines régions de l'espace des couleurs et à grand redshift reste la principale limitation.

Dans toutes les techniques ci-dessus, un facteur limitant est la dépendence aux paramètres d'entrée. L'extraction des couleurs est sensible aux choix de l'ouverture, aux variations de PSF, aux recouvrements de sources et ne capture pas toutes les informations présentes dans les images (brillance de surface, forme, granulosité, gradients de couleur, encombrement, ...). Ces dernieres années les réseaux de neurones convolutionels (CNNs) ont révolutionné le domaine de reconnaissance d'images montrant des performances sans précédent y compris en astronomie (Dieleman et al. 2015; Hoyle 2016). En prenant avantage des derniers développements dans le domaine du *Deep Learning*, il est possible de travailler directement au niveau des pixels avec les images multi-canaux sans dépendre de l'extraction de quantités photométriques. La figure 2.1.2 (droite) montre notre analyse préliminaire avec le sondage SDSS DR12 et ses ~500,000 galaxies avec des redshifts spectroscopiques et plus brillantes que $r \sim 17.8$ (Pasquet et al. 2018). Avec un entrainement du réseau exploitant ~80% des redshifts et un échantillon d'~ 20% pour évaluer les performances, nous montrons qu'une précision infèrieure à σ_z ~0.01 est obtenue. Aucun biais n'est observé avec le redshift, l'inclinaison (b/a), l'extinction galactique. Ces résultats sont notablement meilleurs que les redshifts photométriques fournis par le SDSS (Beck et al. 2016). Bien que préliminaire et que plusieurs problèmes doivent encore être résolus pour des échantillons à plus grand redshift, cette approche offre une alternative très prometteuse pour exploiter les grands relevés d'imagerie en cours (HSC [grizY]) et à venir (J-PASS [50-bands], LSST [ugrizY]) avec des séries de bandes photométriques bien définies.

2.1.3 Paramètres physiques

Le paramètre fondamental recherché pour chaque galaxie est son histoire de formation stellaire pour retracer les évènements majeurs dans son processus de formation/d'évolution. C'est l'objectif de l'archéologie galactique qui s'applique à la Voie Lactée et ses voisines. Pour les autres, où il n'est pas possible d'observer les différentes populations d'étoiles, une première approche est d'estimer la formation d'étoiles récente (SFR) et la masse stellaire (intégrale de l'activité stellaire passée). La luminosité UV (en l'absence d'extinction), la luminosité Far-IR (ou une combinaison des deux), la luminosité radio ou la raie nébulaire H_α sont toutes des traceurs du SFR avec des échelles de temps légèrement différentes (voir revues de Kennicutt 1998; Madau & Dickinson 2014). La sensibilité des mesures UV permet d'estimer la luminosité UV de galaxies faibles et à grand redshift mais elle est affectée par l'atténuation par les poussières dont la correction reste incertaine. La luminosité Far-IR est un indicateur indirect plus fiable mais les performances des instruments actuels limitent les détections aux objets les plus lumineux. Les luminosités Near-IR et Mid-IR quant à elles tracent plus directement la masse stellaire intégrée d'une galaxie mais le rapport $M_\star/L_{NIR}$ est sensible à la population stellaire sous-jacente (Bell et al. 2003) et une estimation précise de la masse stellaire doit prendre cet effet en compte.
Ces quinze dernières années, une méthode plus globale est apparue, qui consiste à utiliser les modèles de synthèse de populations stellaires. Ces modèles incluent notre connaissance des spectres stellaires d'une simple population d'étoiles et de leurs évolutions temporelles. Pour reproduire la diversité de spectres de galaxies, les spectres stellaires sont convolués avec plusieurs histoires de formation d'étoiles et différentes recettes pour le traitement de l'atténuation par les poussières. Le spectre émergent peut être directement comparé aux observations de l'UV au mid-IR, dominées par l'emission stellaire (voir revues de Walcher et al. 2011; Conroy 2013). En exploitant les données Galex, SDSS and Spitzer, Salim et al. (2005); Johnson et al. (2007) ont démontré que l'ajustement des modèles de populations stellaires aux données photométriques multi-longueurs d'onde donné des résultats en bon accord avec les traceurs

individuels (en particulier la raie d'émission H_α et la luminosité L_{IR}). Cette approche toutefois n'est pas exempte d'incertitude. La modélisation incertaine de certaines populations stellaires comme les TP-AGB (Thermally Pulsing Asymptotic Giant Branch, Maraston 2005) peut modifier le rapport $M\star/L_{NIR}$ et induire une sur-estimation de la masse stellaire, la domination d'une jeune population stellaire sur l'ensemble du spectre émergent peut masquer la vieille population stellaire sous-jacente induisant une sous-estimation de la masse stellaire (*outshining problem*, Pforr et al. 2012). Les dégénérescences entre l'atténuation, la métallicité et l'âge persistent même si elles peuvent être réduites avec un plus grand nombre de bande photométrique, une photométrie de qualité et une grande base de longueurs d'onde.
Cette approche adoptée dans le code *Le Phare* permet de dériver des propriétés physiques de manière homogène pour toutes les galaxies d'un échantillon et pour chaque galaxie, de dériver des paramètres internes auto-cohérents. Les modèles de synthèse de populations stellaires utilisés sont basés sur les librairies BC03 (Bruzual & Charlot 2003) et PEGASE (Fioc & Rocca-Volmerange 1997). Plusieurs histoires de formation stellaire sont adoptées (décroissance exponentielle ou retardée) avec trois métallicités différentes et plusieurs lois d'atténuation. L'analyse est effectuée dans un cadre bayésien : $P(M|D) \propto P(M) \times P(D|M)$, où $P(D|M)$ est la probabilité de reproduire les observations (D) pour un modèle donné (M), qui est ce qui est mesuré. En assumant des erreurs gaussiennes, la probabilité est $P(D|M) = e^{-\chi^2/2}$. $P(M|D)$ est la probabilité qu'un modèle s'ajuste aux données, qui est ce qui est recherché. $P(M)$ encode la connaissance à priori de la probabilité d'un modèle (*Le Phare* n'utilise pas de priors sur les modèles, $P(M) = 1$). La fonction de distribution de probabilité (PDF), pour la mesure des paramètres physiques, est obtenue en marginalisant sur l'ensemble des autres paramètres. La médiane de la distribution et son niveau de confiance sont utilisés pour dériver la valeur du paramètre et de son incertitude. Dans le cas de dégénérescences, la forme de la PDF apparaît non gaussienne et très déformée. La méthode est dite bayésienne car elle utilise une librairie prédéfinie de modèles incluant la connaissance à priori des histoires de formation des galaxies.
En exploitant les données multi-longueur d'ondes dans le champs COSMOS, Ilbert et al. (2010) ont montré comment les masses stellaires sont affectées par les différentes hypothèses sur les ingrédients des modèles et notamment le rôle important du choix des lois d'atténuation. Arnouts et al. (2013) ont montré qu'un choix de trois lois d'atténuation, qui représente des atténuations moyennes subies pour des galaxies avec des niveaux d'activités stellaires différents, donnaient une bonne estimation du taux de formation d'étoiles par rapport aux estimations de traceurs individuels $L_{UV} + L_{IR}$ (Fig A1, et A2 Arnouts et al. 2013).

2.2 Evolution du taux de formation d'étoiles cosmique et le rôle des poussières

L'activité stellaire dépend des processus physiques à l'oeuvre affectant le gaz dans le milieu interstellaire, de l'histoire des fusions successives, de l'apport en gaz et des processus de feedback. L'évolution de la densité du taux de formation d'étoiles cosmique (SFRD(z) exprimée en $M_\odot/yr/Mpc^3$) est donc une observable clé pour contraindre les processus dominants impliqués dans l'évolution des galaxies. La première estimation de l'évolution du SFRD a été obtenue par Lilly et al. (1996) avec le sondage spectroscopique du CFRS. Combiné avec de la photométrie optique, la mesure des fonctions de luminosité à 2800Å (rest-frame) a révélé une forte diminution, d'un facteur 10, de la densité de luminosité depuis $z \sim 1$, qu'ils interprétèrent comme le signe d'un déclin majeur de l'activité stellaire cosmique. Ces travaux ont été étendus à plus grand z ($z \geq 3$), à l'aide des observations des champs profonds HDFs en utilisant la densité de luminosité à 1500Å (Madau et al. 1998). Ces résultats combinés suggèrent une augmentation de l'activité stellaire cosmique en remontant le temps, avec un pic d'activité entre $1 \leq z \leq 2$ et une décroissance au delà de $z \sim 2$. Grâce aux observations du satellite GALEX (GALaxy Evolution eXplorer), Arnouts et al. (2005); Schiminovich et al. (2005) ont confirmé et consolidé cette évolution à bas redshift, $0 \leq z \leq 1.2$, avec un même et unique traceur (la luminosité à 1500Å) que les études à grand z. Arnouts et al. (2005) ont mesuré les LFs à 1500Å en utilisant un champ profond de GALEX dans la région du sondage spectroscopique VVDS, à partir d'un échantillon d'$\sim$ 1000 galaxies avec des redshifts spectroscopiques. A l'aide des redshifts photométriques dans les HDFs (Arnouts et al. 2002), l'analyse est étendue à plus grand redshift. La figure 2.3 (gauche) montre les LFs-UV dérivées entre $0 \leq z \leq 3.5$. Une forte évolution de la luminosité caractéristique est observée avec un raidissement de la pente vers les grands redshifts. En décomposant l'échantillon en trois classes spectrales : Sb−Sd, Sd−Irr, Starbursts, ils trouvent que les galaxies Starbursts évoluent peu en luminosité, mais leur densité augmente considérablement avec le redshift (de 15% dans l'univers local à ~50% à z~1). A partir des LFs observées, Schiminovich et al. (2005) ont dérivé l'évolution de la densité de luminosité à 1500Å. Ils confirment la forte évolution, estimée croitre en $(1 + z)^{2.5}$, jusqu'à $z \sim 1$ et un ralentissement, $(1 + z)^{0.5}$, à plus grand z. La conversion da la luminosité à 1500Å en taux de formation d'étoiles est sujette à une correction de l'atténuation par les poussières assez incertaine. En appliquant une atténuation, A_{FUV}, appropriée pour les galaxies de type starburst (Calzetti 1997; Meurer et al. 1999, voir ci-dessous), et une correction minimale typique observée dans l'univers local, une première estimation du SFRD(z) a été possible. Comme illustrée dans la figure 2.3 (droite), le facteur correctif apparait en bon accord avec les quelques mesures de SFRDs dérivées à partir d'échantillons sélectionnés avec la raie nébulaire H_α.

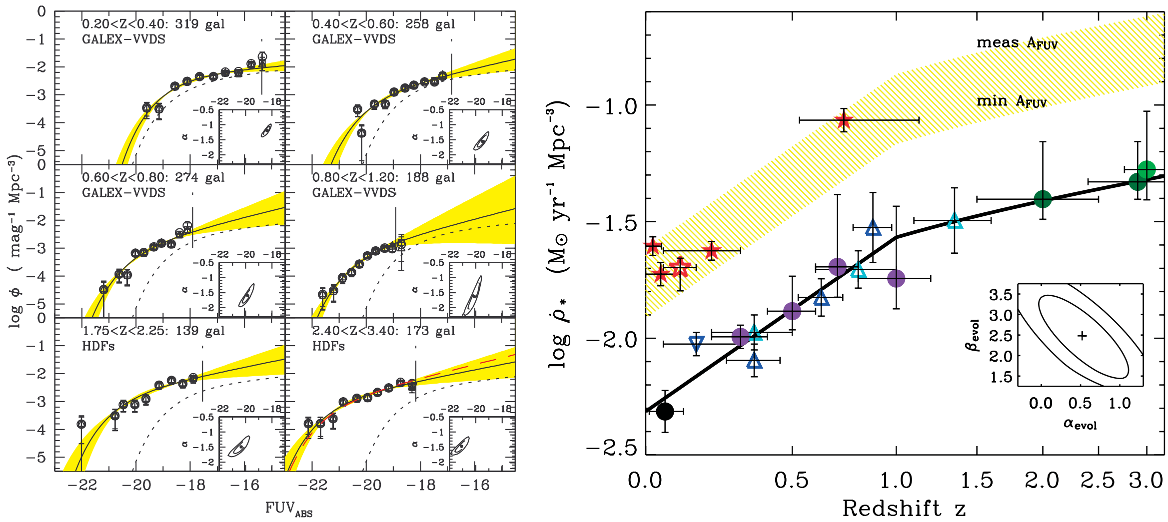


FIGURE 2.3: Gauche : Evolution de la fonction de luminosité FUV depuis $z \sim 3$ dérivée des sondages GALEX-VVDS et HDFs (Arnouts et al. 2005). Droite : Evolution de la densité cosmique du taux de formation d'étoiles à partir de l'émission UV avec (région jaune hachurée) et sans correction (ligne solide) de l'absorption par les poussières durant les dix derniers milliards d'années (Schiminovich et al. 2005).

Cette image globale de l'évolution du SFRD a depuis été consolidée sur un vaste intervalle de redshift $0 \leq z \leq 8$ en combinant l'ensemble des traceurs de l'UV au Far-IR et radio (voir revue de Madau & Dickinson 2014). L'intéret des mesures aux grandes longueurs d'onde est leur capacité à révéler la formation d'étoiles enfouie dans les régions poussiéreuses et donc avec une luminosité stellaire UV trés atténuée. Toutefois ces observations sont limitées par la sensibilité et résolution des instruments actuelles et ont recours à des techniques d'empilage (*stacking*) pour étendre la détection à des sources peu massives (Karim et al. 2011; Heinis et al. 2013). Idéalement, le bilan énergétique de l'activité stellaire peut être obtenu en sommant la formation d'étoiles non éteinte (UV) et celle re-processée en infrarouge (FIR ; $SFR \propto L_{UV} + L_{FIR}$). A défaut d'avoir accés à ces deux quantités simultanément dans les champs profonds, les mesures de SFR reposent essentiellement sur le SED fitting de l'émission stellaire avec des hypothèses sur les lois d'atténuation. La forme de ces lois d'atténuation encode des informations sur la nature des grains de poussière (taille, composition chimique) et la distribution spatiale de la poussière et des étoiles. Les travaux de Calzetti (1997) ont montré qu'il existait une relation étroite entre la pente du continu UV (β) et l'excés infra-rouge ($IRX = L_{IR}/L_{NUV}$) pour les galaxies starbursts et ont proposé une loi d'atténuation typique pour ces galaxies (Calzetti et al. 2000). En combinant les données Herschel ou Spitzer et GALEX, de nombreuses études ont suivi pour élargir ce travail à différents types de galaxies. En particulier Boquien et al. (2009) ont montré la nécéssité d'adopter une large gamme de lois d'atténuation pour reproduire la dispersion observée dans la relation IRX vs β.

Afin de s'affranchir du manque de connaissance sur les lois d'atténuation, Arnouts et al. (2013) ont développé une nouvelle méthode qui exploite l'évolution remarquable de l'excés IR projeté dans l'espace des couleurs "optiques" $(NUV - r)$ vs $(r - K)$, illustrée dans la figure 2.4 (gauche). Une dynamique d'un facteur 1000 est observée à travers cet espace couleur, avec une faible dispersion. Il est ainsi possible de prédire le rapport L_{IR}/L_{UV} de chaque galaxie à l'aide d'un simple vecteur (**NRK**), combinant les luminosités optiques, L_{NUV}, L_r, L_K, accessibles dans la majorité des sondages photométriques profonds. Les prédictions du L_{IR} sont illustrées dans la figure 2.4(droite). Une dispersion inférieure à 0.2dex est observée, meilleure que celle obtenue à l'aide des techniques de SED fitting, révélant l'intéret de cette méthode en l'absence d'informations à grande longueur d'onde. A l'aide d'une librairie de modèles de galaxies avec des prescriptions réalistes d'histoires de formation stellaire, il est possible de reproduire l'évolution de l'IRX observée dans le diagramme $NUVrK$. Pour cela il est nécessaire d'utiliser un modèle de poussières à deux composantes (les régions de formation d'étoiles et le milieu diffus (ISM)) et une distribution complète des inclinaisons des galaxies. En analysant la morphologie des galaxies résolues dans les images du sondage CFHTLS, Moutard et al. (2016b) ont montré que les galaxies avec un fort IRX dans le diagramme $NUVrK$ sont exclusivement des galaxies trés inclinées (vues par la tranche), leurs couleurs $(r - K)$ extrèmes ne pouvant s'expliquer qu'avec une large épaisseur optique (leur Figure 16).

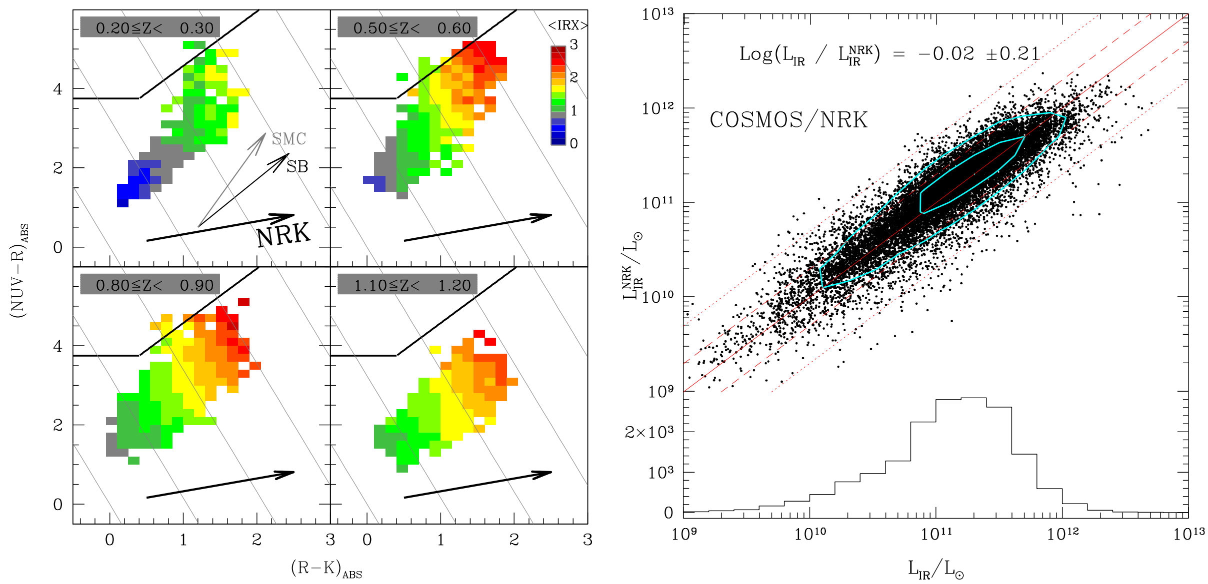


Figure 2.4: Gauche : Evolution de l'excés infrarouge ($IRX = L_{IR}/L_{UV}$) dans le diagramme $NUVrK$ ($(NUV - r)$ vs $(r - K)$) pour les galaxies actives dans le champ COSMOS. Droite : comparaison entre la luminosité infra-rouge mesurée et celle prédite avec le vecteur NrK. Figures extraites d'Arnouts et al. (2013).

La méthode **NRK** reproduit bien la pente et la normalisation de la relation entre SFR et masse stellaire observées dans la littérature (Arnouts et al. 2013) et prédit une évolution du SFR spécifique ($sSFR = SFR/M_\star$) en très bon accord avec les analyses basées sur les données UV +FIR dans les sondages

GOODS et COSMOS (Ilbert et al. 2015). Cette méthode est entrain d'être étendue à plus faibles masses ($M \geq 10^8 M_\odot$) et à plus grand redshift ($z \sim 3-4$), avec des techniques de *stacking* des données PACS et SPIRE d'Herschel (de 24μm à 500μm) afin d'élargir son domaine d'application.

2.3 L'époque d'assemblage des galaxies et les mécanismes de *quenching*

Les premières analyses des champs profonds dans l'infra-rouge proche ont mesuré une diminution de la densité moyenne de masse stellaire au cours du temps, consistante avec une formation hiérarchique des structures. La moitié de la densité de masse stellaire s'est formée à grand z ($z \geq 1$). Les analyses du taux de formation d'étoiles spécifique ($SFR/M_\star$) des galaxies individuelles montrent que les galaxies les plus massives ont formé l'essentiel de leur étoiles plus tôt que les galaxies moins massives. Ce phénomène, appelé *downsizing* (Cowie et al. 1996; Juneau et al. 2005) de la formation d'étoiles avec la masse, peut s'interpréter dans le cadre d'une formation biaisée des galaxies. Le regroupement des galaxies montre que les galaxies à grand z, les plus lumineuses et avec d'intense formation d'étoiles, se forment préférentiellement au sein des halos de matière noire les plus massifs (Adelberger et al. 1998; Arnouts et al. 1999). A plus bas redshift en revanche, l'essentiel de l'activité stellaire a migré vers des halos de plus petite masse, généralisant la notion de *downsizing* aux halos de matière noire (d'apres les analyses de *clustering* avec GALEX et le CFHTLS, Heinis et al. 2007).

En séparant la densité de masse stellaire accumulée dans les galaxies passives et actives, Arnouts et al. (2007) ont mis en évidence que l'époque majeure d'assemblage des galaxies passives se situe entre $z \sim 2$ et $z \sim 1$, avec une augmentation d'un facteur 10 de la densité de masse stellaire, et seulement un facteur 2 entre $z \sim 1$ et $z \sim 0$. La densité de masse stellaire des galaxies actives reste, quant à elle, constante durant les 8 derniers milliards d'années, suggérant qu'une fraction des galaxies actives quitte la séquence $SFR - M_\star$ et devient passive, comme illustré dans la figure 2.5 (gauche). Comprendre l'origine de cette suppression de l'activité stellaire est un enjeux majeur. De nombreux mécanismes ont été mis en avant, qui peuvent être rapides ou lents, internes ou environnementaux, ou une combinaison des deux. • *Morphological quenching* : La présence d'un bulbe galactique stabilise le gaz du disque contre toute fragmentation, nécessaire à la formation d'étoiles (Martig et al. 2009). ce mécanisme est de plusieurs Gyrs. • *Halo quenching* : Dans les halos plus massifs que $10^{12} M_\odot$, le gaz accrété subit un chauffage gravitationnel l'empéchant de refroidir pour alimenter la formation stellaire (Dekel & Birnboim 2006). • *Mass quenching* (Peng et al. 2010) : Terme générique utilisé pour décrire des processus liés à la masse des galaxies. Il inclut les effets rétroactifs liés aux vents stellaires, aux supernovae et aux noyaux actifs, ainsi que le *halo quenching*. • *Environmental quenching* (Peng et al. 2010) : Terme générique qui décrit les processus dans les environnements denses : *Strangulation* quand l'apport en gaz cesse et les galaxies consomment le gaz restant sur des échelles de temps de quelques Gyrs ; *Ram Pressure stripping* lors de la chute de galaxies dans le milieu chaud intra-amas, provoquant la perte de leur réservoir de gaz, processus qui peut être rapide ; *Merging*, processus de fusion, entrainant un pic d'activité, accompagné d'une perte de gaz lié à la formation d'un noyau actif, processus violent et court, accompagné d'une transformation morphologique.

Grâce au sondage VIPERS et sa photométrie multi-longueurs d'onde (Moutard et al. 2016b), Moutard et al. (2016a) ont mesuré les fonctions de masse des populations passives et actives. La taille du sondage a permis de contraindre leurs évolutions, sans être affecté par les effets de variance cosmique, qui dominent les analyses précédentes. La fonction de masse des galaxies actives montre une masse caractéristique constante, à $M_\star \sim 10^{10.65} M_\odot$ entre $0 \leq z \leq 1.5$, confirmant l'existence d'une masse stellaire au delà de laquelle la formation stellaire cesse. Au delà de cette masse, les galaxies migrent vers la population de galaxies passives. Les galaxies plus massives que $10^{11.5} M_\odot$, majoritairement des galaxies passives, subissent une évolution en densité d'un facteur 2 entre z$\sim$ 1 et $z \sim 0$, suggérant que l'assemblage de cette population continue, via un processus de fusion de galaxies passives (*dry mergers*). A l'autre extrémité de la fonction de masse, une population de galaxies passives de faible masse ($M_\star \leq 10^{10} M_\odot$) apparait à bas redshift (voir aussi Ilbert et al. 2013). Les études environnementales suggèrent que cette population est dominée par des galaxies satellites et donc sujettes à un *quenching* environnemental. En exploitant le diagramme $NUVrK$, il est possible de distinguer les deux principaux canaux suivis par les galaxies en cours de *quenching* et d'estimer les temps caractéristiques à l'aide de modèles de populations stellaires comme illustrés dans la figure 2.5 (droite). Les galaxies massives subissent préférentiellement un quenching lent entre 1 et 3.5 Gyrs, compatible avec des processus de strangulation, où l'apport en gaz est progressivement stopé (Peng et al. 2015). Le quenching morphologique, identifié par Haines et al. (2017), est aussi compatible avec ces échelles de temps. Lorsque les galaxies actives atteignent une densité de surface de masse stellaire critique, généralement accompagnée de la prédominance d'un bulbe galactique, elles migrent vers les galaxies passives. Bien que ces analyses ne permettent pas d'identifier un phénomène en particulier, elles pointent préférentiellement vers des processus séculaires lents et progressifs. En revanche, le canal suivi par les galaxies satellites de faibles masses, montre qu'elles subissent un *quenching* plus rapide, inférieur à 1Gyr, qui est compatible avec des mécanismes de *ram-pressure stripping*, ou de fusion, au sein des amas, illustrant deux facettes des processus de *quenching* subis par les galaxies.

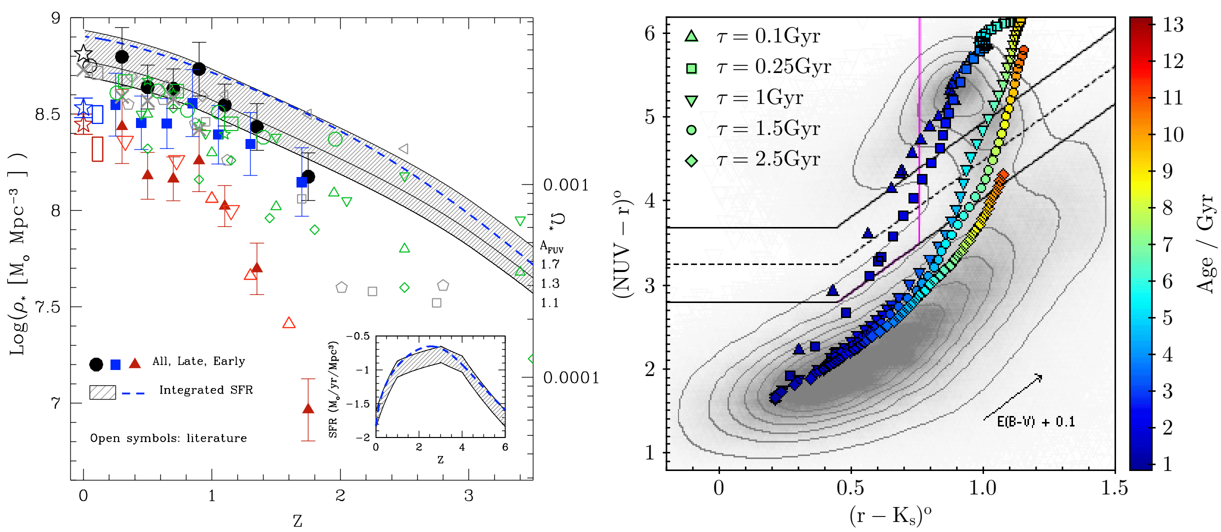


Figure 2.5: Gauche : évolution de la densité cosmique de masse stellaire à $z \leq 2$ (globale : cercles noirs, SF : carrés bleus, Passives : triangles rouges) et comparaison avec l'intégration de la densité de formation d'étoiles cosmique (région grisée) d'après Arnouts et al. (2007). Droite : Tracés évolutifs illustrant les deux principaux canaux suivis par les galaxies en cours de quenching (Moutard et al. 2016a).

2.4 Influence de la toile cosmique sur l'évolution des galaxies

La toile cosmique constitue certainement l'une des caractéristiques les plus remarquables de l'univers. Ces structures émergent de la croissance anisotrope des fluctuations de densité primordiales, sous l'effet de la gravité. Au cours de ce processus, la matière s'écarte des régions sous-denses, s'écoule le long des feuillets puis des filaments et enfin vers les noeuds correspondant aux pics de densité, donnant naissance à la toile cosmique (ou *cosmic web*, CW, Bond et al. 1996). Le gaz suit les gradients de potentiel gravitationnel imposés par la matière noire, formant des filaments d'hydrogène ionisé, le milieu inter-galactique (IGM), dans lequel se forment les galaxies. Au delà des effets bien connus de la densité, dans quelle mesure les effets de marée, induits par l'anisotropie à grande échelle, affecte les propriétés des halos et des galaxies ? Au premier ordre les variations de densité à grande échelle, générées par les grandes structures, modifient la fonction de masse des halos avec plus de halos passant le seuil d'effondrement plus tôt (Kaiser 1984). Ce biais de formation explique pourquoi les galaxies massives, vieilles se trouvent préférentiellement dans les régions denses. Au delà de cet effet, à quel point l'anisotropie du CW et les forces de marée induites influencent l'évolution des galaxies, reste une question ouverte. A densité fixée, y-a-t'il des effets autres que la masse des halos à prendre en compte ? Les récentes simulations hydrodynamiques prédisent l'existence de courants de gaz froid qui pénètrent au coeur des halos pour alimenter les galaxies en gaz et soutenir la formation d'étoiles à grand z. Ces filaments jouent aussi un rôle dans le transfert de moments angulaires aux disques des premières galaxies avec des spins préférentiellement alignés avec la structure filamentaire voisine. Les générations suivantes se forment ensuite par coalescence de galaxies parcourant les filaments en direction des noeuds. Leur mouvement orbital relatif est converti en spin lors de leur fusion, induisant une ré-orientation du spin, perpendiculaire aux filaments. Cette séquence évolutive contribue ainsi à l'émergence de la séquence de Hubble observée aujourd'hui. Ces prédictions sur l'orientation du spin des halos et des galaxies (Codis et al. 2012; Dubois et al. 2014) ont récemment été confirmées observationnellement dans le sondage du SDSS (Tempel et al. 2013). Ces résultats majeurs confirment le rôle joué par l'environnement dynamique à grande échelle dans l'évolution des galaxies, mais négligé dans les modèles de formation des halos basés sur la masse du halo et la densité (Mo et al. 1998). Ce qui reste à démontrer est si les propriétés intégrées telles que la morphologie, la masse stellaire, le SFR et sSFR, dépendent de leur localisation dans le CW, si ces effets ne sont pas effacés par les mécanismes internes tels que les effets rétroactifs. A l'aide des grands sondages spectroscopiques, il est possible désormais d'adresser de telles questions.

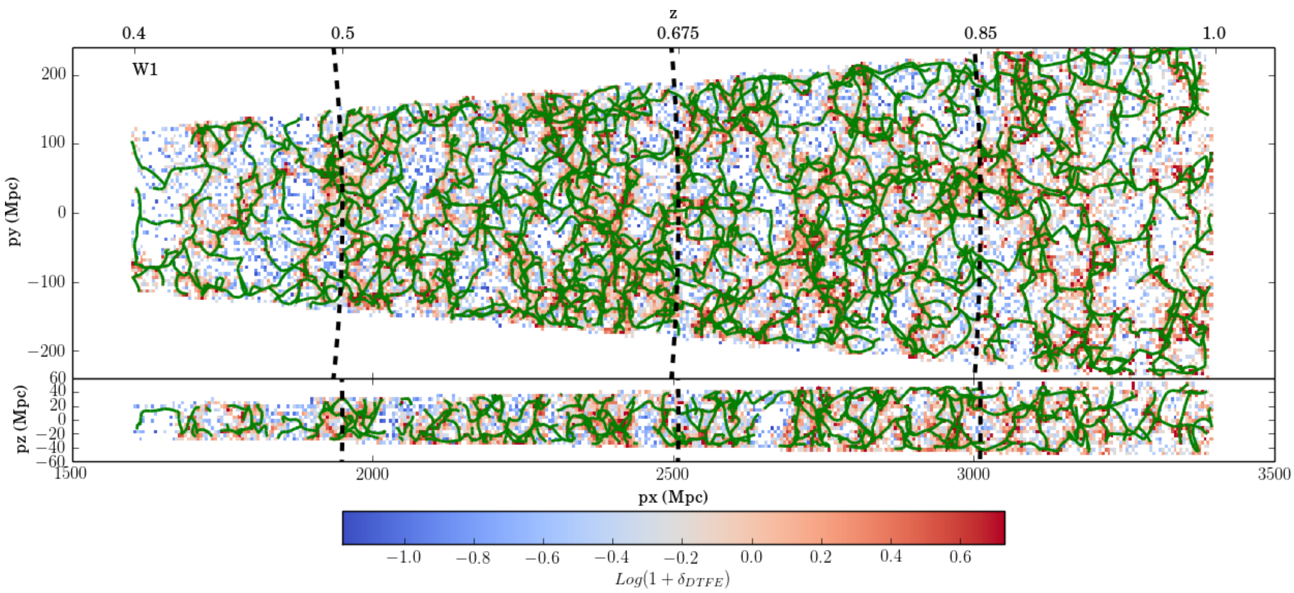


FIGURE 2.6: Recontruction des filaments (lignes vertes) dans le champ VIPERS (W1) entre $0.4 \le z \le 1$, surimposé sur le champ de densité local, codé en couleur (Malavasi et al. 2017).

Malavasi et al. (2017) ont utilisé le sondage VIPERS à grand redshift ($0.4 \le z \le 1$), couvrant 24 deg^2 avec 90,000 galaxies à $i \le 22.5$)). Ils ont pu reconstruire la structure filamentaire de la toile cosmique avec le code DisPerSE (Sousbie 2011), qui traite directement des distributions discrètes. DisPerSE extrait les points critiques du champ de densité reconstruit avec la tesselation de Delaunay et appareille les points critiques via des lignes de champs tangents aux gradients du champ de densité. Cette segmentation géométrique permet d'identifier les murs, les filaments et les noeuds et de filtrer les structures les plus significatives. La figure 2.6 montre comment la structure filamentaire dans le sondage VIPERS suit finement les crêtes du champ de densité. En mesurant la distance de chaque galaxie à son filament le plus proche, Malavasi et al. (2017) ont mis en évidence un effet de ségrégation. Les galaxies les plus massives (passives) sont plus proches du coeur des filaments que les moins massives (actives). Un même effet de masse est observé pour la population de galaxies actives seules (Figure 2.7). Ce résultat suggère que les galaxies moins massives/actives restent en périphérie des filaments, une région riche en vorticité, où l'accrétion s'effectue de manière continue. En revanche, les galaxies massives/passives finissent l'assemblage de leur masse stellaire via des fusions successives, le long des filaments, lors de leur migration vers les noeuds, en accord avec les simulations. Ils n'ont toutefois pas pu séparer les effets de densité, des effets liés aux forces de marée anisotropes.

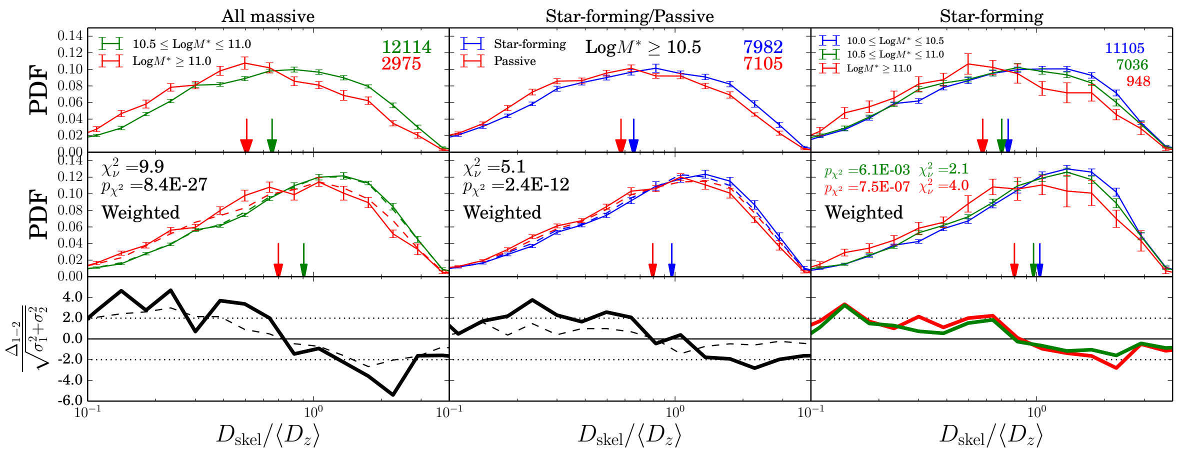


Figure 2.7: Distributions differentielles des distances aux filaments de toutes les galaxies (à gauche) et les galaxies actives (à droite) dans plusieurs bins de masse stellaire et pour les galaxies passives ou actives plus massives que $M_\star = 10^{10.5} M_\odot$ (au centre, Malavasi et al. 2017).

Kraljic et al. (2018) ont étendu l'analyse précédente à l'ensemble des caractéristiques du CW (noeud, filament, mur) à l'aide du sondage GAMA, dans l'univers proche, qui possède un échantillonage spectroscopique plus dense (120 deg^2 avec 150,000 galaxies à $r \leq 19.8$; $z \leq 0.3$). Ils confirment les ségrégations observées avec les filaments de l'étude précédente. D'autre part, ils montrent que la fraction de galaxies rouges augmente en s'approchant des filaments et des noeuds (Figure 2.8, gauche). Cette augmentation démarre à une distance de plusieurs Méga-parsecs des noeuds du CW, suggérant qu'une transformation est déjà en cours dans les filaments, bien avant d'atteindre le rayon de viriel des amas. Pour les galaxies actives, leur activité stellaire montrent deux régimes (Figure 2.8, droite). A distance intermédiaire, un état stable domine, qui peut refléter le bon équilibre entre la conversion de gaz en étoiles et l'apport en gaz, controlé par les filaments environnants. Lorsque la galaxie s'approche du coeur du filament, l'activité stellaire cesse. Ce phénomène peut être associé à une déconnection des écoulements filamentaires dans le coeur plus turbulents, empéchant tout remplissage du réservoir de gaz et conduire à un *quenching* par strangulation (Aragon-Calvo et al. 2016; Peng et al. 2015). En utilisant un estimateur de densité mesuré sur des échelles plus grandes (pour intégrer les effets d'accrétion passés), ils montrent qu'une partie du signal est dûe aux effets de marée induits par l'anisotropie du CW.

L'ensemble de ces résultats montrent l'intérêt de considérer la géométrie de la toile cosmique comme une nouvelle métrique pour interpréter l'évolution des galaxies.

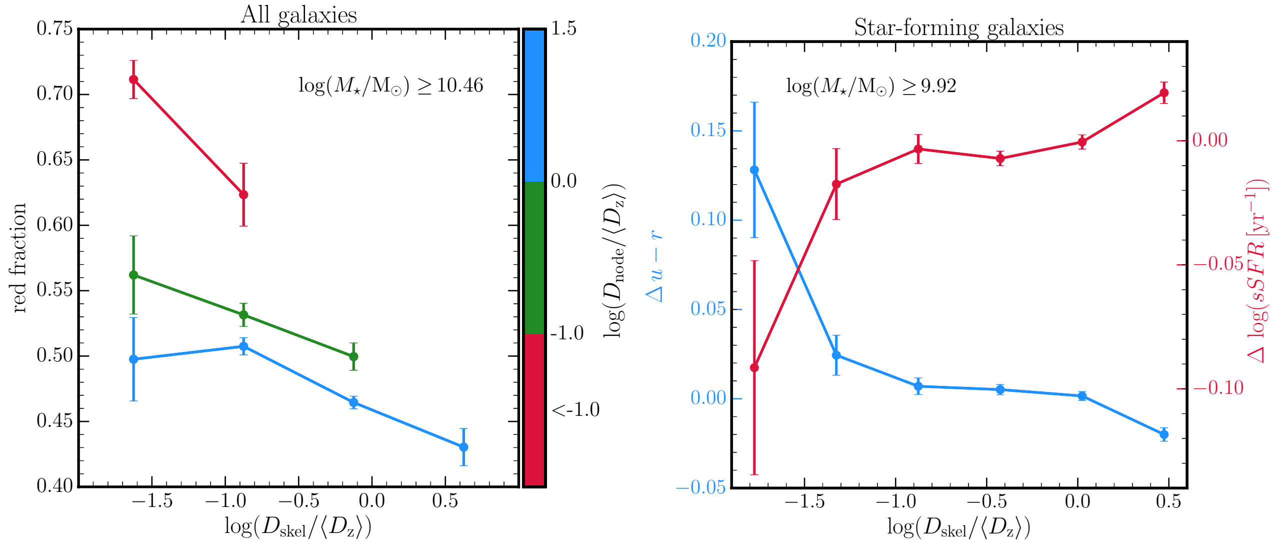


FIGURE 2.8: Evolution des propriétés des galaxies en fonction de leurs positions dans la toile cosmique. Gauche : fraction de galaxies rouges en fonction de la distance aux filaments et aux noeuds. Droite : Evolution de l'excès de couleur intrinsèque $(u - r)$ ou de SFR spécifique pour les galaxies actives en fonction de la distance aux filaments (Kraljic et al. 2018)

3

Perspectives

Au cours de ces vingt dernières années, une vision cohérente des grandes étapes de la formation des galaxies a émergé. Toutefois malgré le succés du modèle de formation hiérarchique, les processus impliqués dans la physique des baryons restent encore incertains. Le pic d'activité stellaire des galaxies de faible masse se produit plus tard que prédit par le modèle standard. L'origine de l'arrêt de la formation stellaire des galaxies massives au delà d'un certaine masse critique est il exclusivement lié au déclenchement de noyaux actifs ? L'efficacité de la conversion gaz − étoiles dans les disques est-elle modifiée lorsque les galaxies sortent de la séquence principale. Cette relation, $M_\star - SFR$, a-t-elle un sens physique ou devrions-nous ne considérer que la masse stellaire impliquée dans le disque ($M_{disque} - SFR$) et le mettre en perspective avec le rôle des bulbes galactiques ainsi que toute transformation morphologique ? Les simulations hydrodynamiques montrent l'importance des courants froids, à grand redshift, mais leurs détections restent encore un challenge observationnel. Quels sont les effets induits par les noyaux actifs sur l'environnement, à proximité des galaxies (le milieu circum-galactique), et sur le milieu inter-galactique (IGM) ? Comment la relation entre l'environnement et les propriétés des galaxies, bien établie dans l'univers local, a-t-elle évoluée au cours du temps ? Au delà du rôle joué par la masse stellaire et la densité locale sur les propriétés des galaxies, les effets de marée induits par les grandes structures ont été mis en évidence à bas z. Qu'en est-il à grand redshift, où les simulations prédisent un lien encore plus étroits ? Les futurs instruments et les grands relevés à venir dans les prochaines décénnies fourniront des éléments de réponses à ces questions.

• Les futurs sondages spectroscopiques comme PFS (Prime Focus Spectrograph), WFIRST et Euclid atteindront une densité de sources spectroscopiques capable de fournir les premières reconstructions de la toile cosmique sur plusieurs dizaines de degrés carrés, à grand redshift, $1 < z < 2$, lorsque l'activité stellaire était encore vigoureuse. Dans cette période, l'univers subit une transition avec le ralentissement de l'activité stellaire et l'apparition des galaxies passives. Il est encore débattu si le rôle de la toile cosmique impacte les propriétés physiques (SFR, sSFR, morphologie, métallicité) des galaxies. Avec son volume équivalent au SDSS à grand z, le sondage PFS sera le seul sondage capable d'explorer le rôle de l'environnement à grand z. Les effets rétro-actifs et les processus d'accrétion sont les paramètres fondamentaux régulant l'évolution des galaxies. Grâce à son domaine spectral, il sera possible de mesurer la métallicité et donc accéder à la fraction de gaz des galaxies et de suivre son évolution avec le redshift et la toile cosmique simultanément pour la première fois. L'échantillonnage élevé permettra d'analyser les populations des amas et proto-amas jusqu'à z∼ 2-3. La haute résolution spectrale permettra d'utiliser les raies d'absorption interstellaires (MgII, SIV, CIV) et mettre en évidence les *outflows* et peut-être les *infalls* très difficiles à détecter mais qui sont la clé de voute du modèle actuel. Le milieu circum-galactique (CGM) est l'interface entre la galaxie et le milieu inter-galactique (IGM). Il contient donc la trace des effets rétro-actifs passés et de l'apport en gaz pour la formation stellaire future. Avec des techniques d'empilage de spectres d'arrière plan, il sera possible d'explorer ces processus en fonction de la distance d'impact aux galaxies d'avant plan et apporter des réponses sur l'importance de ces processus en fonction du redshift, de la masse et de l'environnement. A plus grand redshift, $z \geq 2$, les absorptions de la forêt Lyman-α dans le spectre des galaxies, combinées aux techniques de tomographie, permettront de reconstruire la distribution tri-dimensionnelle du gaz du milieu intergalactique. Il sera alors possible d'analyser directement le lien entre l'environnement du milieu inter-galactique, principal réservoir de gaz, et les galaxies lors du pic de l'activité stellaire cosmique.

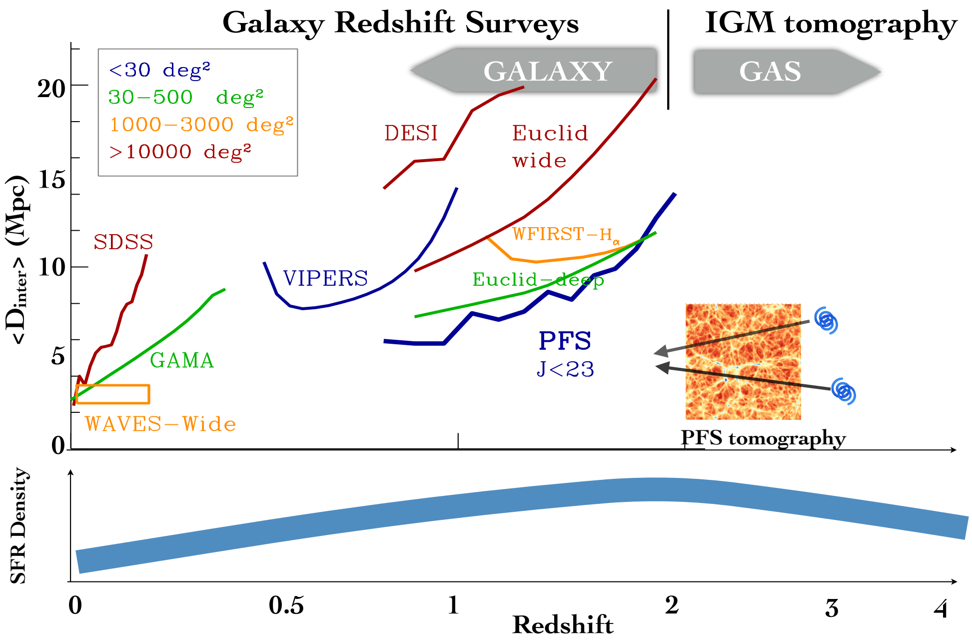


FIGURE 3.1: Mesure des performances, à l'aide de la distance inter-galaxie, des sondages spectroscopiques actuels et futurs pour cartographier la toile cosmique. A grand redshift, l'analyse de la forêt Lyman-α permettra de reconstruire la structuration du milieu intergalactique (IGM).

• Les grands relevés d'imagerie comme HSC (Hyper Suprime Caméra) et LSST offriront une nouvelle approche pour étudier l'évolution des galaxies, grâce à leur combinaison unique entre profondeur $r \sim 27$ et surface. La photométrie couvre le domaine optique (UgrizY) et sera étendue dans l'infrarouge proche (HSC+VISTA, LSST+Euclid). Grâce à l'homogénéité de la photométrie et l'accés à des échantillons spectroscopiques de plus en plus importants, il deviendra possible d'appliquer les techniques de *Deep Learning*, illustrées avec le SDSS, aux données HSC dans un premier temps et préparer la voie pour le LSST. Ces sondages permettront d'assembler de grands échantillons sur une large rangée de propriétés physiques, d'environnements et de redshifts $0 \leq z \leq 7$. Les mesures statistiques, telles que les fonctions de luminosité, les fonctions de corrélation, les analyses de distorsions gravitationnelles faibles, seront faiblement affectées par la variance cosmique et permettront de décomposer les échantillons par masse stellaire, SFR/sSFR afin d'explorer le lien avec leurs halos de matière noire dans différents environnements. Le suivi de leurs évolutions temporelles donnera un éclairage sur les mécanismes de *quenching*. La détection des proto-amas à $z \geq 2$, permettra d'analyser quand prend place les processus de *quenching* des satellites dans ces environnements extrêmes. Les mesures de la masse totale de ces structures cosmiques devient difficile à de tel redshift, car la technique de distorsion gravitationnelle est inéfficace au delà de $z \sim 1$. Grâce à la densité projetée des galaxies à $z \geq 3$, des techniques alternatives comme le biais de magnification pourront être utilisées dans cette tâche. L'analyse de la quantité de flux sortant des galaxies en dessous de la discontinuité de Lyman (λ <912Å), pour les galaxies actives à $z \sim 3-4$, servira de référence pour étudier la contribution à la réionisation des galaxies à $z \geq 6$. Durant l'époque du déclin de la formation stellaire cosmique ($0 \leq z \leq 2-3$), la profondeur unique des données UV (au repos) permettra de suivre l'évolution de la pente des fonctions de luminosité, indicative des phénomènes de rétro-actions dans les halos peu massifs. Le déclin du taux de formation d'étoiles depuis z~1.5 − 2 sera analysé par environnement. A l'aide de redshifts photométriques précis, Laigle et al. (2018) ont montré qu'il était possible de reconstruire de manière fiable la structure filamentaire de la toile cosmique projetée dans de fines tranches de redshift ($\Delta z \leq 0.1$). Une telle reconstruction offre la possibilité d'élargir les analyses actuelles, basées sur les sondages spectroscopiques, à des échantillons statistiques beaucoup plus importants et d'étendre les résultats dans un régime de masse encore inexploré.

• Les récentes analyses de galaxies résolues, dans l'univers locales, apportent des informations sur l'efficacité de la formation stellaire dans les disques et les mécanismes de *quenching* des galaxies. Les modes IFU de JWST et de l'ELT pemettront de telles analyses à grand redshift. Il sera possible de tester si ces mécanismes affectent l'efficacité de la formation d'étoiles sur l'ensemble du disque, dû à un processus de starvation progressif, par exemple, et/ou si ils suivent une évolution radiale *inside-out*/*out-inside* ainsi que le rôle spécifique de la croissance des bulbes galactiques pouvant engendrer un *quenching* morphologique et la dépendance de ces mécanismes avec l'environnement. Ils permettront aussi la recherche de filaments à proximité des galaxies pour détecter comment s'effectue l'alimentation en gaz dans le milieu circum-galactique.

Liste des figures

Bibliographie

Adelberger, K. L., Steidel, C. C., Giavalisco, M., et al. 1998, ApJ, 505, 18

Aragon-Calvo, M. A., Neyrinck, M. C., & Silk, J. 2016, ArXiv 1607.07881

Arnouts, S., Cristiani, S., Moscardini, L., et al. 1999, MNRAS, 310, 540

Arnouts, S., Le Floc'h, E., Chevallard, J., et al. 2013, A&A, 558, A67

Arnouts, S., Moscardini, L., Vanzella, E., et al. 2002, MNRAS, 329, 355

Arnouts, S., Schiminovich, D., Ilbert, O., et al. 2005, The Astrophysical Journal, 619, L43

Arnouts, S., Walcher, C. J., Le Fèvre, O., et al. 2007, A&A, 476, 137

Assef, R. J., Kochanek, C. S., Brodwin, M., et al. 2010, ApJ, 713, 970

Baum, W. A. 1962, in IAU Symposium, Vol. 15, Problems of Extra-Galactic Research, ed. G. C. McVittie, 390

Beck, R., Dobos, L., Budavári, T., Szalay, A. S., & Csabai, I. 2016, MNRAS, 460, 1371

Bell, E. F., McIntosh, D. H., Katz, N., & Weinberg, M. D. 2003, ApJS, 149, 289

Benítez, N. 2000, ApJ, 536, 571

Bertin, E. & Arnouts, S. 1996, A&AS, 117, 393

Bolzonella, M., Miralles, J.-M., & Pelló, R. 2000, A&A, 363, 476

Bond, J. R., Cole, S., Efstathiou, G., & Kaiser, N. 1991, ApJ, 379, 440

Bond, J. R., Kofman, L., & Pogosyan, D. 1996, Nature, 380, 603

Boquien, M., Calzetti, D., Kennicutt, R., et al. 2009, ApJ, 706, 553

Bouché, N., Dekel, A., Genzel, R., et al. 2010, ApJ, 718, 1001

Bruzual, G. & Charlot, S. 2003, MNRAS, 344, 1000

Calzetti, D. 1997, AJ, 113, 162

Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682

Carliles, S., Budavári, T., Heinis, S., Priebe, C., & Szalay, A. S. 2010, ApJ, 712, 511

Chen, Y.-C., Ho, S., Mandelbaum, R., et al. 2017, MNRAS, 466, 1880

Codis, S., Pichon, C., Devriendt, J., et al. 2012, MNRAS, 427, 3320

Colless, M., Peterson, B. A., Jackson, C., et al. 2003, ArXiv Astrophysics e-prints

Collister, A. A. & Lahav, O. 2004, PASP, 116, 345

Conroy, C. 2013, ARA&A, 51, 393

Conseil, S., Vibert, D., Arnouts, S., et al. 2011, in ASPCS, Vol. 442, ADAS, 107

Cowie, L. L., Songaila, A., Hu, E. M., & Cohen, J. G. 1996, AJ, 112, 839

Csabai, I., Dobos, L., Trencséni, M., et al. 2007, Astronomische Nachrichten, 328, 852

Daddi, E., Bournaud, F., Walter, F., et al. 2010, ApJ, 713, 686

Dahlen, T., Mobasher, B., Faber, S. M., et al. 2013, ApJ, 775, 93

Dekel, A. & Birnboim, Y. 2006, MNRAS, 368, 2

Dieleman, S., Willett, K. W., & Dambre, J. 2015, MNRAS, 450, 1441

Dubois, Y., Devriendt, J., Slyz, A., & Teyssier, R. 2012, MNRAS, 420, 2662

Dubois, Y., Pichon, C., Welker, C., et al. 2014, MNRAS, 444, 1453

Eisenstein, D. J., Zehavi, I., Hogg, D. W., et al. 2005, ApJ, 633, 560

Fioc, M. & Rocca-Volmerange, B. 1997, A&A, 326, 950

Frenk, C. S. & White, S. D. M. 2012, Annalen der Physik, 524, 507

Garilli, B., Guzzo, L., Scodeggio, M., et al. 2014, A&A, 562, A23

Guillaume, M., Llebaria, A., Aymeric, D., Arnouts, S., & Milliard, B. 2005, in SPIE, Vol. 6064, Conference Series, 11–1–11–10

Hahn, O., Porciani, C., Dekel, A., & Carollo, C. M. 2009, MNRAS, 398, 1742

Haines, C. P., Iovino, A., Krywult, J., et al. 2017, A&A, 605, A4

Heinis, S., Buat, V., Béthermin, M., et al. 2013, MNRAS, 429, 1113

Heinis, S., Milliard, B., Arnouts, S., et al. 2007, ApJS, 173, 503

Hildebrandt, H., Arnouts, S., Capak, P., et al. 2010, A&A, 523, A31

Hoyle, B. 2016, Astronomy and Computing, 16, 34

Hubble, E. P. 1926, ApJ, 64

Ilbert, O., Arnouts, S., Le Floc'h, E., et al. 2015, A&A, 579, A2

Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, A&A, 457, 841

Ilbert, O., Capak, P., Salvato, M., et al. 2009, ApJ, 690, 1236

Ilbert, O., McCracken, H. J., Le Fèvre, O., et al. 2013, A&A, 556, A55

Ilbert, O., Salvato, M., Le Floc'h, E., et al. 2010, ApJ, 709, 644

Johnson, B. D., Schiminovich, D., Seibert, M., et al. 2007, ApJS, 173, 377

Juneau, S., Glazebrook, K., Crampton, D., et al. 2005, ApJ, 619, L135

Kaiser, N. 1984, ApJ, 284, L9

Karim, A., Schinnerer, E., Martínez-Sansigre, A., et al. 2011, ApJ, 730, 61

Kassin, S. A., Weiner, B. J., Faber, S. M., et al. 2012, ApJ, 758, 106

Kennicutt, Jr., R. C. 1998, ARA&A, 36, 189

Kereš, D., Katz, N., Weinberg, D. H., & Davé, R. 2005, MNRAS, 363, 2

Kim, J.-h., Abel, T., Agertz, O., et al. 2014, ApJS, 210, 14

Kraljic, K., Arnouts, S., Pichon, C., et al. 2018, MNRAS, 474, 547

Laigle, C., Pichon, C., Arnouts, S., et al. 2018, MNRAS, 474, 5437

Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, ArXiv 1110.3193

Lilly, S. J., Le Fevre, O., Hammer, F., & Crampton, D. 1996, ApJ, 460, L1

Llebaria, A., Magnelli, B., Arnouts, S., et al. 2008, in SPIE, Vol. 6812, Conference Series

Madau, P. & Dickinson, M. 2014, ARA&A, 52, 415

Madau, P., Pozzetti, L., & Dickinson, M. 1998, ApJ, 498, 106

Malavasi, N., Arnouts, S., Vibert, D., et al. 2017, MNRAS, 465, 3817

Maraston, C. 2005, MNRAS, 362, 799

Martig, M., Bournaud, F., Teyssier, R., & Dekel, A. 2009, ApJ, 707, 250

Martin, D. C., Fanson, J., Schiminovich, D., et al. 2005, The Astrophysical Journal, 619, L1

McQuinn, M. 2016, ARA&A, 54, 313

Meurer, G. R., Heckman, T. M., & Calzetti, D. 1999, ApJ, 521, 64

Mo, H. J., Mao, S., & White, S. D. M. 1998, MNRAS, 295, 319

Moutard, T., Arnouts, S., Ilbert, O., et al. 2016a, A&A, 590, A103

Moutard, T., Arnouts, S., Ilbert, O., et al. 2016b, A&A, 590, A102

Noeske, K. G., Faber, S. M., Weiner, B. J., et al. 2007a, ApJ, 660, L47

Noeske, K. G., Weiner, B. J., Faber, S. M., et al. 2007b, ApJ, 660, L43

Pasquet, J., Bertin, E., Treyer, M., Arnouts, S., & Fouchez, D. 2018, ArXiv 1806.06607

Peng, Y., Maiolino, R., & Cochrane, R. 2015, Nature, 521, 192

Peng, Y.-j., Lilly, S. J., Kovač, K., et al. 2010, ApJ, 721, 193

Perlmutter, S., Gabi, S., Goldhaber, G., et al. 1997, ApJ, 483, 565

Pforr, J., Maraston, C., & Tonini, C. 2012, MNRAS, 422, 3285

Pichon, C., Pogosyan, D., Kimm, T., et al. 2011, MNRAS, 418, 2493

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 594, A13

Postman, M. & Geller, M. J. 1984, ApJ, 281, 95

Puschell, J. J., Owen, F. N., & Laing, R. A. 1982, ApJ, 257, L57

Salim, S., Charlot, S., Rich, R. M., et al. 2005, ApJ, 619, L39

Schiminovich, D., Ilbert, O., Arnouts, S., et al. 2005, The Astrophysical Journal, 619, L47

Silk, J. & Mamon, G. A. 2012, Research in Astronomy and Astrophysics, 12, 917

Somerville, R. S., Lee, K., Ferguson, H. C., et al. 2004, ApJ, 600, L171

Sousbie, T. 2011, MNRAS, 414, 350

Springel, V. 2010, ARA&A, 48, 391

Springel, V., Frenk, C. S., & White, S. D. M. 2006, Nature, 440, 1137

Tempel, E., Stoica, R. S., & Saar, E. 2013, MNRAS, 428, 1827

Trujillo, I., Carretero, C., & Patiri, S. G. 2006, ApJ, 640, L111

Walcher, J., Groves, B., Budavári, T., & Dale, D. 2011, Ap&SS, 331, 1

White, S. D. M. 1984, ApJ, 286, 38

White, S. D. M. & Rees, M. J. 1978, MNRAS, 183, 341

Xu, C. K., Donas, J., Arnouts, S., et al. 2005, ApJ, 619, L11

Zel'dovich, Y. B. 1970, A&A, 5, 84

4

Travaux sélectionnés

4.1 Le destin des galaxies

Multi wavelength imaging surveys and Photometric redshifts

– Bertin & **Arnouts**, A&A 1996
SExtractor : Software for source extraction
– **Arnouts**, de Lapparent, Mathez, Mazure, Mellier et al., A&AS 124, 1997
The ESO-Sculptor faint galaxy redshift survey : The photometric sample
– **Arnouts**, D'Odorico, Cristiani, Zaggia, Fontana et al., MNRAS 310, 1999
The NTT SUSI deep field
– **Arnouts**, Vandame, Benoist, Groenewegen, da Costa et al., A&A 379, 2001
ESO imaging survey. Deep public survey : Multi-color optical data for the CDFS
– Xu, Donas, **Arnouts**, Wyder, Seibert et al., ApJ 619, 2005
Number Counts of GALEX Sources in Far-UV(1530Å) & Near-UV (2310Å) bands
– Ilbert, **Arnouts**, Mc Cracken et al., A&A 2006
Accurate photometric redshifts for the CFHTLS calibrated using the VVDS
– Hildebrandt, **Arnouts**, Capak, Moustakas, Wolf et al., A&A 523, 2010
PHAT : PHoto-z Accuracy Testing
– Moutard, **Arnouts**, Ilbert, Coupon, Hudelot et al., A&A 590, 2016
The VIPERS Multi-Lambda Survey I : UV and NIR observations, catalogues & photo-z
– Pasquet, Bertin, Treyer, **Arnouts** & Fouchez, 2018, soumis à A&A, ArXiv 1806.06607
Photometric redshifts from SDSS images using a Convolutional Neural Network

From Luminosity Functions to cosmic Star Formation Rate and Stellar Mass Assembly

– **Arnouts**, Schiminovich, Ilbert, Tresse, Milliard et al., ApJ 619, 2005
The GALEX VIMOS-VLT Deep Survey Measurement of the Evolution of the 1500 Å LF
– Schiminovich, Ilbert, **Arnouts**, Milliard, Tresse et al., ApJ 619, 2005
The GALEX-VVDS Measurement of the Evolution of the Far-UV LD and cosmic SFR
– **Arnouts**, Walcher, Le Fèvre, Zamorani, Ilbert et al., A&A 476, 2007
The SWIRE-VVDS-CFHTLS surveys : stellar mass assembly over the last 10 Gyr.
– Walcher, Lamareille, Vergani, **Arnouts** , Buat et al., A&A 491, 2008
The VVDS-SWIRE-GALEX-CFHTLS surveys : physical properties of galaxies at z≤1.2
– Goto, **Arnouts**, Malkan, Takagi, Inami et al., MNRAS 414, 2011
Infrared luminosity functions of AKARI Sloan Digital Sky Survey galaxies
– Goto, **Arnouts**, Inami, Matsuhara, Pearson et al., MNRAS 410, 2011
LFs of IR galaxies with AKARI : implications for the cosmic SFH & AGN evolution
– **Arnouts**, Le Floc'h, Chevallard, Johnson, Ilbert et al., A&A 558, 2013
Encoding of the infrared excess in the NUVrK color diagram for star-forming galaxies
– Ilbert, **Arnouts**, Le Floc'h, Aussel, Bethermin et al., A&A 579, 2015
Evolution of the specific star formation rate function at z≤ 1.4
– Moutard, **Arnouts**, Ilbert, Coupon, Davidzon et al., A&A 590, 2016
The VIPERS MLS. II. Diving with massive galaxies in 22 square degrees since z=1.5

4.2 Agrégation des galaxies et Toile cosmique

Galaxy Clustering

– **Arnouts**, Cristiani, Moscardini, Matarrese, Lucchin al. MNRAS 310, 1999
"Measuring and modelling the redshift evolution of clustering : the HDF North"
– **Arnouts**, Moscardini, Vanzella, Colombi, Cristiani et al., MNRAS 329, 2002
"Measuring the redshift evolution of clustering : the HDF South"
– Heinis, Treyer, **Arnouts**, Milliard, Donas A&A 424, 2004
"The clustering of ultraviolet-selected galaxies at z~0.1"
– Heinis, Milliard, **Arnouts**, Blaizot, Schiminovich et al., ApJS 173, 2007
"Clustering Properties of Rest-Frame UV-Selected Galaxies II. Migration of Star Formation Sites with Cosmic Time"
– Heinis, Budavari, Szalay, **Arnouts**, Aragon-Calvo et al., ApJ 698, 2009
"Spatial Clustering from GALEX-SDSS Samples : Star Formation History and Large-Scale Clustering"
– Coupon, **Arnouts**, Van Waerbeke, Moutard, Ilbert et al., MNRAS 449, 2015
"The galaxy-halo connection : a lensing, clustering and abundance analysis"
– Moutard, Sawicki, **Arnouts**, Golob, Malavasi et al., 2018, MNRAS 479, 2147
"On the fast quenching of young low-mass galaxies up to z~0.6. New spotlight on the lead role of environment"

Cosmic Web

– Scoville, **Arnouts**, Aussel, Benson, Bongiorno et al., ApJS 206, 2013
Evolution of Galaxies and Their Environments at z = 0.1-3 in COSMOS
– Malavasi, **Arnouts**, Vibert, de la Torre, Moutard et al., MNRAS 465, 2017
VIPERS : Galaxy segregation inside filaments at z=0.7
– Kraljic, **Arnouts**, Treyer, Pichon, Vibert et al., MNRAS 474, 2018
The Cosmic Web Reconstruction in GAMA survey at $z \leq 0.3$
– Laigle, Pichon, **Arnouts**, McCracken, Dubois et al., MNRAS 474, 2018
Photometric redshifts probe the impact of filaments on galaxy properties
–Treyer, Kraljic, **Arnouts**, de La Torre, Pichon et al., MNRAS 477, 2018
Group quenching and galactic conformity at low redshift

4.3 Articles sélectionnés

Abstracts des articles choisis suivis des manuscrits intégraux.

Article 1 : **The SWIRE-VVDS-CFHTLS surveys : stellar mass assembly & build up of the red sequence**
Résumé : We present an analysis of the stellar mass growth over the last 10 Gyr (z≤ 2) using a unique large sample of galaxies selected at 3.6μ?m. We have assembled accurate photometric and spectroscopic redshifts for ~21,200 and 1500 galaxies, respectively, with F(3.6 μm)≥9.0μJy by combining data from Spitzer-SWIRE IRAC, the VIMOS VLT Deep Survey (VVDS), UKIDSS and very deep optical CFHTLS photometry. We split our sample into quiescent (red) and active (blue) galaxies on the basis of an SED fitting procedure that we have compared with the strong rest-frame color bimodality $(NUV - r)_{ABS}$. The present sample contains ~4400 quiescent galaxies. Our measurements of the K-rest frame luminosity function and luminosity density evolution support the idea that a large fraction of galaxies is already assembled at z ~1.2, with almost 80% and 50% of the active and quiescent populations already in place, respectively. Based on the analysis of the evolution of the stellar mass-to-light ratio (in K-band) for the spectroscopic sub-sample, we derive the stellar mass density for the entire sample. We find that the global evolution of the stellar mass density is well reproduced by the star formation rate derived from UV based measurements when an appropriate dust correction is applied, which supports the idea of an initial mass function that is on average universal. Over the last 8 Gyr (z ≤1.2) we observe that the stellar mass density of the active population shows a modest mass growth rate ($\rho \sim 0.005(\pm 0.005) M_\odot/Mpc^3/yr$), consistent with a constant stellar mass density, $\rho^{active} \sim 3.1\ 10^8 M/Mpc^3$. In contrast, an increase by a factor of 2 for the quiescent population over the same timescale is observed. As a consequence, the growth of the stellar mass in the quiescent population must be due to the shutoff of star formation in active galaxies that migrate into the quiescent population. We estimate this stellar mass flux to be $\rho_{A \to Q} \sim 0.017(\pm 0.004) M_\odot/Mpc^3/yr$, which balances the major fraction of new stars born according to our best SFR estimate ($\rho = 0.025(\pm 0.003) M_\odot/Mpc^3/yr$). From z = 2 to z = 1.2, we observe a major build-up of the quiescent population with an increase by a factor of ?10 in stellar mass (a mass growth rate of $\sim 0.063 M_\odot/Mpc^3/yr$). This rapid evolution suggests that we are observing the epoch when, for the first time in the history of the universe, an increasing fraction of galaxies end their star formation activity and start to build up the red sequence.

Article 2 : **Encoding of the infrared excess in the NUVrK color diagram for star-forming galaxies**
Résumé : We present an empirical method of assessing the star formation rate (SFR) of star-forming galaxies based on their locations in the rest-frame color?color diagram $(NUV - r) vs. (r - K)$. By using the Spitzer 24μm sample in the COSMOS field (~16400 galaxies with 0.2< z <1.3) and a local GALEX-SDSS-SWIRE sample (~700 galaxies with z≤0.2), we show that the mean infrared excess $< IRX >=< L_{IR}/L_{UV} >$ can be described by a single vector, NRK, that combines the two colors. The calibration between IRX and NRK allows us to recover the IR luminosity, L_{IR}, with an accuracy of $\sigma \sim 0.21$ for the COSMOS sample and 0.27 dex for the local one. The SFRs derived with this method agree with the ones based on the observed (UV+IR) luminosities and on the spectral energy distribution (SED) fitting for the vast majority (~85%) of the star-forming population. Thanks to a library of model galaxy SEDs with realistic prescriptions for the star formation history, we show that we need to include a two-component dust model (i.e., birth clouds and diffuse ISM) and a full distribution of galaxy inclinations in order to reproduce the behavior of the IRX stripes in the NUVrK diagram. In conclusion, the NRK method, based only on the rest-frame UV/optical colors available in most of the extragalactic fields, offers a simple alternative of assessing the SFR of star-forming galaxies in the absence of far-IR or spectral diagnostic observations.

Article 3 : **The VIPERS-MLS – II : Evolution of massive galaxies at** $z < 1.5$
Résumé : We investigate the evolution of the galaxy stellar mass function and stellar mass density from redshift $z = 0.2$ to $z = 1.5$ of a $K_s < 22$-selected sample with highly reliable photometric redshifts and over an unprecedentedly large area. Our study is based on near-infrared observations carried out with the WIRCam instrument at CFHT over the footprint of the VIPERS spectroscopic survey and benefits from the high-quality optical photometry from the CFHTLS and ultraviolet observations with the GALEX satellite. The accuracy of our photometric redshifts is $\sigma_{\Delta z/(1+z)}$ <0.03 and 0.05 for the bright ($i_{AB} < 22.5$) and faint ($i_{AB} > 22.5$) samples,

respectively. The galaxy stellar mass function is measured with ~760,000 galaxies down to $K_s \sim 22$ and over an effective area of ~ 22.4 deg^2, the latter of which drastically reduces the statistical uncertainties (i.e. Poissonian error and cosmic variance). We point out the importance of carefully controlling the photometric calibration, whose effect becomes quickly dominant when statistical uncertainties are reduced, which will be a major issue for future cosmological surveys with EUCLID or LSST, for instance. By exploring the rest-frame $(NUV - r)$ vs $(r - K_s)$ colour-colour diagram with which we separated star-forming and quiescent galaxies, (1) we find that the density of very massive $\log(M_*/M_\odot) > 11.5$ galaxies is largely dominated by quiescent galaxies and increases by a factor 2 from z~ 1 to $z \sim 0.2$, which allows for additional mass assembly through dry mergers. (2) We also confirm the scenario in which star formation activity is impeded above a stellar mass $\log(\mathcal{M}^{\star}_{\rm SF}/M_\odot) = 10.64 \pm 0.01$. This value is found to be very stable at $0.2 < z < 1.5$. (3) We discuss the existence of a main quenching channel that is followed by massive star-forming galaxies, and we finally (4) characterise another quenching mechanism that is required to explain the clear excess of low-mass quiescent galaxies that is observed at low redshift.

Article 4 : **Measuring and modelling the redshift evolution of clustering : the HDF North**
Résumé : The evolution of galaxy clustering from z=0 to z=4.55 is analysed using the angular correlation function and the photometric redshift distribution of galaxies brighter than $I_{AB} \leq 28.5$ in the Hubble Deep Field North. The reliability of the photometric redshift estimates is discussed on the basis of the available spectroscopic redshifts, comparing different codes and investigating the effects of photometric errors. The redshift bins in which the clustering properties are measured are then optimized to take into account the uncertainties of the photometric redshifts. The results show that the comoving correlation length r_0 has a small decrease in the range $0 < z < 1$ followed by an increase at higher z. We compare these results with the theoretical predictions of a variety of cosmological models belonging to the general class of Cold Dark Matter scenarios, including Einstein-de Sitter models, an open model and a flat model with non-zero cosmological constant. Comparison with the expected mass clustering evolution indicates that the observed high-redshift galaxies are biased tracers of the dark matter with an effective bias b strongly increasing with redshift. Assuming an Einstein-de Sitter universe, we obtain b=2.5 at z=2 and b=5 at z=4. These results support theoretical scenarios of biased galaxy formation in which the galaxies observed at high redshift are preferentially located in more massive haloes. Moreover, they suggest that the usual parameterization of the clustering evolution as $\xi(r,z) \sim \xi(r,0)(1+z)^{-(3+\epsilon)}$ is not a good description for any value of ϵ. Comparison of the clustering amplitudes that we measured at z =3 with those reported by Adelberger et al. and Giavalisco et al., based on a different selection, suggests that the clustering depends on the abundance of the objects : more abundant objects are less clustered, as expected in the paradigm of hierarchical galaxy formation. The strong clustering and high bias measured at z=3 are consistent with the expected density of massive haloes predicted in the frame of the various cosmologies considered here. At z=4, the strong clustering observed in the Hubble Deep Field requires a significant fraction of massive haloes to be already formed by that epoch. This feature could be a discriminant test for the cosmological parameters if confirmed by future observations.

Article 5 : **The galaxy-halo connection from a joint lensing, clustering and abundance analysis**
Résumé : We present new constraints on the relationship between galaxies and their host dark matter haloes, measured from the location of the peak of the stellar-to-halo mass ratio (SHMR), up to the most massive galaxy clusters at redshift z~0.8 and over a volume of nearly 0.1 Gpc^3. We use a unique combination of deep observations in the CFHTLenS/VIPERS field from the near-UV to the near-IR, supplemented by ~60,000 secure spectroscopic redshifts, analyzing galaxy clustering, galaxy ?galaxy lensing and the stellar mass function. We interpret our measurements within the halo occupation distribution (HOD) framework, separating the contributions from central and satellite galaxies. We find that the SHMR for the central galaxies peaks at $M_{h,peak} = 1.9 \times 10^{12} M_\odot$ with an amplitude of 0.025, which decreases to 0.001 for massive haloes ($M_h > 10^{14} M_\odot$). Compared to central galaxies only, the total SHMR (including satellites) is boosted by a factor of 10 in the high-mass regime (cluster-size haloes), a result consistent with cluster analyses from the literature based on fully independent methods. After properly accounting for differences in modeling, we have compared our results with a large number of results from the literature up to z = 1 : we find good general agreement, independently of the method used, within the typical stellar-mass systematic errors at low to intermediate mass ($M_\star < 10^{11} M_\odot$) and the statistical errors above. We have also compared our SHMR results to semi-analytic simulations and found that the SHMR is tilted compared to our

measurements in such a way that they over- (under-) predict star formation efficiency in central (satellite) galaxies.

Article 6 : **The cosmic Web reconstruction in the GAMA survey at** $z \leq 0.3$
Résumé : The role of the cosmic web in shaping galaxy properties is investigated in the GAMA spectroscopic survey in the redshift range $0.03 < z < 0.25$. The stellar mass, $(u - r)$ dust corrected colour and specific star formation rate (sSFR) of galaxies are analysed as a function of their distances to the 3D cosmic web features, such as nodes, filaments and walls, as reconstructed by DisPerSE. Significant mass and type/colour gradients are found for the whole population, with more massive and/or passive galaxies being located closer to the filament and wall than their less massive and/or starforming counterparts. Mass segregation persists among the starforming population alone. The red fraction of galaxies increases when closing in on nodes, and on filaments regardless of the distance to nodes. Similarly, the star-forming population reddens (or lowers its sSFR) at fixed mass when closing in on filament, implying that some quenching takes place. These trends are also found in the state-of-the-art hydrodynamical simulation HORIZON-AGN. These results suggest that on top of stellar mass and large-scale density, the traceless component of the tides from the anisotropic large-scale environment also shapes galactic properties. An extension of excursion theory accounting for filamentary tides provides a qualitative explanation in terms of anisotropic assembly bias : at a given mass, the accretion rate varies with the orientation and distance to filaments. It also explains the absence of type/colour gradients in the data on smaller, non-linear scales.

4.3.1 Article 1 : The SWIRE-VVDS-CFHTLS surveys : stellar mass assembly & build up of the red sequence

A&A 476, 137–150 (2007)
DOI: 10.1051/0004-6361:20077632

Astronomy & Astrophysics

The SWIRE-VVDS-CFHTLS surveys: stellar mass assembly over the last 10 Gyr. Evidence for a major build up of the red sequence between $z = 2$ and $z = 1$ ⋆

S. Arnouts[1,2], C. J. Walcher[2], O. Le Fèvre[2], G. Zamorani[3], O. Ilbert[4], V. Le Brun[2], L. Pozzetti[3], S. Bardelli[3], L. Tresse[2], E. Zucca[3], S. Charlot[5], F. Lamareille[6], H. J. McCracken[5], M. Bolzonella[3], A. Iovino[7], C. Lonsdale[8], M. Polletta[8], J. Surace[9], D. Bottini[10], B. Garilli[10], D. Maccagni[10], J. P. Picat[6], R. Scaramella[3], M. Scodeggio[10], G. Vettolani[3], A. Zanichelli[3], C. Adami[2], A. Cappi[3], P. Ciliegi[3], T. Contini[6], S. de la Torre[2], S. Foucaud[14], P. Franzetti[10], I. Gavignaud[15], L. Guzzo[7], B. Marano[3], C. Marinoni[13], A. Mazure[2], B. Meneux[7], R. Merighi[3], S. Paltani[12], R. Pellò[6], A. Pollo[2], M. Radovich[11], S. Temporin[7], and D. Vergani[10]

(Affiliations can be found after the references)

Received 12 April 2007 / Accepted 18 September 2007

ABSTRACT

We present an analysis of the stellar mass growth over the last 10 Gyr ($z \leq 2$) using a unique large sample of galaxies selected at 3.6 μm. We have assembled accurate photometric and spectroscopic redshifts for ∼21 200 and 1500 galaxies, respectively, with $F(3.6\,\mu\mathrm{m}) \geq 9.0\,\mu$Jy by combining data from Spitzer-SWIRE IRAC, the VIMOS VLT Deep Survey (VVDS), UKIDSS and very deep optical CFHTLS photometry. We split our sample into quiescent (red) and active (blue) galaxies on the basis of an SED fitting procedure that we have compared with the strong rest-frame color bimodality $(NUV - r')_{\mathrm{ABS}}$. The present sample contains ∼4400 quiescent galaxies. Our measurements of the K-rest frame luminosity function and luminosity density evolution support the idea that a large fraction of galaxies is already assembled at $z \sim 1.2$, with almost 80% and 50% of the active and quiescent populations already in place, respectively.

Based on the analysis of the evolution of the stellar mass-to-light ratio (in K-band) for the spectroscopic sub-sample, we derive the stellar mass density for the entire sample. We find that the global evolution of the stellar mass density is well reproduced by the star formation rate derived from UV based measurements when an appropriate dust correction is applied, which supports the idea of an initial mass function that is on average universal.

Over the last 8 Gyr ($z \leq 1.2$) we observe that the stellar mass density of the active population shows a modest mass growth rate ($\dot{\rho} \sim 0.005(\pm 0.005)\ M_{\odot}/\mathrm{Mpc}^3/\mathrm{yr}$), consistent with a constant stellar mass density, $\rho_{\star}^{\mathrm{active}} \sim 3.1 \times 10^8\ M_{\odot}/\mathrm{Mpc}^3$. In contrast, an increase by a factor of ∼2 for the quiescent population over the same timescale is observed. As a consequence, the growth of the stellar mass in the quiescent population must be due to the shutoff of star formation in active galaxies that migrate into the quiescent population. We estimate this stellar mass flux to be $\dot{\rho}_{A \rightarrow Q} \sim 0.017(\pm 0.004)\ M_{\odot}/\mathrm{Mpc}^3/\mathrm{yr}$, which balances the major fraction of new stars born according to our best SFR estimate ($\dot{\rho} = 0.025(\pm 0.003)\ M_{\odot}/\mathrm{Mpc}^3/\mathrm{yr}$).

From $z = 2$ to $z = 1.2$, we observe a major build-up of the quiescent population with an increase by a factor of ∼10 in stellar mass (a mass growth rate of ∼0.063 $M_{\odot}/\mathrm{Mpc}^3/\mathrm{yr}$). This rapid evolution suggests that we are observing the epoch when, for the first time in the history of the universe, an increasing fraction of galaxies end their star formation activity and start to build up the red sequence.

Key words. galaxies: luminosity function, mass function – galaxies: formation

1. Introduction

The strong decline of the cosmic star formation rate (SFR) since $z \sim 1$ is now well established (Schiminovich et al. 2005; Hopkins et al. 2006), and has been shown to be accompanied by a decrease of faint star forming galaxies and the decline of luminous ultra-violet galaxies (Lilly et al. 1996; Arnouts et al. 2005). One fundamental question is to understand the link between the global decline of the star formation rate and the evolution of the mass assembly. Inspection of the specific star formation rate (SFR per unit of stellar mass) in individual galaxies reveals that the preferred site of star formation activity has migrated from massive systems at high z to low mass systems at low-z. This is usually referred to as the downsizing effect (Cowie et al. 1996; Juneau et al. 2005). Evidences of such an effect have been seen in the fundamental plane relation of elliptical galaxies (Treu et al. 2005) as well as in the analysis of spectra of local galaxies (Heavens et al. 2004; Kauffmann et al. 2003) where both analyses show that massive galaxies have formed their stars earlier than less massive ones. Clustering properties of star-forming galaxies at high and low redshifts reveal that the bulk of star formation activity has migrated from massive dark matter halos (DMH) at high z to low mass DMH at low z,

⋆ Based on data obtained with the European southern observatory Very Large Telescope, Paranal, Chile, program 070A-9007(A) and on observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/DAPNIA, at the Canada-France-Hawaii Telescope (CFHT) which is operated by the National Research Council (NRC) of Canada, the Institut National des Science de l'Univers of the Centre National de la Recherche Scientifique (CNRS) of France, and the University of Hawaii. This work is based in part on data products produced at TERAPIX and the Canadian Astronomy Data Centre as part of the Canada-France-Hawaii Telescope Legacy Survey, a collaborative project of NRC and CNRS and on data obtained as part of the UKIRT Infrared Deep Sky Survey.

Article published by EDP Sciences and available at http://www.aanda.org or http://dx.doi.org/10.1051/0004-6361:20077632

generalizing the downsizing effect from stellar mass to the dark matter (Heinis et al. 2007).

While the ΛCDM hierarchical scenario successfully describes the clustering properties of galaxies on large scales (Mo & White 1996; Springel et al. 2006), the quenching of star formation in massive systems is less well understood. It relies on the complex physics of baryons and there is not yet enough constraints on the mechanisms involved in the regulation of star formation. These mechanisms may act on galactic scales like AGN, Supernovae feedback (Benson et al. 2003; Croton et al. 2006) or on large scales via merging or gas heating in dense environments (Naganime et al. 2001; Yoshikawa et al. 2001). For instance, to reproduce the properties of local massive elliptical galaxies, de Lucia et al. (2006) show that the stars could form first at high redshift in sub-galactic units, while the galaxies would continue their stellar mass assembly through merging onto low redshift.

In this context, the most stringent constraint for the semi-analytical models is the direct observation of the number density of massive galaxies at high redshift. Evidences of the existence of massive galaxies at high redshift are becoming numerous. An important population of EROs at $z \geq 1$ has been discovered (K20 survey: Cimatti et al. 2002), with roughly half of them being old stellar systems (Cimatti et al. 2004). They could be the already assembled progenitors of local massive ellipticals, as supported by their clustering signals and number density (Daddi et al. 2002; Arnouts 2003). The measures of the evolution of the stellar mass density from NIR surveys found that half of the stellar mass is already in place at $z \leq 1{-}1.5$ (Fontana et al. 2003, 2004; Pozzetti et al. 2003; Drory et al. 2004, 2005; Caputi et al. 2005) moving the formation epoch close to the peak of the cosmic SFR. The analysis of the galaxy mass function (GMF) at different redshifts provides additional information on how the process of mass assembly acts on different mass scales. Such analysis is becoming common among the recent surveys (Fontana et al. 2004, 2006; Drory et al. 2004, 2005; Bundy et al. 2005; Borch et al. 2006; Franceschini et al. 2006; Pozzetti et al. 2007). While the different studies do not necessarily agree with each other, it emerges that the most massive galaxies have undergone less evolution than less massive ones since $z \sim 1.0$ and 50 to 70% are already in place at such a redshift, revealing a faster assembly for the most massive systems.

An alternative approach to measure the density of massive galaxies at high redshift is to investigate the evolution of the local elliptical galaxies that dominate the massive end of the GMF (Bell et al. 2003; Baldry et al. 2004; Bundy et al. 2006; Cirasuolo et al. 2006). This has been done by following the redshift evolution of the red sequence, defined by the $(U - V)$ or $(U - B)$ colors, from today up to $z \sim 1.1$. Bell et al. (2004) and Faber et al. (2005) find that the density of red galaxies drops by a factor 4 or 2.5 respectively, suggesting that their stellar mass density roughly doubles, while at the same time the blue population shows little evolution. They interpret this evolution by the migration of blue galaxies that quenched their star formation and migrated into the red sequence. Bell et al. (2004) also reported a significant evolution of the most massive ellipticals. Because there is a shortage of massive blue galaxies able to produce those massive ellipticals, they introduce the idea of "dry" or purely stellar mergers (merging between red galaxies) as a possible scenario to produce these massive local ellipticals.

Analysis from the VVDS survey by Zucca et al. (2006), who define the early-type class based on SED fitting, report a more modest decline (∼40%) in the number density of this population up to $z \sim 1.1$, which is also consistent with the analysis by Ilbert et al. (2006b) based on a morphologically selected elliptical sample in the CDF South.

Similarly to VVDS results, Brown et al. (2007), who have selected ellipticals with same optical color criteria than Bell et al. (2004), in an area of 7 deg^2, do not observe any evolution of comoving density ($\Phi^\star$) up to $z = 1$ and according to the modest luminosity density increase (∼36%), predicts that the stellar mass has roughly doubled since $z \sim 1$. Moreover, they observe a modest evolution of luminous ellipticals (with $L \geq 4\ L^\star$) with 80% of their stellar mass in place at $z \sim 0.8$, suggesting that "dry" mergers should not play a dominant role in the evolution of ellipticals over the last 8 Gyr.

Complementary analysis by Bundy et al. (2006) find signatures of a downsizing effect with quenching affecting first the massive galaxies and then moving to lower mass systems. This can be interpreted as an anti-hierarchical process with massive early types ending first their mass assembly while the less massive ones are still assembling their mass (see also Cimatti et al. 2006). If such a process is confirmed, it would be a strong observational constraint for the models of galaxy formation within the standard paradigm.

A large fraction of the surveys discussed above are limited to $z \leq 1{-}1.2$, due to optical selection, while the deepest surveys are still covering modest sizes in particular when the sample is splitted by galaxy types. The impact of cosmic variance on results coming from small fields remains a major problem. Even in the three fields covered by COMBO17, it still appears to play a significant role (Bell et al. 2004; Somerville et al. 2004). The selection of the samples varies from optical to IR band and the stellar mass is estimated from various approaches including SED fitting, rest-frame colors, with or without infrared information, introducing significant dispersions in the final estimates. Additionally, the selection of elliptical galaxies based on optical colors has been found to show a significant contamination by Sey2 AGN and star forming galaxies up to 40% (Franzetti et al. 2007).

In this paper we explore the redshift evolution of the stellar mass assembly for a unique large sample selected in the observed Mid-Infrared, via the analysis of the K-rest luminosity functions, the luminosity density and the evolution of the mass to light ratio. We take advantage of a native IRAC selection that allows us to detect massive red objects up to $z \sim 2$, and we isolate a large sample of red/quiescent galaxies according to their low level of star formation activity. We then discuss implications for the stellar mass assembly of the active and quiescent populations. Throughout the paper we adopt the concordance cosmology $\Omega_\mathrm{m} = 0.3$ and $\Omega_\Lambda = 0.7$ and $H_0 = 70\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$. All magnitudes are in the AB system, and for convenience we adopt a Salpeter IMF unless otherwise specified.

2. The data

In this work we make use of the large amount of data collected in the VVDS-0226-04 field. The present sample is based on a 3.6 μm flux limited sample from the SWIRE survey which overlaps the deep multi-colour imaging survey from the CFHTLS. The common area between the two surveys corresponds to 0.85 square degree. We complete this dataset by including the deep spectro-photometric data from the VVDS survey and the infrared data from the UKIDSS survey.

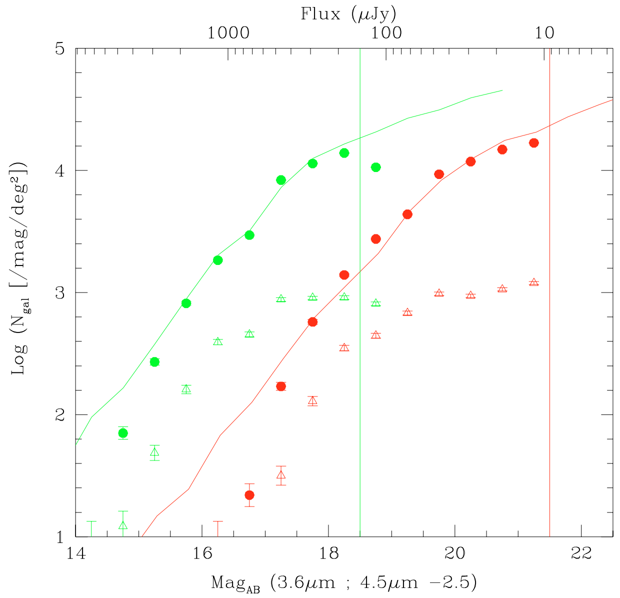


Fig. 1. Stars (triangles) and Galaxy (circles) number counts for the 3.6 μm (red symbols) and 4.5 μm (green symbols) samples. Solid lines show the IRAC number counts from Fazio et al. (2005).

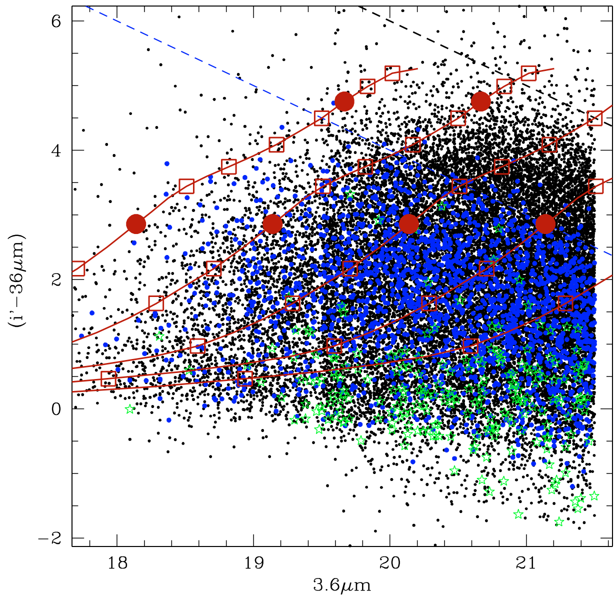


Fig. 2. Color–magnitude distribution with (i – 3.6 μm) vs. 3.6 μm. We show the spectroscopic sample for galaxies (filled blue symbols) and for stars (green stars symbols). The spectroscopic flux limit ($I \leq 24$) and the 5σ detection limit for the CFHTLS ($i' \sim 26$) are shown as dashed lines (lower and upper line respectively). The solid lines show the behaviour of an elliptical galaxy (with τ = 0.1 Gyr and $z_{\rm form}$ = 6) moving from z = 2.5 to z = 0 (from red to blue color) and for different absolute K magnitudes between $K_{\rm ABS} = -25$ (left solid line) and $K_{\rm ABS} = -21$ (right solid line). The large circles denote the redshifts: z = 1.0, 2.0, while open squares are spaced by $\delta z = 0.2$.

2.1. The SWIRE sample

The present sample corresponds to a 3.6 μm flux limited sample with $F(3.6\ \mu\text{m}) \geq 9\ \mu$Jy (or $m_{\rm AB}(3.6) \leq 21.5$), based on the SWIRE survey (Lonsdale et al. 2003). The SWIRE photometry is based on the band-merged catalog including 3.6, 4.5, 5.6, 8.0 μm and 24 μm passband (Surace et al. 2005). with a typical 5σ depth of 5.0, 9.0, 43, 40 and 311 μJy respectively We use the flux measurements derived in 3 arcsec apertures for faint sources as suggested by Surace et al. (2005), while we adopt the adaptive apertures (Kron magnitude; Bertin & Arnouts 1996) for the bright sources ($m_{\rm AB}(3.6) \leq 19.5$). In Fig. 1, we show the IRAC 3.6 and 4.5 μm number counts. When comparing with the number counts from Fazio et al. (2005), the 3.6 μm galaxy number counts are ~80% complete at 9 μJy. At this depth within the common area with CFHTLS, we have ~25 500 sources.

2.2. The CFHTLS data

The deep multi-colour photometry ($u^*g'r'i'z'$) from the Canada-France-Hawaii Telescope Legacy Survey is based on the T0003 release (CFHTLS-D1). These data cover one square degree, with sub-arcsecond seeing in all bands and reach the limiting magnitudes (corresponding to 50% completeness) of 26.6, 26.5, 26.0, 26.0 and 25.2 in u^*, g', r', i', z' respectively. A full description of the data will be presented in a forthcoming paper by Mc Cracken et al. (in prep.).

Thanks to the very deep optical data, almost all the IRAC sources have an optical counterpart, except a negligible fraction (~0.5%), due to unmatched positions within 1.5 arcsec.

2.3. The VVDS data

The VIMOS VLT Deep Survey consists of deep photometry and spectroscopy (Le Fèvre et al. 2005):

- Deep B, V, R, I imaging with a depth (50% completeness) of 26.5, 26.2, 25.9, 25.0, respectively (Mc Cracken et al. 2003). In addition, J and K observations with NTT-SOFI have been obtained at the depth (50% completeness) of 24.2 and 23.8, respectively over 172 arcmin^2 (Iovino et al. 2005).
- The first epoch VVDS spectroscopic sample is based on a randomly selected sample of ~9000 sources in the magnitude range $17 \leq I \leq 24.0$ (Le Fèvre et al. 2004). The spectroscopic area overlapping the SWIRE-CFHTLS data is 0.42 deg^2 and provides ~1500 secure spectra for sources with $m_{\rm AB}(3.6\ \mu\text{m}) \leq 21.5$.

2.4. The UKIDSS data

Finally, we complete our dataset with the J and K photometry from the UKIDSS Ultra Deep Survey (Lawrence et al. 2006) based on the DR1 release (Warren et al. 2007). These data reach a depth (5σ limits) of 22.5 and 22.0 in J and K, respectively. The overlap with SWIRE-CFHTLS is 0.55 and 0.76 deg^2 for J and K bands, respectively. We match the UDS sources with the optical data within a 1 arcsec radius. Within the common areas, 89% (or 93%) of the 3.6 selected sources have a J (or K) flux measurement.

2.5. The combined sample

In Fig. 2, we show the (i – 3.6 μm) vs. 3.6 μm color–magnitude diagram for the whole sample (small dots). The galaxies and stars from the spectroscopic sample are shown as blue and green symbols respectively. We plot the cut-offs introduced in the spectroscopic sample due to the magnitude limit, $I_{\rm AB} \leq 24$ (lower dashed line) and in the CFHTLS by adopting the 5σ detection limit, $i' \sim 26$ (upper dashed line). The evolutionary tracks of an elliptical galaxy, formed at z = 6 and with an e-folding parameter τ = 0.1 Gyr (using the PEGASE model; Fioc et al. 1997)

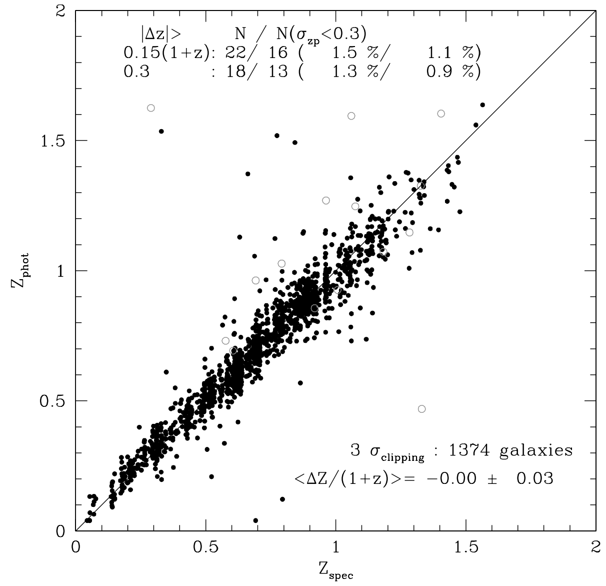


Fig. 3. Photometric vs. spectroscopic redshift comparison, for ~1400 secure VVDS spectroscopic redshifts. Open symbols refer to objects with internal photo-z uncertainties $\sigma(z_{\rm phot}) \geq 0.3$.

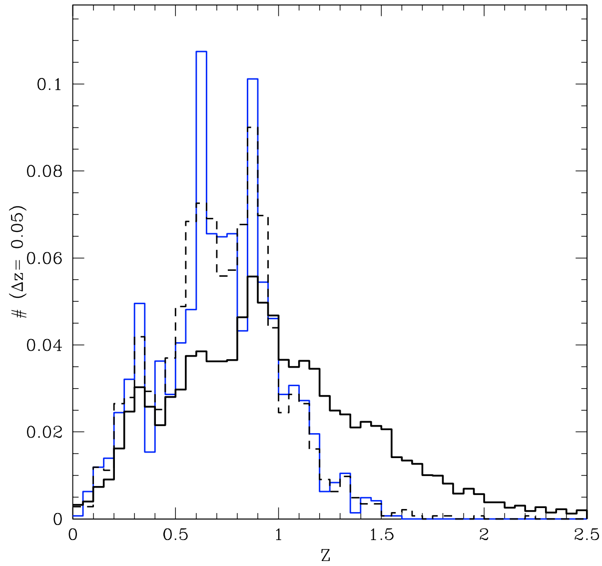


Fig. 4. The redshift distributions for the 3.6 μm sample with $F(3.6\ \mu\text{m}) \geq 9\ \mu$Jy, for the full photo-$z$ sample (solid thick line), the spectroscopic sub-sample (solid thin line) and the photo-z of spectroscopic objects (dashed line). All three curves are normalized to unity.

are shown as solid lines for different K band absolute magnitudes ranging from $K_{\rm ABS} = -25$ to -21. While an optically selected sample with $I \leq 24$, can detect ellipticals brighter than $K_{\rm ABS} \geq -23.0$ up to $z \sim 1.2$, a 3.6 μm selected sample with $m(3.6\ \mu\text{m}) \leq 21.5$ can detect them up to $z \sim 2$ if very deep optical photometry is available. The present sample is therefore well adapted to investigate the evolution of old and/or massive galaxies up to $z \sim 2$, providing a major step with respect to previous analyses based on optical selection.

3. Photometric redshifts

3.1. The method

We measure the photometric redshifts and we classify the whole population in galaxy/quasar/star based on the χ^2 fitting analysis of the spectral energy distributions, using the photometric redshift code "Le Phare"[1]. In this work, we adopt a similar procedure to the one described by Ilbert et al. (2006b). We use empirical templates based on the four observed spectral types from Coleman et al. (1980) and add two starburst templates from Kinney et al. (1996). Templates are extrapolated into ultraviolet and infra-red wavelengths using the GISSEL synthetic models (Bruzual & Charlot 2003) and we refine the set of SEDs with a linear interpolaton amongst the original SEDs. For Scd and later spectral types, we allow for different amounts of dust attenuation with a reddening excess $E(B-V)$ varying from 0 to 0.6 and an interstellar extinction law from Prevot et al. (2004). The spectroscopic sample is used to perform a template optimization and estimate possible systematic shifts amongst the different passbands as introduced by Ilbert et al. (2006b). The SED fitting is performed from the u^* to the 3.6 and 4.5 μm photometric passbands (the two latter passbands being still dominated by stellar light). All objects have at least 4 passbands with measured fluxes which ensures reliable SED fitting analysis.

In Fig. 3, we compare the photometric and spectroscopic redshifts for 1400 galaxies observed by the VVDS. We obtain an accuracy in the photo-z of $\sigma[\Delta z/(1+z)] \sim 0.031$ with no systematic shift and a small fraction of catastrophic errors (1.5% with $|\Delta z| \geq 0.15(1+z)$).

We consider as stars and QSOs (or AGN) the objects with a high SExtractor stellarity index parameter in i' band (CLASS_STAR ≥ 0.97) and for which either the star or the QSO template provide a best χ^2 fitting. After removing these stars and QSOs, we end up with a total sample of ~21 200 galaxies out of 25 500 sources.

3.2. The redshift distributions

The redshift distributions are shown in Fig. 4 for the whole sample with photometric redshifts (thick line), for the spectroscopic sub-sample ($I \leq 24$, thin line) and for the photo-z's of the spectroscopic objects (dashed line). The two latter distributions, both optically selected, are in excellent agreement, but do not show the high redshift tail (up to z = 2–2.5) observed for the 3.6 μm sample. We find that ~40% of galaxies with $F(3.6\ \mu\text{m}) \geq 9\ \mu$Jy are at $z \geq 1$. A similar fraction, ~35%, is obtained by Franceschini et al. (2006) for galaxies with $F(3.6\ \mu\text{m}) \geq 10\ \mu$Jy in the GOODS-CDFS field. Note however that both samples rely partially on photometric redshifts. Since the high-z tail ($z \geq 1.4$) cannot be verified by our spectroscopic sample, we have investigated the reliability of the redshifts for this high-z population through color–color diagnostics. In Fig. 5, we show the $(g'-z')$ vs. $(z'-36\ \mu\text{m})$ color–color diagram as an indicator for galaxies at $z \geq 1.4$. This is similar to the BzK diagnostic proposed by Daddi et al. (2004). The locus of galaxies with $z_{\rm phot} \geq 1.4$ is well separated from the low redshift population as illustrated by the VVDS spectroscopic sample. This was also reported by Daddi et al. (2004) for the spectroscopic follow-up of BzK candidates. We conclude that our MIR selection at 3.6 μm allows us to probe a significant population of galaxies in the redshift range $1 \leq z \leq 2$ with no indication of any bias from the lack of spectroscopic redshifts.

The spectroscopic redshift distribution also shows strong peaks at $z \sim 0.35$, 0.6 and 0.85. Two of them (z = 0.35 and 0.85) are still observed in the photo-z distribution although

[1] http://www.oamp.fr/people/arnouts/LE_PHARE.html

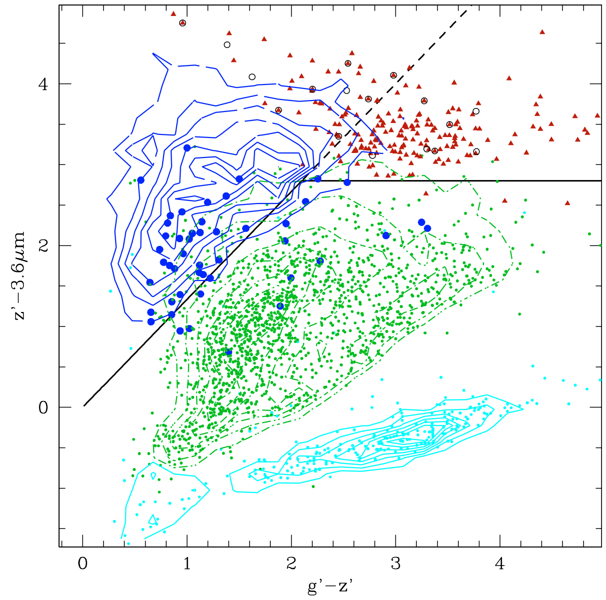


Fig. 5. Color–color diagram with ($z' - 3.6\ \mu$m) vs. ($g' - z'$) for 3.6μ sample. The filled dots show the spectroscopic objects while contours show the distributions of the photo-z sample for $z \leq 1.4$ (dashed-dotted green contours), $z \geq 1.4$ for the active/blue (long dashed blue contours) and stars (thin cyan contours). The quiescent/red populations with $z_{\rm phot} \geq 1.4$ are shown as triangles and the open circles show dusty galaxies observed at 24 μm (F(24 μm) $\geq$ 300 μJy) and falling above the $(NUV - r')_{\rm ABS}$ criteria described in the text. The lines show a typical separation between the three classes in a similar way as the BzK selection.

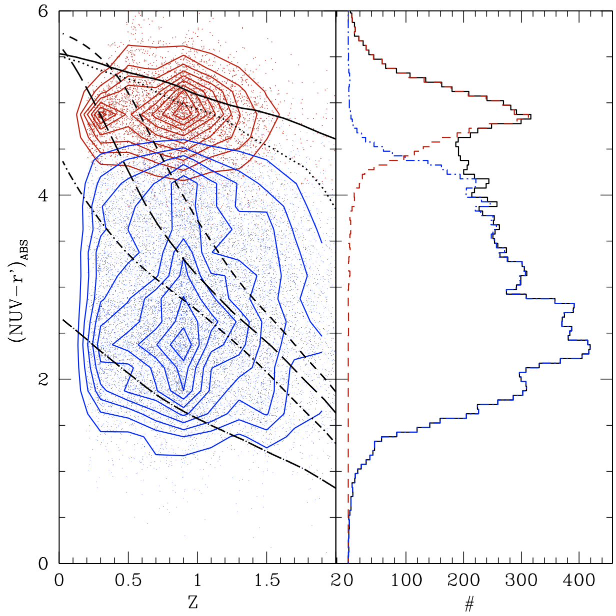


Fig. 6. *Left panel*: $(NUV - r')_{\rm ABS}$ color vs. redshift for the quiescent/red (red contours) and active/blue (blue contours) populations. Color evolution for models with $z_{\rm f} = 3, 6$ and $\tau = 0.1$ (dotted and solid lines resp.); $z_{\rm f} = 3$ and $\tau = 1, 1.5$ (short and long dashed lines); $\tau = 2, \infty$ (dot-short dashed and dot-long dashed lines resp). *Right panel*: histogram for the $(NUV - r')_{\rm ABS}$ color (solid line). The dashed lines show the distributions of the two populations as separated by SED templates (see text).

the area is two times larger and the distribution is somewhat smeared by the photo-z errors. These two peaks have been identified as large scale structures extending through the entire field (Marinoni et al. 2005; Adami et al. 2007) and we estimate the number density to be in excess by ∼20–25% in these structures (based on a smoothed redshift distribution derived by applying a moving window with a width of $\Delta z = 0.2$). Despite the large field of view, cosmic variance remains a significant source of uncertainty as previously pointed out by Bell et al. (2004) in the case of the COMBO17 survey. Following the recipes described by Somerville et al. (2004), we estimate the cosmic variance from large scale structure fluctuations in the ΛCDM paradigm. In the low and high redshift bins (with volumes of 2.7 and 12.5×10^5 Mpc^3 respectively), we expect the cosmic variance to produce density variations of 30% and 10% respectively, which appears to be in good agreement with our own estimates. We decide to account for this effect as an additional source of uncertainties in the measurements presented hereafter.

3.3. Estimates of absolute magnitudes

Throughout this paper we use absolute luminosity quantities. In practice, to derive the absolute magnitudes we use an adaptive filter method as introduced and tested on simulations by Ilbert et al. (2005, Appendix A). To reduce the uncertainties in the k-corrections used to derive the absolute magnitude at $\lambda_{\rm rest}$, we choose the apparent magnitude in the filter ($\lambda_{\rm obs}$) that best matches the redshifted $\lambda_{\rm rest}$. In the case where $\lambda_{\rm obs} = \lambda_{\rm rest} * (1+z)$, the k-correction is reduced to its redshift component ($2.5 * \log(1+z)$) with no template dependency.

All absolute magnitudes are derived according to this scheme except for the specific case of the K absolute magnitude ($K_{\rm ABS}$). We always use use the 3.6 μm magnitudes for $z \geq 0.4$. This choice yields null k-correction at $z \sim 0.6$ and requires small k-correction in highest redshifts with little template dependency. When available, at $z \leq 0.4$, we use the observed K band.

In the next section we use the NUV and r' absolute luminosities. Over the redshift range $0.3 \leq z \leq 1.75$ the rest-frame NUV (∼2350 Å) and r' passbands correspond to observed wavelengths of $0.3\ \mu\text{m} \leq \lambda \leq 0.65\ \mu$m and $0.8\ \mu\text{m} \leq \lambda \leq 1.7\ \mu$m respectively. We therefore need only small extrapolations for the NUV absolute magnitudes in the lowest z-bin. For the r' filter, the near infrared photometry is required at high redshift, which is available for a large fraction of our sample. We note however, that even without near-infrared photometry the shape of the SED is always well constrained in the NIR domain, because all our objects have per definition a measurement at 3.6 μm. We therefore obtain a good estimate of the r' absolute magnitudes even at high redshift.

We have derived the absolute magnitudes based on the empirical SEDs used to measure the photo-z. As a external check, we have also measured the absolute magnitudes based on a library of SEDs from PEGASE2 code (Fioc et al. 1997). Thanks to the use of the adaptive filter method, the results from the two libraries are fully consistent with an rms dispersion for the $(NUV - r')_{\rm ABS}$ color of $\sigma[\Delta(NUV - r')] \sim 0.2$. This is small compared to the large dynamical range of this color (see Fig. 6).

3.4. The selection of quiescent galaxies

We split our sample in two main classes in order to distinguish the red/quiescent and blue/active galaxies. To do so, we explore and compare two approaches, where one is based on the color bimodality and the other on an SED fitting procedure. For the color bimodality we consider the excellent separation provided by the $(NUV - r')_{\rm ABS}$ color (Fig. 6; right panel) as suggested

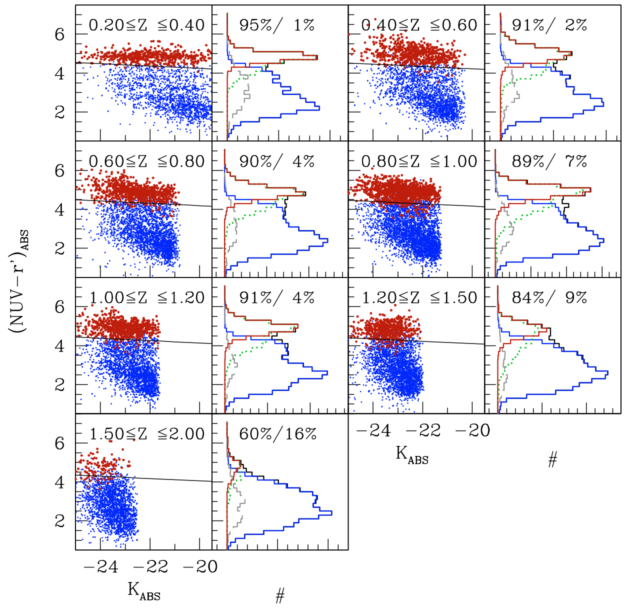


Fig. 7. $(NUV-r')_{\rm ABS}$ color vs. $K_{\rm ABS}$ distribution (left side) as a function of redshift (as specified in each panel). The two colored set of points correspond to the quiescent and active galaxies as separated by SED templates. The solid lines denote the valley between the two populations as seen in the projected histograms (right side). The green dotted histograms show the selection of red galaxies based on $U-B$ rest-frame color (see text). The dashed histograms show the galaxies detected at 24 μm ($F(24\ \mu{\rm m}) \geq 300\ \mu$Jy).

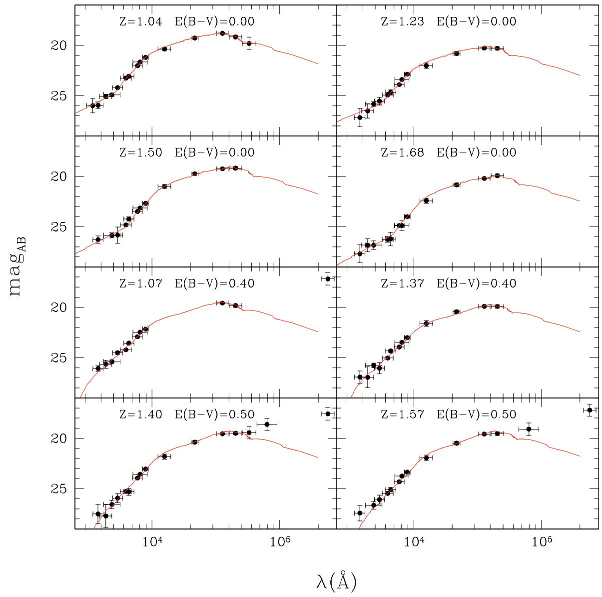


Fig. 8. Examples of SEDs for a sample of high-z galaxies with $(NUV-r')_{\rm ABS}$ above of the color criterion (see text). The fit is performed only in the wavelength domain 0.3 μm $\leq \lambda \leq$ 4.5 μm. The top four objects are best fitted by an old/passive elliptical, while the lower four galaxies are best fitted by dusty star-forming SEDs.

by recent GALEX results (Salim et al. 2005; Wyder et al. 2007). This is in contrast to the often used rest-frame $(U-V)_{\rm ABS}$ color suggested by Bell et al. (2004).

One of the advantages of the $(NUV-r')_{\rm ABS}$ color is that the NUV passband is sensitive to a stellar population with a light-weighted age of $\langle t \rangle \sim 10^8$ yr while the r' passband is probing $\langle t \rangle \geq 10^9$ yr (Martin et al. 2005), making the $(NUV-r')_{\rm ABS}$ color an excellent indicator of the current over past star formation activity. Indeed, the GALEX-SDSS sample shows that $(NUV-r')_{\rm ABS}$ is tightly correlated with the birthrate b parameter ($b = SFR(t < 10^8\ {\rm yr})/\langle SFR \rangle$) with $(NUV-r')_{\rm ABS} \geq 4.3$ being associated to galaxies with $b \leq 0.1$ and morphologically with de Vaucouleur profiles (Salim et al. 2005).

To illustrate the redshift evolution in $(NUV-r')_{\rm ABS}$, we use the stellar synthesis model PEGASE2, with smooth star formation histories, described by different e-folding times and formation redshifts (τ(Gyr), $z_{\rm f}$). In Fig. 6, left panel, we show the evolution of $(NUV-r')_{\rm ABS}$ vs. z for different SF histories (τ) and redshift of formation. Galaxies with $\tau \leq 0.1$ Gyr are in the red clump at all z, while models with higher τ move progressively into the red sequence at decreasing redshifts ($z \sim 0.7$ and 0.4 for $\tau = 1$ and 1.5 Gyr respectively). All models with $\tau \geq 2$ Gyr have ongoing star formation that prevents them from reaching the red sequence unless their star formation is quenched prematurely. Therefore the $(NUV-r')_{\rm ABS}$ color bimodality (Fig. 6, right panel) appears to be a good way to separate active and quiescent galaxies according to their birthrate parameter, with $b \leq 0.1$ for quiescent galaxies.

In Fig. 7, we show the galaxy distribution in the diagram $(NUV-r')_{\rm ABS}$ vs. $K_{\rm ABS}$ and the projected histograms for the different redshift bins. A strong bimodality is observed at almost all redshifts providing a natural way to split the sample. We define an adaptive color cut-off located in the valley and characterized by: $(NUV-r')_{\rm ABS} = -0.06 \times (K_{\rm ABS} + 22) + b(z)$, where $b(z) = 4.4{-}0.13 \times z$ (solid lines). However, some uncertainties remain on the nature of the population in the red sequence due to the contamination by dusty star-forming galaxies. A reddening excess of $E(B-V) = 0.2$ produces a color reddening of $\Delta(NUV-r)_{\rm ABS} = 0.94$ (or 1.04), assuming starburst (or SMC) extinction curve. We also show in Fig. 7, the locations of those galaxies that are most likely dusty star-forming galaxies, as indicated by their detection at 24 μm (with $F(24\ \mu{\rm m}) \geq 300\ \mu$Jy). Between 5 to 9% (depending on z) of this bright 24 μm population lies above our adopted colour cut and is a source of contamination for our quiescent population.

We then take a different approach, by classifying the galaxies according to the best-fit SED, on the basis of the six original templates (but allowing for dust extinction for the later types), in a similar way as Zucca et al. (2006). We consider as quiescent galaxies only those classified as elliptical and all other galaxies as active.

In Fig. 7, we compare the SED-based classification with the $(NUV-r')_{\rm ABS}$ color distribution (quiescent as red points; active as blue points). To quantify the overlap between the two different classifications, in each panel, we measure the fraction of ellipticals in the sample above and below the color cut previously defined. We find that the SED-ellipticals dominate (with ~90%) the red peak population up to $z \leq 1.5$ with only a small fraction falling below the cut. At $z \geq 1.5$, the fraction of ellipticals above the cut represents only 60% of the reddest objects while the remaining 40% are best classified as dusty star-forming galaxies. In Fig. 8, we show examples of the two types of SEDs (Ell or dusty star-forming) for galaxies above the $(NUV-r')$ cut and with $z_{\rm phot} \geq 1.0$.

In Fig. 5, we show the distribution of SED-elliptical galaxies with $z \geq 1.4$ (red triangles). As for the BzK selection with a spectroscopic validation (Daddi et al. 2004), they are reasonably well separated from the active galaxies with $z \geq 1.4$ (blue contours). We also show the bright 24 μm population with $(NUV-r')_{\rm ABS}$ in the red sequence and $z \geq 1.4$ (open circles). They are spread all over the upper part of the diagram. Amongst this population,

~50% of them has been correctly adjusted by dusty star-forming SEDs. This shows that SED fitting can get rid partially of the dusty galaxies falling in the red sequence. As a last test to quantify the residual contamination of the red/quiescent sample, we project our elliptical sample in the $(J - K)$ vs. $(r' - K)$ color diagram. This is susceptible to distinguish the early-types from the dusty starbursts (Mannuzzi & Pozzetti 2000; Daddi et al. 2004). We find that ~15% of the sample lies in the dusty starburst locus. If this color separation were to be efficient, this would mean that our quiescent sample may still be moderately contaminated by dusty galaxies.

In conclusion, we have analyzed two methods to select the red/quiescent galaxies based either on the $(NUV-r')_{\rm ABS}$ color or on the selection of quiescent galaxies according to SED fitting. We have shown that the SED-quiescent population represents the large majority of galaxies in the red peak of the bimodality distribution. We decide to adopt the SED-fitting classification because it allows to reduce the contamination of dusty starbursts in our red/quiescent sample. At $z \geq 1.4$, we estimate the residual contamination to be around ~15%. Only deeper Spitzer-MIPS observations could allow to better disentangle the two populations. However we note that the results discussed in the following sections are marginally affected by the chosen selection.

We end up with a sample of 4425 quiescent and 16 770 active galaxies.

4. The stellar mass to light ratio (M/L_K)

We now derive the evolution of the stellar mass to light ratio (M/L_K) with redshift for the active and quiescent populations. This will allow us to assess stellar mass quantities. Thanks to the large VVDS spectroscopic sample, we can measure how the M/L_K varies with redshift. In particular, the use of spectroscopic redshifts allows us to avoid mass uncertainties relative to photo-z and we thus obtain accurate estimates of the scatter for these relations.

We derive the stellar mass $M^\star$ for each galaxy by fitting the multi-band photometry. In the present work, we base ourselves on the Bruzual & Charlot (2003) stellar population model. We have used this model to construct a library of star formation histories that includes stochastic bursts (see Kauffmann et al. 2003; and Salim et al. 2005, for similar libraries). Although this library uses a Chabrier (2003) IMF, we have decided to adopt a classical Salpeter IMF throughout this paper. We therefore apply a systematic shift of +0.24 dex to the masses derived with a Chabrier IMF. In brief, our method uses the Bayesian probability distribution functions as described in Appendix A of Kauffmann et al. (2003) to derive parameter estimates for each observed SED, including errors on each derived parameters, in this case the stellar mass. A full description of our mass estimation routines will appear in forthcoming papers (Lamareille et al. 2007; Walcher et al. 2007). For now we refer to previous works for discussion on the accuracy of the stellar mass estimates by SED fitting methods, in particular the choice of star formation histories with or without episodic bursts (Borch et al. 2006; Pozzetti et al. 2007) and the addition of Mid-IR bands (Fontana et al. 2006). We note, however, that our stellar mass estimates may be affected by the established problems pertaining to the contribution of short-lived, luminous, infrared-bright stars in intermediate age population (~1 Gyr) (Maraston et al. 2006). Our mass estimates would possibly need to be changed by 0.1 to 0.2 dex (see Pozzetti et al. 2007 for an in-depth discussion). A final solution to this problem affecting all publicly available stellar population synthesis models is however not imminent.

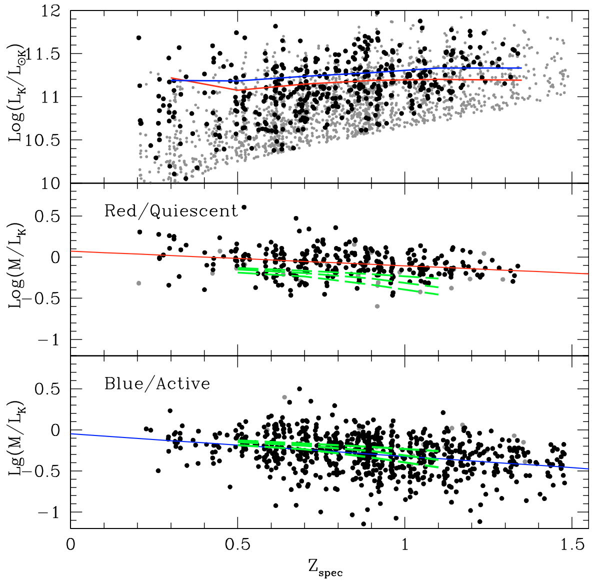


Fig. 9. *Upper panel*: distribution of the K luminosity ($L_K/L_{\odot K}$) for the spectroscopic sample (quiescent galaxies: black filled circles; active galaxies: grey small circles). The characteristic luminosity $L^\star$ for the Red/quiescent and Active/blue samples are shown as solid lines. *Lower panels*: behavior of the Mass to light ratio (M/L_K) for the quiescent (*middle panel*) and active (*lower panel*) samples with the best fit shown as solid lines. Comparison with results from Drory et al. (2004; long dashed lines) based on three different stellar mass cuts: $M/M_\odot \geq 2\times10^{11}$, $10^{11} \leq M/M_\odot \leq 2\times10^{11}$, and $4\times10^{10} \leq M/M_\odot \leq 10^{11}$.

In Fig. 9 (upper panel), we show the redshift distribution of the K luminosity ($L_K/L_{\odot K}$) for the quiescent (large black circles) and active (small grey circles) populations. The solid lines show the evolution of the characteristic luminosity, $L^\star$, derived below in Sect. 5.1 and reported in Table 1. The distribution of M/L_K as a function of redshift for the quiescent and active samples are shown in the middle and lower panels respectively. Before deriving the M/L_K versus redshift relations from the spectroscopic sample we have tested several subsets: the first one includes only objects brighter than $L_K \geq 10^{11}\ L_{\odot K}$ to define an unbiased luminosity sample up to $z = 1.5$. The second one considers objects with a luminosity between $0.4\,L^\star(z) \leq L_K(z) \leq 2.5\,L^\star(z)$, allowing to probe the evolution of the M/L_K for galaxies around $L^\star(z)$ which are the main contributors to the total luminosity density with 70%, 45% and 55% for the quiescent, active and total samples, respectively. Because the spectroscopic sample is optically selected ($I \leq 24$), we have also considered a sub-sample with $m3.6 \leq 20.5$ (which means one magnitude brighter than the full sample) to insure a fair sampling of the $(i' - 3.6\,\mu{\rm m})$ colors at all redshifts. All these subsamples are found to provide consistent results within the parameter's uncertainties, therefore we only report the values for the three samples with the selection centered around their respective $L^\star(z)$. For the three samples, we fit the redshift evolution of the M/L ratio by a power-law defined as $\log\langle M/L_K\rangle = a\times Z + b$, with a global rms given by σ.

For the global sample, based on 999 galaxies:

$$a = -0.30 \pm 0.03, b = 0.03 \pm 0.03, \sigma = 0.22. \quad (1)$$

For the active/blue sample, based on 753 galaxies:

$$a = -0.27 \pm 0.03, b = -0.05 \pm 0.03, \sigma = 0.21. \quad (2)$$

For the quiescent/red sample, based on 298 galaxies:

$$a = -0.18 \pm 0.04, b = +0.07 \pm 0.04, \sigma = 0.15. \quad (3)$$

Table 1. K-LF parameters (α, $M^\star$, $\Phi^\star$) and K luminosity densities (ρ_L) for the whole, active and quiescent samples. Errors in $M^\star$, $\Phi^\star$ refer to Poisson errors and slope uncertainties. Additional uncertainties due to photometric redshift (PdZ) and cosmic variance (CV) are reported separately. Errors in ρ_L include all of them.

$\langle z \rangle$	#	α	$M^\star$	$\Phi^\star$ (10^{-3} Mpc^{-3})	$(\frac{d\Phi}{\Phi})_{PdZ}$	$(\frac{d\Phi}{\Phi})_{CV}$	ρ_L ($10^8 L_{\odot K}$ Mpc^{-3})
			ALL				
0.3	2180	-1.1 ± 0.2	$-22.84^{+0.4}_{-0.6}$	$4.19^{+2.2}_{-1.9}$	0.05	0.23	6.97 ± 2.2
0.5	2680	-1.1 ± 0.2	-22.83 ± 0.3	3.50 ± 1.3	0.05	0.20	5.77 ± 1.3
0.7	3336	-1.1 ± 0.2	-22.96 ± 0.2	3.36 ± 1.1	0.05	0.16	6.24 ± 1.3
0.9	4545	-1.1 ± 0.2	-23.08 ± 0.2	4.22 ± 1.2	0.05	0.26	8.75 ± 2.5
1.1	3027	-1.1 ± 0.2	-23.22 ± 0.2	2.75 ± 0.7	0.11	0.13	6.48 ± 1.4
1.35	3077	-1.1 ± 0.2	-23.18 ± 0.17	2.56 ± 0.5	0.09	0.12	5.82 ± 1.5
1.75	2189	-1.1 ± 0.2	-23.42 ± 0.14	1.47 ± 0.2	0.07	0.10	4.17 ± 1.1
			active				
0.3	1666	-1.3 ± 0.2	$-22.83^{+0.43}_{-0.56}$	$2.38^{+1.6}_{-1.3}$	0.05	0.27	4.76 ± 1.7
0.5	2063	-1.3 ± 0.2	-22.82 ± 0.3	2.22 ± 1.0	0.05	0.21	4.40 ± 1.3
0.7	2569	-1.3 ± 0.2	-22.95 ± 0.2	2.21 ± 0.9	0.05	0.17	4.94 ± 1.3
0.9	3428	-1.3 ± 0.2	-23.06 ± 0.2	2.82 ± 1.0	0.05	0.28	6.97 ± 2.7
1.1	2358	-1.3 ± 0.2	-23.19 ± 0.2	1.99 ± 0.6	0.12	0.14	5.55 ± 1.8
1.35	2609	-1.3 ± 0.2	-23.19 ± 0.18	2.08 ± 0.5	0.09	0.13	5.80 ± 2.0
1.75	2081	-1.3 ± 0.2	-23.49 ± 0.15	1.27 ± 0.3	0.07	0.11	4.66 ± 1.8
			quiescent				
0.3	514	-0.6 ± 0.2	$-22.91^{+0.34}_{-0.49}$	$1.78^{+0.42}_{-0.5}$	0.09	0.33	2.63 ± 1.1
0.5	617	-0.3 ± 0.2	-22.55 ± 0.19	1.47 ± 0.17	0.07	0.24	1.60 ± 0.4
0.7	767	-0.3 ± 0.2	-22.73 ± 0.15	1.29 ± 0.13	0.09	0.21	1.65 ± 0.4
0.9	1117	-0.3 ± 0.2	-22.83 ± 0.14	1.58 ± 0.14	0.07	0.34	2.22 ± 0.8
1.1	669	0.0 ± 0.3	-22.86 ± 0.17	0.90 ± 0.08	0.12	0.17	1.42 ± 0.3
1.35	468	0.3 ± 0.3	-22.84 ± 0.16	0.41 ± 0.03	0.21	0.18	0.74 ± 0.2
1.75	108	0.6 ± 0.3	-23.27 ± 0.15	0.05 ± 0.02	0.28	0.17	0.17 ± 0.05

The fits to the quiescent and active samples are shown as solid lines in Fig. 9. All subsamples show a similar trend with a gradual decline of the $\langle M/L_K \rangle$ by a factor ~1.5 to 2 up to $z = 1$. Similar results were obtained by Drory et al. (2004) for samples selected with different mass limits, based on the NIR MUNICS survey, and are shown in Fig. 9 (long dashed lines) The behavior of the M/L is mainly driven by the stellar activity. The youngest and bluest stellar population have the smallest M/L_λ (Bell & de Jong 2001; Drory et al. 2004). Similarly, the decline of the M/L_λ with redshift reflects the well established decrease of the specific star formation rate (downsizing effect; Cowie et al. 1996: for a galaxy with same mass, the SF activity is higher in the past). The smaller decline of M/L_K with redshift observed for the quiescent sample reflects an older mean age of the stellar population, pushing the epoch of formation of the stellar component to high redshift. We find that the quiescent and active samples show a rather small scatter along this relation, with only 0.15 dex and 0.22 dex scatter respectively. The main origin for the scatter is the dependance of the M/L_K on the star formation history (SFH). The more complex SFH for the active population could be responsible for the larger scatter.

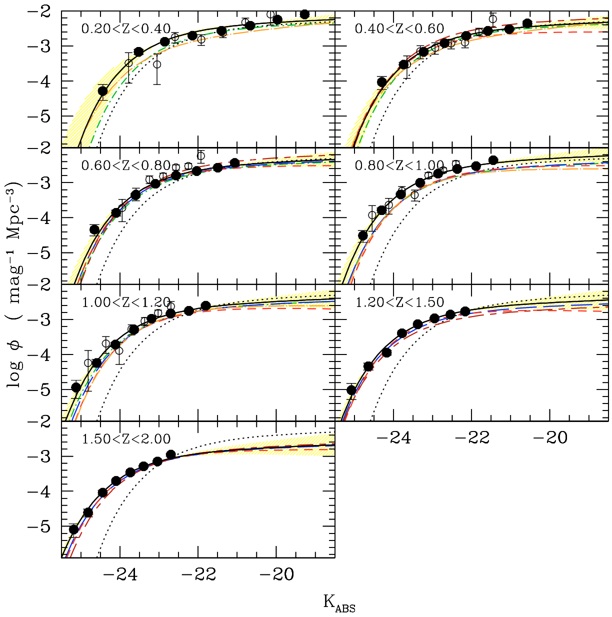


Fig. 10. K-band rest-frame luminosity functions, in different redshift slices, for the whole photo-z sample (V_{max}: filled dark circles, STY: solid lines and shaded area based on slope uncertainties) and the spectroscopic sample (V_{max}: open circles). Also shown are the local LF from Kochanek et al. (2001; dotted line) and high-z LFs from Drory et al. (2003; green dahed-dot line), Caputi et al. (2005; blue long dashed line), Pozzetti et al. (2004; orange long dashed-dot line), Saracco et al. (2006; dark red long-short dahed line, $\alpha = -1.1$), Cirasuolo et al. (2006; red dashed line).

5. The *K* band luminosity function and density

5.1. The K rest-frame luminosity function

5.1.1. K-LF measurements

We measure the K rest-frame luminosity function (LF) by adopting the V_{max} and STY estimators from the VVDS LF tool (ALF; Ilbert et al. 2004). In Fig. 10, we show the LFs for the whole and spectroscopic samples and in Fig. 11 for the red/quiescent and blue/active samples. To measure the LFs, a weight has been applied as a function of apparent magnitude to account for the incompleteness in number counts. In the case of the spectroscopic sample, we derive the LFs only for galaxies with $m_{AB}(3.6\mu) \leq 20.5$ and $0.2 \leq z \leq 1.2$. This restriction is due to the I band selection effect (see Fig. 2). An additional weight is applied to the

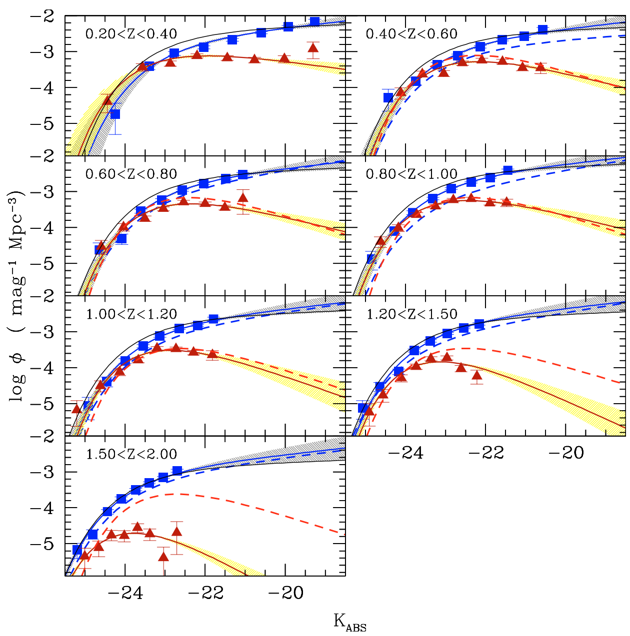


Fig. 11. K-rest luminosity functions for the blue/active (blue squares and solid lines) and the red/quiescent (red triangles and solid lines) samples. The thin black lines refer to the global sample from Fig. 10. We compare our LFs with the red and blue samples from the UKIDSS survey (Cirasuolo et al. 2006; red and blue dashed lines).

spectroscopic galaxies to take into account the sampling rate of the VVDS observations and its optical selection (as described in Ilbert et al. 2005).

In Table 1 we report the values of the STY parameters for the different samples. The errorbars on the parameters refer to the Poisson errors and slope uncertainties. For the global and blue samples, we have fixed the faint-end slopes of the STY estimator at low and high redshifts to the values observed between $0.6 \leq z \leq 1.0$, where we have the best constraints in the bright and faint-end simultaneously. This choice is consistent with local (Kochanek et al. 2001) and high redshift analysis (Drory et al. 2004; Caputi et al. 2005; Cirasuolo et al. 2006). For the red sample, we adopt a variable slope, although the low and high-z slopes are less well constrained. At low-z ($z \leq 0.4$), we use $\alpha = -0.6$, consistent with the slope observed by Bell et al. (2004) from a local elliptical sample. At $z \geq 1$, we adopt a gradual flattening of the slope that best reflects the evolution of the $V_{\rm max}$ estimator. However the depth of the current data does not allow any statement on the reality of this apparent flattening and deeper observations are required to confirm it. We note that the choice of the slope does not affect the discussion in next sections because changes to the estimates of the luminosity densities are marginal ($\rho_L(\alpha) - \rho_L(\alpha = -0.3) = -0.01, -0.02, -0.04$ dex for $\langle z \rangle = 1.1, 1.35, 1.75$, resp.).

5.1.2. Additional source of uncertainties

Additional sources of uncertainties on the LFs come from the photo-z and cosmic variance. We have estimated the uncertainties relative to the photo-zs from 100 mock samples based on the redshift probability distribution (PdZ) of each source (as in Bolzonella et al. 2000) and by recomputing the type separation and luminosity functions. Considering only the variations of $\Phi^\star$, we find uncertainties in the density to be between 5% to 12% for the total and active samples, depending on the redshift bin. For the quiescent sample, the uncertainties vary between 7% and 28%. This method allows to account simultaneously for the large photo-z uncertainties of the ellipticals at high-z and to estimate the stability of the quiescent vs. active separation at high redshift. We report these uncertainties in Table 1 ($(\mathrm{d}\Phi/\Phi)_{PdZ}$). Except for the quiescent sample at high z, the photo-z uncertainties are not a dominant source of errors in the present work.

It is more difficult to evaluate the effect of cosmic variance. As a first attempt we have measured the global LFs in 10 non overlapping fields covering the whole area and find variations of $\Phi^\star$ of between 0.07 dex to 0.2 dex. However, the fact that the 10 fields are located in a single area of the sky prevents us from using them as cosmic variance uncertainties over the whole survey (adopting a simple rescaling as $\sqrt{10}$). Therefore, as mentioned in Sect. 2, we adopt the formal approach discussed by Somerville et al. (2004). The cosmic variance is estimated as $\sigma = b\sigma_{\rm DM}$, where $\sigma_{\rm DM}$ is the variance of the dark matter that depends on the comoving volume (their Fig. 3, right panel) and b is the bias relative to each specific sample that depends on the comoving number density (their Fig. 3, left panel). Theses uncertainties are reported in Table 1 ($(\mathrm{d}\Phi/\Phi)_{\rm CV}$). Cosmic variance remains the dominant source of uncertainty for the whole and active samples at all redshift, while for the quiescent population the photo-z uncertainty is the major source of uncertainty at $z \geq 1.35$.

5.1.3. Evolution of the K-band LFs and comparison with other studies

In Fig. 10, we compare the global LFs with other studies. Beyond the fluctuations of $\Phi^\star$ due to the presence of large scale structures, our LFs appear slightly brighter than previous studies, by about 10%. This is however consistent with the expected calibration uncertainty of the SWIRE-IRAC flux (Surace et al. 2005). Comparing $K^\star$ to the local measurement by Kochanek et al. (2001), we observe a brightening of $\Delta K^\star \sim -0.8(-1.0) \pm 0.2$ from $z = 0$ to $z = 1.5(2)$. Regarding the comoving density parameter ($\Phi^\star$) we observe a decline by a factor 1.7 ± 0.2 up to $z = 1.2$ and 2.5 ± 0.5 up to $z = 2.0$ as compared to local estimate. This mild luminosity brightening and modest density decline agree with previous studies (Caputi et al. 2005; Saracco et al. 2006; Cirasuolo et al. 2006).

In Fig. 11, we compare the LF per types with our global LF and results from UKIDSS based on $(U - B)_{\rm rest}$ color selection (Cirasuolo et al. 2006). Our active sample shows a similar brightening than the global sample. The density, $\Phi^\star$, appears stable between $0.2 \leq z \leq 1.5$ and starts to decline at higher z. For the quiescent sample a different behaviour is observed. Up to $z = 1.2$, the density declines by a factor $2^{+1}_{-0.7}$, followed by a sharp drop, by a factor 16^{+9}_{-6}, between $1.2 \leq z \leq 2$. While the behaviour of the active and quiescent samples agrees with UKIDSS results on blue and red galaxies (concerning faint end slopes and global evolution), we obtain a different ratio between the blue/active and red/quiescent populations and they do not observe the strong decline for the red population at $z \geq 1.2$. We suggest two differences in the selection criterion for red/quiescent galaxies as possible source to this discrepancy:

- In Fig. 7, we show the $(NUV - r')$ color distribution (green dotted histograms) for galaxies selected on the basis of the optical color $(U - B)$, in a similar way as Cirasuolo et al. While the Cirasuolo selection includes our quiescent sample, it also includes galaxies with $(NUV - r)$ as blue as 2.5–3.

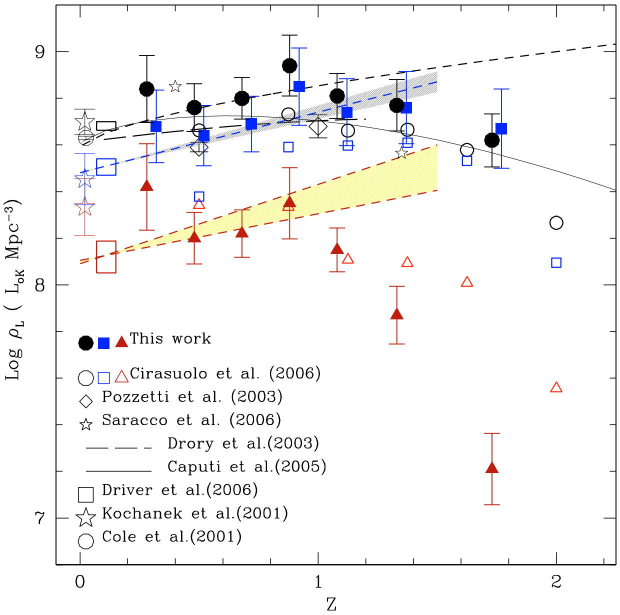


Fig. 12. Evolution of the luminosity density, ρ_{L_K} (expressed in $L_{\odot K}$), for the total (filled circles), active (filled squares) and quiescent (filled triangles) populations. Results are compared with previous works as specified in the caption and with a pure luminosity evolution models (dashed lines).

This means that the optical criterion yields a number of red galaxies that is larger by ~40 to 50% depending on redshift;

– the modest decline of the red LF at $z \geq 1.2$ for UKIDSS can also takes its origin in the color criterion and in the inclusion of the larger fraction of red dusty starbursts at high-z. Regarding our selection of a quiescent sample, we were conservative in the sense that we required that the galaxies do not reveal signs of star formation activities according to their $(NUV - r')$ color and by using multi-color SED fitting to exclude dusty starbursts. In our red population at $1.5 \leq z \leq 2$, only 60% of galaxies is considered as quiescent, the rest being dusty starbursts. We expect this effect to be even stronger in the case of an optical color selection.

The strong evolution between $z = 2$ and $z = 1.2$ of the LF parameters of the quiescent sample alone suggests that we are probing the epoch when an increasing number of galaxies stop their star formation activities and turn quiescent. However, the LF parameters are strongly correlated with each others and their interpretation shaky. We therefore now turn to discussing the evolution of the different samples via the measurements of integrated quantities (luminosity densities and stellar mass densities).

5.2. The K band luminosity density

A sensitive test for galaxy formation is the measurement of the luminosity density which provides the total amount of light emitted per unit of volume and is estimated as: $\rho_{L_K} = \int_0^\infty L\Phi(L)\mathrm{d}L = \Phi^\star L_K^\star \Gamma(\alpha + 2)$, where Γ is the gamma function and $(\Phi^\star, L^\star, \alpha)$ are the Schechter parameters listed in Table 1. The measurements of ρ_{L_K} (expressed in solar unit with $M_{K\odot}^{AB} = 5.14$) are listed in Table 1 and plotted in Fig. 12. The quoted errorbars include all the sources of errors discussed in previous sections.

The total sample shows a luminosity density increase by a factor ~2 (starting from local values) up to $z \sim 1{-}1.2$ followed by a fall off of a similar amount up to $z = 2$. This global behavior agrees, within the errorbars, with previous K band luminosity density studies and is well described by the fit proposed by Caputi et al. (2005, solid line). Looking at the SED selected samples, the active sample dominates the K band luminosity density at all z and follows a similar evolution as the global sample. For the quiescent sample, the luminosity density remains roughly flat up to $z \sim 1.2$ then drops significantly. As discussed in previous section, this trend is less pronounced in the UKIDSS's red sample measurements.

To interpret the evolution of the luminosity density for the different samples, we need to set local references. Since no similar K-band based selection exists, we make use of existing optical catalogs with information that can be converted to yield a K band local reference. We have used the results reported by Driver et al. (2006) based on the Millennium Galaxy Catalog (MGC) where they provide the LFs in B band and stellar mass density ($\rho_\star$) for a large variety of classification, based on morphology, colors, SED fitting and spectral classes (η parameter from 2dF: Madgwick et al. 2002). We adopt two criteria for the local reference that can reflect the definition of our quiescent sample:

– as noted by Madgwick et al. (2003), the η parameter is highly correlated with the birthrate b parameter. In particular their first class with $\eta \leq -1.4$ corresponds to $b \leq 0.1$ and is thus similar to what we expect locally for our criterion. We adopt the η_1 class as representative of our quiescent sample and define all other classes as belonging to the active sample as reported in Table 2 of Driver et al. (2006);
– the majority of galaxies in the $(NUV - r')$ red sequence observed by GALEX are dominated by de Vaucouleur profiles (Salim et al. 2005). We adopt the E/SO(red) morphological class of the MGC as reported in their Table 2 as another representation of our quiescent sample.

Finally, we converted $\rho_\star$ to ρ_{L_K} for each sample by adopting the Bell & de Jong (2001) relations between $(B - R)$ and M/L_K. The resulting ranges of local estimates are shown in the Fig. 12 as colored rectangles. We additionally plot the morphologically separated luminosity densities from 2MASS (Kochanek et al. 2001) as colored open stars. The low-z estimates agree reasonably well between themselves except for the quiescent sample from 2MASS. This is most likely due to the inclusion of the blue compact ellipticals in the sample of Kochanek et al. (2001). The MGC with its red E/S0 sample appears to be most suited to provide our local reference in this case.

Using the local normalization, we can now compare the evolution of ρ_{L_K} for pure luminosity evolution (PLE) models, i.e. with no merging involved. We normalize to the local density $\Phi^\star(z = 0)$ and use $\Delta \mathrm{Log}\rho_{L_K}(z) = -0.4\Delta K_{\mathrm{ABS}}(z)$. We adopt the luminosity evolution as derived in Sect. 5.1. Specifically, for the total sample we adopt the luminosity evolution from Caputi et al. (2005), for the active sample we adopt $\Delta K_{\mathrm{ABS}}(z) = -0.65(\pm 0.07) \times Z$, while for the quiescent sample we use two extrem PEGASE models that encircles the observed luminosity evolution and can be described by $\Delta K_{\mathrm{ABS}}(z) = -0.5| - 0.85 \times Z$. In Fig. 12, the PLE predictions are shown with dashed lines and shaded areas. The evolution of the rest-frame K-band luminosity density is consistent with the PLE model for the total and active samples up to $z = 1.2$, while the model overpredicts ρ_{L_K} at higher redshift. This shows that the number density of galaxies has to drop. For the quiescent sample, the PLE model fails

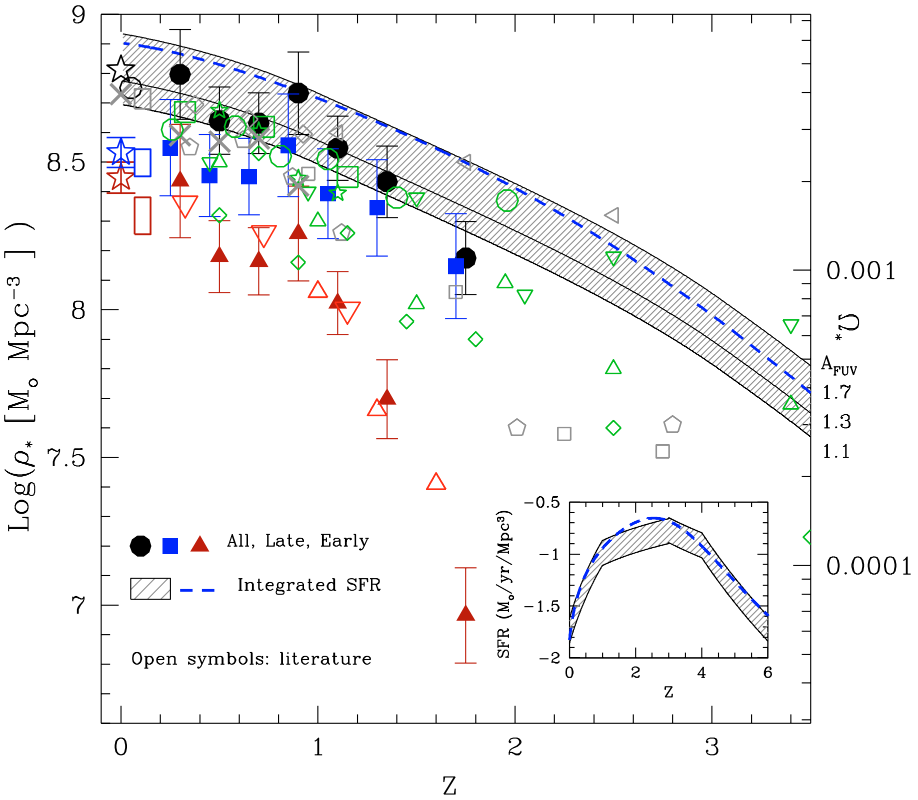


Fig. 13. Evolution of the stellar mass density as a function of cosmic time (assuming a Salpeter IMF). The total, active and quiescent stellar mass densities from this work are shown with large filled circles, blue squares and red triangles respectively. For reference, the right-hand axis gives the stellar density parameter. The integrated Star Formation Rates for different dust attenuation corrections ($A_{\rm FUV}$ = 1.1, 1.3, 1.7), based on the SFR derived by GALEX (Schiminovich et al. 2005), are plotted as solid lines, and the one from the compilation of Hopkins and Beacom as a dashed line. The dust corrected SFRs are shown in the inset. High z measurements from the literature for total samples have been splitted between analysis based on optical information only (grey symbols) and including Near or Mid IR data (green symbols) for the mass estimates. with Optical: Brinchmann & Ellis (2000; pentagons); Cohen et al. (2002; losanges); Dickinson et al. (2003; squares); Gwyn et al. (2005; inclined triangles); Borch et al. (2006; crosses). with NIR data: Drory et al. (2004; open stars); Drory et al. (2005; up and down triangles); Fontana et al. (2006; losanges); Franceschini et al. (2006; ellipticals: red down triangles and global samples: open squares); Pozzetti et al. (2007; open circles); Abraham et al. (2007; ellipticals: red up triangles). Local values are from Kochanek et al. (2001) for morphologically selected ellipticals (red star) and spirals (blue star) and whole sample (black star); Cole et al. (2001; circle) and Driver et al. (2006; colored rectangles). For clarity we do not show errorbars for other surveys.

to reproduce the observed trend at low and high redshifts. It predicts an increase in luminosity density by a factor 1.6 to 2.2, up to $z \sim 1.2$, while the observations suggest a modest increase by a factor ranging from 1.0 to 1.4. The disagreement with the PLE model at highest redshift is even more pronounced than for the active sample suggesting that the number density of the quiescent galaxies must drop even faster.

6. The stellar mass density up to z = 2

We now derive the stellar mass density, $\rho_\star$, up to z = 2. To that end we convert our K-band luminosity densities (Table 1), to stellar mass densities via the relation $\rho_\star(z) = \rho_{L_K}(z)\times\langle M/L_K\rangle(z)$ and using the mass to light ratio equations determined in Sect. 4.

Our measurements of $\rho_\star$ for the three samples are shown in Fig. 13. The errorbars account for Poisson, photo-z and cosmic variance uncertainties (as for ρ_L) and an additional uncertainty of 0.05 dex from M/L estimates (assuming that the paramaters a and b in the M/L relation are un-correlated).

6.1. The evolution of the stellar mass density

When moving back in time, the global population shows a small but regular decline up to $z \sim 1.1$ which accelerates at higher z. By comparing with local estimates, we find that the total stellar mass has decreased by roughly a factor of $\sim$1.5, 2 and 4 up to $z \sim$ 1.0, 1.5, 2.0. The compilation of previous surveys shows a large scatter by roughly a factor two and our estimates are located in the upper envelope. While the scatter can be in part due to mass estimates, for example the use or not of near IR data, the cosmic variance is most likely to be the dominant factor as discussed here and by Bell et al. (2003).

We quantify the stellar mass evolution of the active and quiescent samples with a simple linear fit with redshift up to z = 1.2. Including local measurements, we get: $\Delta{\rm Log}\rho_\star^{\rm Active} = -0.05(\pm0.09)z+8.51(\pm0.04)$ and $\Delta{\rm Log}\rho_\star^{\rm Quiescent} = -0.31(\pm0.07)z + 8.38(\pm0.02)$. The active population shows a modest evolution, consistent with no evolution with a mean value $\rho_\star^{\rm Active} = 10^{8.49\pm0.04}$. This constancy of the active (blue) sequence has been pointed out by Borch et al. (2006) and by Martin et al. (2007) who derived a similar value. On the other hand, we observe that the stellar mass of the quiescent galaxies, $\rho_\star^{\rm Quiescent}$, has increased by a factor 2 ± 0.3 between z = 1.2 and z = 0. Between z = 2 and z = 1.2 the evolution in stellar mass is even stronger, it increases by a factor of $\sim$10. This suggests two different regimes in the build-up of the quiescent population.

6.2. Comparison to the integrated star formation rate estimates

Thanks to the recent constraints on the cosmic star formation rate, we can compare the observed stellar mass density ($\rho_\star$) to the predictions for the cosmic star formation rate ($\dot{\rho}_\star$) over time, since the two quantities are simply related as follows:

$$\rho_\star(t) = (1-R)\int_0^t \dot{\rho}_\star(t)\mathrm{d}t \tag{4}$$

where R is the recycling factor (fraction of stellar mass released in the interstellar medium) which is assumed, in the case of a Salpeter IMF, to be $R = 0.3$ (Madau et al. 1998). For the SFR, $\dot{\rho}_\star(z)$, we use the uncorrected UV SFR as measured by Schiminovich et al. (2005) between $0 \le z \le 3$ that we have extended up to $z \sim 6$, based on measurements from Steidel et al. (1999), Bowens et al. (2006) and Bunker et al. (2004). To derive the dust corrected UV-SFR, we adopt a range of dust attenuation that varies between $1.1 \le A_{\rm FUV} \le 1.7$ which is consistent with local estimates (Buat et al. 2005) and with other dust-free SFR estimators up to $z \sim 1$ (Schiminovich et al. 2005; Hopkins & Beacom 2006). We also use the SFR estimate from the analysis of Hopkins & Beacom (2006) who compile most of the recent surveys from the UV to the Far-Infrared to derive a global, dust corrected SFR[2]. The two SFR histories are shown in an inset in Fig. 13 and are in good agreement within the range of dust attenuations we have considered. The result of integrating $\dot{\rho}_\star$ is shown as a shaded area. We emphasize that we did not impose beforehand that our integrated SFR be consistent with the stellar mass density at $z = 0$. Nevertheless, it provides a remarkable fit in the redshift range probed in this work and encloses well the local estimates (see also Rudnick et al. 2006).

We find that our measurements are well represented by an attenuation value ranging from $1.1 \le A_{\rm FUV} \le 1.3$) between $0 \le z \le 1.5$. This constancy of the mean dust attenuation up to $z \sim 1.5$, can be qualitatively interpreted as a trade-off between two competing factors: -1- star-forming galaxies at high-z have less dust attenuation than their local counterparts at the same total luminosity (Reddy et al. 2006; Xu et al. 2007) -2- star-forming galaxies at high-z are intrinsically more luminuous (Arnouts et al. 2005; Le Floch et al. 2006). We can quantify these effects by looking at the relations between the ($L_{\rm Dust}/L_{\rm UV}$) ratio and the total luminosity ($L_{\rm TOT} = L_{\rm Dust} + L_{\rm UV}$), all quantities expressed as νL_ν. From the local relation (Martin et al. 2005), a typical $L^\star_{\rm UV}(z = 0)$ galaxy has a $L_{\rm Dust}/L_{\rm UV} \sim 3.75$, corresponding to $A_{\rm FUV} \sim 1.3$ (Buat et al. 2005). If we let this galaxy evolve according to the evolution of the luminosity functions ($L_{\rm FUV} \sim (1+z)^{2.5}$; Arnouts et al. 2005 and $L_{\rm FIR} \sim (1+z)^{3.2}$; Le Floch et al. 2006), at $z \sim 1.35$ we obtain a typical luminosity of $\log(L_{\rm TOT}/L_\odot) \sim 11.34$. We can finally estimate the amount of dust attenuation for this galaxy based on the relation from Reddy et al. (2006, their Eq. (5)) that holds in the redshift range $1 \le z \le 3$. We obtain $L_{\rm Dust}/L_{\rm UV} \sim 4.3$, corresponding to $A_{\rm FUV} \sim 1.4$. This rough estimate shows that a typical $L^\star(z)$ galaxy that contributes to the bulk of the SFR at all z shows, on average, a small evolution of the amount of dust attenuation, in agreement with the observation of the integrated dust corrected SFR in Fig. 13.

To summarize, with simplified but not unrealistic estimates of the dust correction, we observe that the two complementary quantities, the star formation rate and the stellar mass density, are remarkably consistent with each other. This result is encouraging, and may provide some more fundamental informations about the initial mass function (IMF). The observations of the two quantities explore two different parts of the IMF (dominated by massive stars, a few $M_\odot$, for the SFR, and low mass stars, for stellar mass) and are extrapolated through the adopted IMF. The good agreement of these complementary observations over the last 10 Gyr, is a good support for an IMF that is on average universal, as already pointed out in previous work (e.g. Franceschini et al. 2006). We note that this remark is not specific to the Salpeter IMF adopted throughout this paper, but holds also for other IMFs suggested in the literature, as long as the quantities ($\rho_\star$, $\dot{\rho}_\star$) are appropriately corrected for IMF effects.

[2] We adopt their SFR based on the modified Salpeter IMF which is converted to a Salpeter IMF by applying +0.2 dex in stellar mass.

7. Discussion and conclusions

In this work, we have used a unique large sample of 21 200 galaxies selected at 3.6 μm (based on the VVDS and SWIRE surveys) to investigate the evolution of the luminosity functions, luminosity densities, stellar mass to light relations and stellar mass densities up to $z \sim 2$. We have separated the active and quiescent galaxies based on an SED fitting procedure. We define as quiescent those galaxies which are best fit by an elliptical template. We have shown that this sample reproduces well the red sequence of the color bimodality ($NUV - r'$) (known to be a good separator between active and quiescent galaxies with $b \le 0.1$) and at the same time minimizes the contamination by dusty starbursts. We find that the active and quiescent populations follow different behaviours. In particular, there is a clear transition between two regimes in the evolution of the quiescent population at a redshift of $z = 1.2$ (i.e. 8 Gyr ago).

7.1. The last 8 Gyr

The active population appears to be in place at $z \sim 1.2$ with a small evolution over this time laps except the aging of its stellar population. In contrast the quiescent sequence shows a gradual increase with a doubling of its stellar mass in line with previous optical studies (Bell et al. 2003; Faber et al. 2006; Brown et al. (2006).

The lack of evolution of the stellar mass density for the active, $\rho_\star^{\rm active}$, and the doubling for quiescent, $\rho_\star^{\rm quiescent}$, is surprising. The gradual increase of the global stellar mass density, over the last 8 Gyr can be easily explained by the formation of new stars by star forming galaxies at a rate described by the cosmic star formation rate. Or, similarly, it is found that intermediate mass galaxies ($M/M_\odot \sim 10.5$) have specific SFRs allowing them to increase their stellar mass by a factor two since $z = 1$ (Juneau et al. 2005). On the other hand, the increase of $\rho_\star^{\rm quiescent}$ is in apparent contradiction with the definition of quiescent galaxies that cannot produce more than 5% of additional stellar mass over the last 8 Gyr (due to the absence of star formation). The most plausible explanation for the increase of $\rho_\star^{\rm quiescent}$ is a progressive migration of galaxies from the active to the quiescent population at a similar rate than that with which new stars are formed.

We quantify this evolution by a rough estimate of the stellar mass growth per unit of time defined as:

$$\dot{\rho}_X = \frac{\rho_\star^X(z_l) - \rho_\star^X(z_h)}{t_{\rm univ}(z_l) - t_{\rm univ}(z_h)}, \tag{5}$$

where X refers to stellar mass of the considered sample in the redshift range $z_l \le z \le z_h$. We find that the stellar mass growth of the active population evolves very little with $\dot{\rho}_{\rm Active} = 0.005 \pm 0.003\ M_\odot/{\rm Mpc}^3/{\rm yr}$, while the quiescent population has a much higher stellar mass growth with $\dot{\rho}_{\rm Quiescent} = 0.017 \pm 0.004\ M_\odot/{\rm Mpc}^3/{\rm yr}$. We can compare this value with the mass growth expected by integrating the star formation history over the same period of time. We find that $\dot{\rho}_{SFR} = 0.025 \pm 0.003\ M_\odot/{\rm Mpc}^3/{\rm yr}$ (assuming $1.1 \leq A_{\rm FUV} \leq 1.3$). Under the assumption that the quiescent population has negligible residual star formation, its mass growth can be attributed to the mass flux of active galaxies moving into a quiescent mode ($\dot{\rho}_{A\to Q} = \dot{\rho}_{\rm Quiescent}$), and which appear to account for most of the global stellar mass growth derived from the SFR.

Additional evidence for this scenario is also presented by Vergani et al. (2007) who split the VVDS spectroscopic sample in active and quiescent galaxies based on 4000 Å break. In order to satisfy the constraints from the stellar mass functions split in active and quiescent galaxies, they find that the star formation in some massive blue galaxies must have been quenched, moving these galaxies into the "red sequence".

7.2. The major epoch of build-up for quiescent galaxies:

The present analysis allows us to extend previous work to the redshift range between $2 \geq z \geq 1.2$ (3.2 Gyr $\leq T_{\rm univ} \leq$ 5 Gyr). This appears to be a transition epoch, as evidenced by the strong increase in stellar mass, by a factor of ~10, of the quiescent population between $2 \geq z \geq 1.2$, while the active population increases still by a factor ~2.5. This translates into a stellar mass growth for the quiescent populations of $\sim 0.065\ M_\odot/{\rm Mpc}^3/{\rm yr}$, which is more than 3 times faster than the evolution at $z \leq 1.2$. In contrast to $z \leq 1.2$, where most of the galaxies seem to be in place, at $z \geq 1.2$, galaxies are still in an active phase of mass assembly. In particular, the evolution of the quiescent population suggests that we are observing the epoch when, for the first time in the history of the universe, a large number of active galaxies are ending their star formation activity and start to build up a quiescent population. While the mechanism acting in this process is not clear, gas exhaustion, merging or other effects, this build-up happens a few Gyr after the peak of the cosmic SFR (Hopkins & Beacom 2006). An interesting related information is the similar evolution followed by morphologically selected elliptical galaxies (Franceschini et al. 2006; Abraham et al. 2007). If not by chance, this coincidence could suggest that the build-up of the quiescent sequence is closely followed or preceded by a morphological transformation.

Acknowledgements. We are grateful to the referee for his careful reading of the manuscript and his suggestions. S.A. wants to thank B. Milliard for useful discussions. This work has been partially supported by the CNRS-INSU and the "Progamme National de Cosmologie" (France), the "Programme National Galaxies" (France) and a grant awarded for "EFIGI project" (grant # 45500) from the French research ministry. C.J.W. is supported by the MAGPOP Marie Curie EU Research and Training Network.

References

Abraham, R. G., McCarthy, P. J., Glazebrook, K., et al. 2007 [arXiv:astroph/0701779]
Arnouts, S. 2003, in Multiwavelength Cosmology, proceeding, Mykonos, June 2003
Arnouts, S., Schiminovich, D., Ilbert, O., et al. 2005, ApJ, 619, L43
Baldry, I. K., Glazebrook, K., Brinchmann, J., et al. 2004, ApJ, 600, 681
Bell, E., & de Jong, R. S. 2001, ApJ, 550, 212
Bell, E. F., McIntosh, D. H., Katz, N., & Weinberg, M. D. 2003, ApJ, 585, 117
Bell, E., Wolf, C., Meisenheimer, K., et al. 2004, ApJ, 608, 752
Bell, E. F., Papovich, C., Wolf, C., et al. 2005, ApJ, 625, 23
Benson, A. J., Bower, R. G., Frenk, C. S., et al. 2003, ApJ, 599, 38
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Bolzonnella, M., Miralles, J. M., & Pello, R. 2000, A&A, 363, 476
Borch, A., Meisenheimer, K., Bell, E., et al. 2006, A&A, 453, 869
Bouwens, R. J., Illingworth, G. D., Blakeslee, J. P., & Franx, M. 2006, ApJ, 653, 53
Brinchmann, J., & Ellis, R. S. 2000, ApJ, 536, L77
Brown, M. J. I., Dey, A., Jannuzi, B. T., et al. 2007, ApJ, 654, 858
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Buat, V., Iglesias-Paramo, J., Seibert, M., et al. 2005, ApJ, 619, L51
Bundy, K., Ellis, R. S., & Conselice, C. J. 2005, ApJ, 625, 621
Bundy, K., Ellis, R. S., Conselice, C. J., et al. 2006, ApJ, 651, 120
Bunker, A. J., Stanway, E. R., Ellis, R. S., & Mc Mahon, R. G. 2004, MNRAS, 355, 374
Caputi, K. I., Dunlop, J. S., Mc Lure, R. J., & Roche, N. D. 2005, MNRAS, 361, 607
Chabrier, G. 2003, PASP, 115, 763
Cimatti, A., Mignoli, M., Daddi, E., et al. 2002, A&A, 392, 395
Cimatti, A., Daddi, E., Renzini, A., et al. 2004, Nature, 430, 184
Cimatti, A., Daddi, E., & Renzini, A. 2006, A&A, 453, 29
Cirasuolo, M., McLure, R. J., Dunlop, J. S., et al. 2006 [arXiv:astro-ph/0609287]
Cohen, J. G. 2002, ApJ, 567, 672
Cole, S., Norberg, P., & Baugh, C. M. 2001, MNRAS, 326, 255
Cowie, L. L., Songaila, A., Hu, E., & Cohen, J. G. 1996, AJ, 112, 839
Croton, D. J., Springel, V., White, S. D. M., et al. 2006, MNRAS, 365, 11
Daddi, E., Cimatti, E., Broadhurst, T., et al. 2002, A&A, 384, 1
Daddi, E., Cimatti, E., Renzini, A., et al. 2004, ApJ, 617, 746
de Lucia, G., Springel, V., White, S. D. M., et al. 2006, MNRAS, 366, 499
Dickinson, M., Papovich, C., Ferguson, H. C., et al. 2003, ApJ, 587, 25
Driver, S. P., Allen, P. D., Graham, A. W., et al. 2006, MNRAS, 368, 414
Drory, N., Bender, R., Feulner, G., et al. 2003, ApJ, 595, 698
Drory, N., Bender, R., Feulner, G., et al. 2004, ApJ, 608, 742
Drory, N., Salvato, M., Gabash, A., et al. 2005, ApJ, 619, L131
Faber, S. M., Willmer, C. N. A., Wolf, C., et al. 2005 [arXiv:astro-ph/0506044]
Fazio, G. G., Hora, J. L., & Allen, L. E. 2004, ApJS, 154, 39
Fioc, M., & Rocca-Volmerange, B. 1997, A&A, 326, 950
Fontana, A., Donnarumma, I., Vanzella, E., et al. 2003, ApJ, 594, 9
Fontana, A., Pozzetti, L., Donnarumma, I., et al. 2004, A&A, 424, 23
Fontana, A., Salimbeni, S., Grazian, A., et al. 2006, A&A, 459, 745
Franceschini, A., Rodighiero, G., & Cassata, P. 2006, A&A, 453, 397
Franzetti, P., Scodeggio, M., Garilli, B., et al. 2007, A&A, 465, 711
Gwyn, S. D. J., & Hartwick, F. D. A. 2005, AJ, 130, 1337
Heavens, A., Panter, B., Jimenez, R., & Dunlop, J. 2004, Nature, 428, 625
Heinis, S., Milliard, B., Arnouts, S., et al. 2007, ApJ, accepted [arXiv:0706.1076]
Hopkins, A., & Beacom, J. F. 2006, ApJ, 651, 142
Ilbert, O., Tresse, L., Arnouts, S., et al. 2004, MNRAS, 351, 541
Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006a, A&A, 457, 841
Ilbert, O., Lauger, S., Tresse, L., et al. 2006b, A&A, 453, 809
Juneau, S., Glazebrook, K., Crampton, D., et al. 2005, ApJ, 619, L135
Kauffmann, G., Heckman, T. M., White, S. D. M., et al. 2003, MNRAS, 341, 33
Kochanek, C. S., Pahre, M. A., & Falco, E. E. 2001, ApJ, 560, 566
Lawrence, A., Warren, S. J., Almaini, O., et al. 2006 [arXiv:astro-ph/0604426]
Le Fèvre, O., Vettolani, G., & Garilli, B. 2005, 439, 845
Lilly, S. J., Le Fèvre, O., Hammer, F., & Crampton, D. 1996, ApJ, 460, L1
Lonsdale, C. J., Smith, H. E., Rowan-Robinson, M., et al. 2003, PASP, 115, 897
Madau, P., Pozzetti, L., & Dickinson, M. 1998, ApJ, 498, 106
Madgwick, D. S., Lahav, O., Baldry, I. K., et al. 2002, MNRAS, 333, 133
Madgwick, D. S., Somerville, R., Lahav, O., & Ellis, R. 2003, MNRAS, 343, 871
Maraston, C., Daddi, E., Renzini, A., et al. 2006, ApJ, 652, 85
Marinoni, C., Le Fèvre, O., Meneux, B., et al. 2005, A&A, 442, 801
Martin, C. D., Fanson, J., Schiminovich, D., et al. 2005, ApJ, 619, L1
Martin, C. D., Wyder, T. K., Schiminovich, D., et al. 2007, ApJ, accepted [arXiv:astro-ph/0703281]
Mo, H. J., & White, S. D. M. 1996, MNRAS, 282, 347
Nagamine, K., Cen, R., & Ostriker, J. P. 2001, Birth and Evolution of the Universe, 383
Pozzetti, et al. 2007, A&A, submitted
Pozzetti, L., & Mannucci, F. 2000, MNRAS, 317, L17
Reddy, N. A., Steidel, C. C., Fadda, D., et al. 2006, ApJ, 644, 792
Rudnick, G., Labbé, I., Forster Schreiber, N. M., et al. 2006, ApJ, 650, 624
Saracco, P., Fiano, A., Chincarini, G., et al. 2006, MNRAS, 367, 349
Salim, S., Charlot, S., Rich, M. R., et al. 2005, ApJ, 619, L39
Schiminovich, D., Ilbert, O., Arnouts, S., et al. 2005, ApJ, 619, L47

Somerville, R., Lee, K., Ferguson, H. C., et al. 2004, ApJ, 600, 171
Springel, V., Frenk, C. S., & White, S. D. M. 2006, Nature, 440, 1137
Steidel, C. C., Adelberger, K. L., & Giavalisco, M. 1999, ApJ, 519, 1
Surace, J. A., Shupe, D. L., Fang, F., et al. 2005, AAS, 207.6301
Takeuchi, T. T., Buat, V., & Burgarella, D. 2005, A&A, 440, L17
Treu, T., Ellis, R. S., Liao, T. X., et al. 2005, ApJ, 633, 174
van Dokkum, P. G. 2005, AJ, 130, 2647
van Dokkum, P. G., & Stanford, S. A. 2003, ApJ, 585, 78
Warren, S. J., Hambly, N. C., Dye, S., et al. 2007, MNRAS, 375, 213
Wyder, T., Martin, C. D., & Schiminovich, D. 2007, ApJ, accepted [arXiv:0706.3938]
Xu, C. K., Shupe, D., Buat, V., et al. 2007 [arXiv:astro-ph/0701737]
Yoshikawa, K., Taruya, A., Jing, Y. P., & Suto, Y. 2001, ApJ, 558, 520
Zucca, E., Ilbert, O., Bardelli, S., et al. 2006, A&A, 455, 879

[1] Canada-France-Hawaii Telescope Corporation, Kamuela, HI-96743, USA
[2] Laboratoire d'Astrophysique de Marseille, UMR 6110, BP 8, 13376 Marseille Cedex 12, France
e-mail: arnouts@cfht.hawaii.edu
[3] INAF-Osservatorio Astronomico di Bologna, via Ranzani 1, 40127 Bologna, Italy
[4] Institute for Astronomy, 2680 Woodlawn Dr., University of Hawaii, Honolulu 96822, Hawaii
[5] Institut d'Astrophysique de Paris, UMR 7095, 98 bis Bvd Arago, 75014 Paris, France
[6] Laboratoire d'Astrophysique de Toulouse, UMR 5572, 14 av. E. Belin, 31400 Toulouse, France
[7] INAF-Osservatorio Astronomico di Brera, via Brera 28, Milan, Italy
[8] University of California, San Diego 9500 Gilman Dr. La Jolla, CA 92093-0424, USA
[9] Spitzer Science Center, Mail Stop 314-6, 1200 East California Boulevard, Pasadena, CA 91125, USA
[10] IASF-INAF, via Bassini 15, 20133 Milano, Italy
[11] INAF-Osservatorio Astronomico di Capodimonte, via Moiariello 16, 80131 Napoli, Italy
[12] Geneva Observatory, ch. des Maillettes 51, 1290 Sauverny, Switzerland
[13] Centre de Physique Théorique, UMR 6207 CNRS-Université de Provence, 13288 Marseille, France
[14] School of Physics & Astronomy, University of Nottingham, University Park, Nottingham, NG72RD, UK
[15] Astrophysical Institute Potsdam, An der Sternwarte 16, 14482 Potsdam, Germany

4.3.2 Article 2 : Encoding of the infrared excess in the NUVrK color diagram for star-forming galaxies

A&A 558, A67 (2013)
DOI: 10.1051/0004-6361/201321768

Astronomy & Astrophysics

Encoding of the infrared excess in the *NUVrK* color diagram for star-forming galaxies⋆

S. Arnouts[1,4], E. Le Floc'h[2], J. Chevallard[3], B. D. Johnson[3], O. Ilbert[4], M. Treyer[4], H. Aussel[2], P. Capak[5], D. B. Sanders[6], N. Scoville[7], H. J. McCracken[3], B. Milliard[4], L. Pozzetti[8], and M. Salvato[9,10]

[1] Canada-France-Hawaii Telescope Corporation, 65-1238 Mamalahoa Hwy, Kamuela, HI 96743, USA
e-mail: arnouts@cfht.hawaii.edu
[2] Laboratoire AIM, CEA/DSM-CNRS-Université Paris Diderot, IRFU/Service d'Astrophysique, Bât. 709, CEA-Saclay, 91191 Gif- sur-Yvette Cedex, France
[3] UPMC-CNRS, UMR7095, Institut d'Astrophysique de Paris, 75014 Paris, France
[4] Aix-Marseille Université, CNRS, Laboratoire d'Astrophysique de Marseille, UMR 7326, 13388 Marseille, France
[5] Spitzer Science Center, MS 314-6, California Institute of Technology, Pasadena, CA 91125, USA
[6] Institute for Astronomy, 2680 Woodlawn Dr., University of Hawaii, Honolulu, HI 96822, USA
[7] California Institute of Technology, MC 249-17, 1200 East California Boulevard, Pasadena, CA 91125, USA
[8] INAF – Osservatorio Astronomico di Bologna, via Ranzani 1, 40127 Bologna, Italy
[9] MPE, Giessenbachstrasse 1, 85748 Garching, Germany
[10] Excellence Cluster, Boltzmann Strasse 2, 85748 Garching, Germany

Received 25 April 2013 / Accepted 25 July 2013

ABSTRACT

We present an empirical method of assessing the star formation rate (SFR) of star-forming galaxies based on their locations in the rest-frame color–color diagram ($NUV - r$) vs. ($r - K$). By using the *Spitzer* 24 μm sample in the COSMOS field (∼16 400 galaxies with $0.2 \leq z \leq 1.3$) and a local GALEX-SDSS-SWIRE sample (∼700 galaxies with $z \leq 0.2$), we show that the mean infrared excess $\langle \mathrm{IRX} \rangle = \langle L_{\mathrm{IR}}/L_{\mathrm{UV}} \rangle$ can be described by a single vector, ***NRK***, that combines the two colors. The calibration between ⟨IRX⟩ and NRK allows us to recover the IR luminosity, L_{IR}, with an accuracy of $\sigma \sim 0.21$ for the COSMOS sample and 0.27 dex for the local one. The SFRs derived with this method agree with the ones based on the observed (UV+IR) luminosities and on the spectral energy distribution (SED) fitting for the vast majority (∼85%) of the star-forming population. Thanks to a library of model galaxy SEDs with realistic prescriptions for the star formation history, we show that we need to include a two-component dust model (i.e., birth clouds and diffuse ISM) and a full distribution of galaxy inclinations in order to reproduce the behavior of the ⟨IRX⟩ stripes in the NUVrK diagram. In conclusion, the NRK method, based only on the rest-frame UV/optical colors available in most of the extragalactic fields, offers a simple alternative of assessing the SFR of star-forming galaxies in the absence of far-IR or spectral diagnostic observations.

Key words. infrared: galaxies – ultraviolet: galaxies – galaxies: evolution

1. Introduction

Star formation activity is a key observable for understanding the physical processes in the build up of galaxies. The SFR in galaxies depends on the physics of star-forming regions, merger history, gas infall and outflows and on stellar and AGN feedbacks. The SFR distribution and its evolution therefore offer a crucial test for any model of galaxy evolution. There are different indicators of the ongoing star formation (e.g., Kennicutt 1998), such as ultraviolet continuum ($\lambda \sim 912$–3000 Å) produced by massive, young stars ($t \sim 10^8$ yr, Martin et al. 2005); nebular recombination lines from gas ionized by the hot radiation from early-type stars ($\lambda \leq 912$ Å, $t \sim 10^7$ yr); far-infrared (FIR) emission from dust heated by UV light; and non-thermal radio emission, such as synchrotron radiation from supernova remnants (see Kennicutt 1983, 1998 for their respective calibrations with SFR).

Recently, measurements of SFRs in large samples of galaxies based on the above indicators have provided interesting insights in the dominant mode of baryon accretion for star-forming galaxies: the tight correlation, with small scatter, between stellar mass and SFR observed up to $z \sim 2$ (e.g., Salim et al. 2007; Noeske et al. 2007b; Daddi et al. 2007; Wuyts et al. 2012) suggests that a secular, smooth process, such as gas accretion, rather than merger-induced starbursts, may be the dominant mechanism governing star formation in galaxies. Also, the relation between specific SFR and stellar mass (e.g., Brinchmann et al. 2004; Salim et al. 2007; Noeske et al. 2007a) reveals that less massive galaxies have their onset of star formation occurring later than more massive ones (e.g., Cowie et al. 1996), and that a simple mass-dependent gas exhaustion model can reproduce the observed decline of the cosmic SFR since $z \sim 1$–1.5 (e.g., Lilly et al. 1996; Schiminovich et al. 2005; Le Floc'h et al. 2005). Such a scenario is also consistent with recent evidence of a higher fraction of molecular gas in massive star-forming galaxies at $z \geq 1$ with respect to nearby galaxies (e.g., Erb et al. 2006; Daddi et al. 2008, 2010; Tacconi et al. 2010).

Extending such measurements to high redshift for large samples of galaxies poses several challenges: optical recombination lines are often too weak and are shifted to near-IR wavelengths, where current spectroscopic capabilities are limited.

⋆ Appendices are available in electronic form at http://www.aanda.org

Far-IR indicators are also of limited use at high redshift, since the modest sensitivity and resolution of infrared telescopes make these observations sparse and restricted to the most massive galaxies, unless stacking techniques are used (e.g., Zheng et al. 2007; Pannella et al. 2009; Karim et al. 2011; Lee et al. 2010; Heinis et al. 2013). On the other hand, with increasing redshift, the ultraviolet continuum emission is progressively shifted to optical and near-IR bands, making it easily accessible with the largest ground-based telescopes (e.g., Burgarella et al. 2007; Daddi et al. 2007; Reddy et al. 2006; Steidel et al. 1996; Bouwens et al. 2011).

In the absence of dust, the ultraviolet luminosity of a galaxy is proportional to the mean SFR on a timescale $t \sim 10^8$ yr (Donas & Deharveng 1984; Leitherer & Heckman 1995). The presence of dust in the interstellar medium of galaxies hampers such a direct estimate of the SFR from UV observations. Starlight, especially in the UV, is efficiently absorbed and scattered by dust grains, which heat up and re-emit the absorbed energy at FIR wavelengths. This means that a fraction of UV photons will not escape the galaxy, and that neglecting this effect will lead to severe underestimates of the "true" SFR (e.g., Salim et al. 2007). Calzetti (1997) and Meurer et al. (1999) proposed a way to correct the UV continuum for dust attenuation that relies on the existence of a tight correlation between the slope of the UV continuum in the region $1300 \leq \lambda/\text{Å} \leq 2600$ (i.e., β-slope) and the ratio between the total IR ($8 \leq \lambda/\mu\text{m} \leq 1000$) and ultraviolet luminosities (i.e., infrared excess, $IRX \sim L_{\text{IR}}/L_{\text{UV}}$). This relation allows the estimation of the total IR luminosity from the shape of the ultraviolet spectral region, providing a way to quantify the amount of attenuation of UV starlight (i.e., A_{UV}) reprocessed by dust. It has been abundantly used to derive the SFR of high redshift galaxies, for which the UV slope is the only accessible quantity (Reddy et al. 2006; Smit et al. 2012).

To account for departures of quiescent star-forming galaxies from the star-burst IRX-β relation, several authors (e.g., Kong et al. 2004; Seibert et al. 2005; Salim et al. 2007) proposed a modified version of this relation, expressed in terms of $A_{\text{FUV}} - \beta$, which leads to a smaller dust correction at a given slope β, though with a significant scatter ($\sigma(A_{\text{FUV}}) \sim 0.9$). It is worth mentioning that such a large scatter is not surprising in view of the steepness of the $A_{\text{FUV}} - \beta$ relation, where a small uncertainty in UV color (e.g., $\Delta(\text{FUV} - \text{NUV}) \sim 0.1$) produces a large variation in A_{FUV} ($\Delta A_{\text{FUV}} \sim 0.8$). Johnson et al. (2007a) proposed another method calibrated for quiescent star-forming galaxies which relies on the combination of the D_{4000} break, a spectral feature sensitive to the age of a galaxy, and a long baseline color. They explored different colors and functions and obtain the smallest residuals by using the (NUV − 3.6 μm) color ($\sigma(IRX) \sim 0.3$ for their star-forming population). Treyer et al. (2007) compared the SDSS SFR estimates of Brinchmann et al. (2004), based on nebular recombination lines, with those obtained from the UV continuum corrected for the effect of dust attenuation with the methods of Seibert et al. (2005); Salim et al. (2007), and of Johnson et al. (2007a). They found that the method of Johnson et al. (2007a) leads to the smallest scatter [$\sigma(SFR) \sim 0.22$ vs. 0.33]. However, the method of Johnson et al. (2007a) depends on spectral diagnostics, such as the D_{4000} break, which are difficult to obtain for high redshift galaxies.

In this work, we analyze the behavior of the IRX within the rest-frame color–color diagram (NUV − r) vs. ($r - K$). We describe a new relationship between IRX and a single vector, called ***NRK***, defined as the combination of the colors (NUV − r) and ($r - K$). This new diagnostic provides an effective way to assess the total IR luminosity L_{IR} and then the SFR for individual galaxies with photometric information widely available in large surveys, and does not require complex modeling such as SED fitting techniques. The (NUV − r) vs. ($r - K$) color diagram adopted in this work is similar to the ($U - V$) vs. ($V - J$) diagram proposed by Williams et al. (2009) to separate passive or quiescent from star-forming galaxies, but better leverages the role of SFH and dust by extending into the extreme wavelengths of the SED. Total infrared luminosities were estimated from the deep MIPS-24 μm observations of the COSMOS field, with standard technics extrapolating mid-IR flux densities with luminosity-dependent IR galaxy SED templates as described in e.g., Le Floc'h et al. (2005). Our approach is motivated by the tight correlations that exist between mid-IR and total IR luminosities, both at low redshifts (Chary & Elbaz 2001; Takeuchi et al. 2003) and in the more distant Universe (Bavouzet et al. 2008). In fact, stacking analysis and direct individual identifications of distant sources in the far-IR with facilities like *Spitzer* and *Herschel* (e.g. Lee et al. 2010; Elbaz et al. 2010) have shown that these extrapolations from mid-IR photometry provide reliable estimates of total IR luminosities up to $z \sim 1.3$ (dispersion of ∼0.25 dex, no systematic offset), at least for star-forming galaxies initially selected at 24 μm. Also, given the relative depths of the different IR and submillimeter observations carried out in the COSMOS field with e.g., *Spitzer*, *Herschel* or JCMT, we note that the deep 24 μm data in COSMOS provide the largest sample of star-forming galaxies with measurable IR luminosities. We thus decide to limit our analysis to massive ($M \geq 10^{9.5}\,M_{\odot}$) star-forming sources first selected at 24 μm and lying at $z \leq 1.3$ so as to obtain reliable L_{IR} estimates. In a companion paper Le Floc'h et al. (in prep.), we will extend our analysis to higher redshifts and lower masses by stacking along the ***NRK*** vector the MIPS and *Herschel* data available in the COSMOS field.

The paper is organized as follows. In Sect. 2 we describe the dataset, the sample selection and the estimates of physical parameters. In Sect. 3 we discuss the behavior of the mean IRX (i.e., ⟨IRX⟩) in the (NUV − r) vs. ($r - K$) diagram and the calibration of the ⟨IRX⟩ vs. NRK relation. In Sect. 4, we compare our predicted IR luminosities with the ones based on the 24 μm observations and SED fitting technique. We also investigate the dependence with galaxy physical parameters and apply our method to the well established SFR-mass relation. In Sect. 5 we discuss the method, the possible origin of the relation with a complete library of model galaxy SEDs and dust models. We draw the conclusions in Sect. 6.

Throughout the paper we adopt the following cosmology: $H_0 = 70$ km s^{-1} Mpc^{-1} and $\Omega_{\text{M}} = 0.3$, $\Omega_{\Lambda} = 0.7$. We adopt the initial mass function of Chabrier (2003), truncated at 0.1 and 100 $M_{\odot}$. All magnitudes are given in the AB system (Oke 1974).

2. The dataset

2.1. Observations

2.1.1. The *Spitzer*-MIPS 24 μm observations and COSMOS catalog

The region of the sky covered by the *Cosmic Evolution Survey* (i.e., COSMOS Scoville et al. 2007) is observed by *Spitzer*-MIPS (Multi Imaging Photometer) at 24, 70 and 160 μm over a 2 deg^2 area (Sanders et al. 2007). In this work, we use the deep 24 μm observations and the catalogue of sources extracted by Le Floc'h et al. (2009). This catalog is 90% complete down to the flux limit density $S_{24} \sim 80$ μJy. The sources are cross-matched with the multi-wavelength photometric COSMOS catalog. This includes deep ultraviolet GALEX imaging (Zamojski et al. 2007), ground based optical observations with intermediate and broad band filters, near IR photometry (Capak et al. 2007; McCracken et al. 2010; Taniguchi et al. 2007) and deep IRAC photometry (Sanders et al. 2007; Ilbert et al. 2010) for a total of 31 filter passbands. The combined photometry, as well as accurate photometric redshifts, are available from the catalog of Ilbert et al. (2010, version 1.8). The typical depths at 5σ are 25.2, 26.5, 26.0, 26.0, 25.0, 23.5 and 24.0 in the NUV, u, g, r, i, K and 3.6 μm pass-bands, respectively. A vast majority of 24 μm sources is also detected in the optical bands (~95% with $u \leq 26.5$ and $i_{AB} \leq 24.5$).

We adopt spectroscopic redshifts from the bright and faint zCOSMOS sample (Lilly et al. 2009), when available, otherwise we use the photometric redshifts computed by Ilbert et al. (2009). The photo-z accuracy for the 24 μm sample is $\sigma = 0.009$ at $i_{AB} \leq 22.5$ (with 5700 z_{spec}) and 0.022 for the fainter sample (with 470 z_{spec})[1].

We limit our analysis to a maximum redshift $z \sim 1.3$ in order to derive reliable IR luminosities from the 24 μm flux density (see next section). We reject AGN dominated sources according to their mid-IR properties based on the diagnostic of Stern et al. (2007). The selected sample consist of ~16 500 star-forming galaxies and ~400 evolved/passive galaxies, with $S_{24} \geq 80$ μJy and $z \leq 1.3$. The separation between star-forming and passive galaxies is based on the position of the galaxies in the rest-frame (NUV $- r$) versus ($r - K$) diagram as discussed in Sect. 2.3 and Appendix B.

2.1.2. A low-z sample

We complement the COSMOS field with a lower redshift sample selected from Johnson et al. (2007a), which includes SWIRE observations (Lonsdale et al. 2003) in the Lockman Hole and the FLS regions. The multi-wavelength observations combine the GALEX, SDSS and 2Mass photometry. We restrict the sample to the sources detected in the three MIPS passbands (24, 70 and 160 μm). The final sample consists of ~730 galaxies with $z \leq 0.2$ (mean redshift: $\overline{z} \sim 0.11$) of which ~560 are star-forming galaxies (with the same separation criterion as above).

2.2. Rest-frame quantities and physical parameters

2.2.1. The infrared luminosity

The total Infrared luminosity (L_{IR}) refers to the luminosity integrated from 8 to 1000 μm and is derived by using the code "Le Phare"[2] (Arnouts et al. 1999; Ilbert et al. 2006) combined with infrared SED templates of Dale & Helou (2002). For the low-z sample, the IR luminosity is estimated by fitting the 8, 24, 70 and 160 μm bands with a free scaling of the SEDs (see Goto et al. 2011, for details). For the COSMOS sample, the L_{IR} is derived by extrapolating the 24 μm observed flux density with the star-forming galaxy templates of Dale & Helou (2002). While such SEDs are provided as a function of 60/100 μm luminosities ratio, we rescale the templates following the locally-observed dust temperature-luminosity relationship, so as to mimic a luminosity dependence similar to the SEDs of the libraries proposed by Chary & Elbaz (2001) and Lagache et al. (2004). In this scheme, the monochromatic luminosity at any given IR wavelength is linked to L_{IR} by a monotonic relation, similar to the correlations that have been observed between the mid-IR emission and the total IR luminosity of galaxies in the local Universe (e.g., Spinoglio et al. 1995; Chary & Elbaz 2001; Takeuchi et al. 2003; Treyer et al. 2010; Goto et al. 2011). In this way, a 24 μm flux density observed at a given redshift is associated to a unique L_{IR}. Note that we did not consider any AGN SEDs in this work, since AGN-dominated sources were removed from our initial sample (see Sect. 2.1.1).

Such extrapolations from mid-IR photometry have been widely used in the past, especially for the interpretation of the mid-IR deep field observations carried out with the infrared Space Observatory and the *Spitzer* Space Telescope. Given the large spectral energy distribution (SED) variations between the rest-frame emission probed at 24 μm and the peak of the IR SED where the bulk of the galaxy luminosity is produced, the estimate of L_{IR} with this method yet depends on the assumed SED library. To quantify this effect, we use different IR star-forming galaxy templates of the literature, finding systematic differences of only $\lesssim$0.2 dex up to $z \sim 1$ (e.g., see Fig. 7 of Le Floc'h et al. 2005). Furthermore, stacking analysis of mid-IR selected sources with MIPS-70 μm and MIPS-160 μm data in COSMOS (Lee et al. 2010) and other fields (e.g., Bavouzet et al. 2008) has allowed accurate determinations of average IR luminosities for galaxies stacked by bin of 24 μm flux. These studies have shown very good agreement with the 24 μm extrapolated luminosities of star-forming galaxies up to $z \sim 1.5$, thus confirming the robustness of the technic. Even more convincingly, direct far-IR measurements of the total IR luminosities for a 24 μm flux-limited sample of star-forming galaxies at $z \leq 1.5$ with the *Herschel* Space Observatory have recently revealed a tight correlation with the luminosities extrapolated from 24 μm photometry and star-forming SED templates (Elbaz et al. 2010). The comparison between the two estimates shows a dispersion less than 0.15 dex at $z \leq 1$, demonstrating again the relevance of the method at least for mid-IR selected sources at low and intermediate redshifts, such as in our current study. For these reasons we restrict our analysis to redshifts lower than $z \sim 1.3$, where the method discussed above to derive the L_{IR} is widely tested and robust.

2.2.2. The other physical parameters and luminosities

We use the code "Le Phare" combined with the population synthesis code of Bruzual & Charlot (2003, hereafter BC03) to derive the physical parameters for each galaxy. Details regarding the SED fitting are given in Appendix A. We perform a maximum likelihood analysis, assuming independent Gaussian distributed errors, including all the available photometric bands from 0.15 to 4.5 μm. The physical parameters are derived by considering the median value of the likelihood marginalized over each parameter, while errors correspond to the 68% credible region.

We adopt the same approach as Ilbert et al. (2006) to derive the rest-frame luminosities (or absolute magnitudes). We use the photometry in the nearest rest-frame broadband filter to minimize the dependency to the k-correction. These luminosities are consistent within 10% with those derived according to the best fit template (smallest χ^2) or those provided by Ilbert et al. (2009)

[1] The accuracy of the photometric redshifts is based on the normalized median absolute deviation (Hoaglin et al. 1983): $1.48 \times Median(|z_s - z_p|/(1+z_s))$, where z_s and z_p are the spectroscopic and photometric redshifts, respectively.

[2] http://www.cfht.hawaii.edu/~arnouts/lephare.html

based on a smaller set of empirical SEDs. Our results are not significantly affected by the adopted choices of luminosities[3].

We note also that the great accuracy of the COSMOS photometric redshifts in our redshift domain allows us to neglect the impact of the photo-z errors in the quantities discussed here. However, we adopt spectroscopic redshifts when available.

Throughout this paper, the stellar masses always refer to the estimates from the SED fitting technique, assuming a Chabrier IMF truncated at 0.1 and 100 $M_{\odot}$, while the total SFR is defined as the sum of the unobscured ultraviolet and total IR luminosities. For the latter, we adopt a relation similar to that proposed by Bell et al. (2005) and adjusted for a Chabrier (2003) IMF:

$$SFR^{\rm tot}[M_{\odot}\,{\rm yr}^{-1}] = 8.6\times10^{-11}\times(L_{\rm IR}+2.3\times\mathcal{L}_{\rm NUV}), \tag{1}$$

where the total IR luminosity is defined as $L_{\rm IR}/L_{\odot} \equiv \int_8^{1000} {\rm d}\lambda L(\lambda)$ and $\mathcal{L}_{\rm NUV}$ is the monochromatic NUV luminosity: $\mathcal{L}_{\rm NUV}/L_{\odot} = \nu L_{\nu}(2300\,{\rm Å})$. We adopt a factor of ~2.3 instead of 1.9 as proposed by Bell et al. (2005) who uses the FUV luminosity. According to stellar population models, our factor is more appropriated to correct the L(2300 Å) in total UV luminosity. The use of the FUV luminosity better traces the emission of short-lived, massive stars, while in this work we use the NUV luminosity. The reason is that we can obtain a more accurate rest-frame NUV than FUV luminosity, thanks to the GALEX NUV and CFHT u-band observations at different redshifts. As shown by Hao et al. (2011), this choice does not impact the reliability of the SFR estimates.

It is worth noting that, as shown in Appendix A, the instantaneous SFR derived from the SED fitting is consistent within less than a factor of two with the SFR estimated with Eq. (1). Considering that the SED fitting relies only on the 0.15–4.5 μm bands, this agreement between the different SFR estimates over at least 2 orders of magnitudes is remarkable.

Figure 1 shows the SFR (top panel; as estimated with Eq. (1)) and stellar mass (bottom panel) as a function of redshift for the 24 μm star-forming galaxies. As already shown by Le Floc'h et al. (2005), below $z \leq 0.5$ the population is composed of moderately star-forming galaxies with $SFR \leq 10\,M_{\odot}\,{\rm yr}^{-1}$. The fraction of luminous IR galaxies (LIRGs) gradually increases from $z = 0.5$ to $z = 1$ and becomes dominant at $z \gtrsim 1$ in our sample. At all redshift, the fraction of ultra luminous IR galaxies (ULIRGs) is negligible.

Our sample is dominated by galaxies with $\log(M/M_{\odot}) \gtrsim 9.5$. To characterize how representative the 24 μm population is with respect to the entire star-forming (hereafter SF) population at a given mass and redshift, we define a 50% completeness mass limit ($M_{50\%}(z)$) as the stellar mass where the ratio $\Phi^{24\,\mu}_{\rm SF}(M,z){\rm d}M/\Phi^{\rm All}_{\rm SF}(M,z){\rm d}M \sim 0.5$; where $\Phi^{24\,\mu}_{\rm SF}(M)$ and $\Phi^{\rm All}_{\rm SF}(M)$ are the $V_{\rm max}$ weighted comoving volume densities of 24 μm- and K-selected ($K \leq 23.5$) samples of star-forming galaxies respectively. Above this limit (~10^{10} [$10^{10.5}$] $M_{\odot}$ at $z \sim 1.1$ [1.3]), we consider the physical properties of the 24 μm population to be representative of the whole SF sample.

[3] We remind the readers that systematic shifts in photometric passbands can propagate into the absolute magnitudes and then in the Eqs. (2) and (3) of this work. A change in calibration of the IRAC 3.6 μm of ~0.1 mag will affect $K_{\rm ABS}$ by the same amount and the SFR by ~0.06 dex.

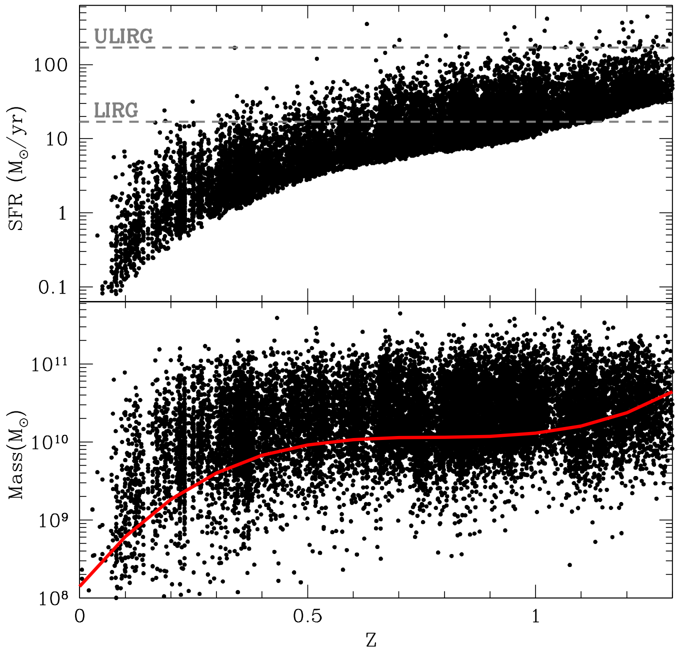


Fig. 1. Star formation rate (*top panel*) and stellar mass distribution (*bottom panel*) as function of redshift for the 24 μm-selected star-forming population. The horizontal gray lines in the *top panel* indicate the SFR thresholds corresponding to luminous and ultra-luminous infrared galaxies (i.e., LIRGs and ULIRGs). The solid red line in the *bottom panel* indicates the 50% completeness limit (i.e., $M_{50\%}(z)$, see Sect. 2.2.2).

2.3. The (NUV – r) versus (r – K) color–color diagram

Recently, Williams et al. (2009) have shown that quiescent and star-forming galaxies occupy two distinct regions in the rest-frame $(U-V)$ versus $(V-J)$ color–color diagram (i.e., UVJ) and validated their separation scheme with a morphological analysis (Patel et al. 2011). In the present work, we adopt the (NUV $-$ r) versus $(r-K)$ color–color (i.e., $NUVrK$) diagram to increase the wavelength leverage between current star formation activity and dust reddening. The $NUVrK$ diagram provides also an efficient way to separate passive and star-forming galaxies. In Fig. 2, we show the mean sSFR derived from the SED fitting for the entire COSMOS spectroscopic sample with $0.2 \leq z \leq 1.3$. The density contour levels (black solid lines) reveal the presence in the top left part of the diagram of a population with low specific SFR [$\log(sSFR^{\rm SED}/{\rm yr}^{-1}) \lesssim -10.5$], well separated from the rest of the sample. We define this region with the following criteria: (NUV $-$ r) > 3.75 and (NUV $-$ r) > 1.37 × $(r-K)$ + 3.2. In Appendix B we discuss further the star-forming and passive galaxy separation based on the morphological information from HST imaging of the COSMOS field.

In Fig. 3, we show the density distribution of the 24 μm sample in this $NUVrK$ diagram and the color tracks for five models with exponentially declining star formation histories. The colors indicate models with different e-folding times τ = 0.1, 1, 3, 5, 30 Gyr (red, purple, orange, green, blue lines respectively), while the filled and open squares mark different ages (t = 0.1, 0.5, 1, 3, 5 and t = 6.5, 9.0, 12 Gyr, respectively).

The (NUV $-$ r) is a good tracer of the specific SFR, since the NUV band is sensitive to recent (i.e., $t \leq 10^{8.5}$ yr) star formation and the r-band to old stellar populations (e.g., Salim et al. 2005). Models with short star formation timescales quickly move to the top side of the diagram, since star formation ceases early

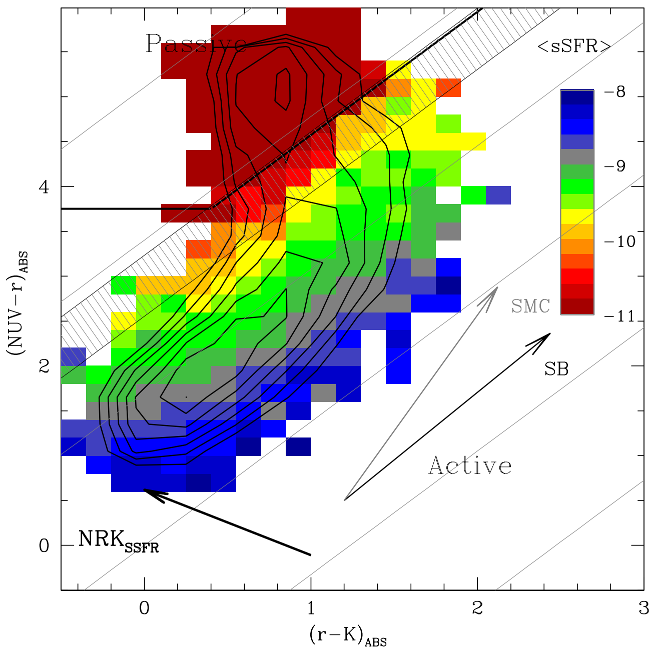


Fig. 2. Mean value of specific SFR (derived from the SED fitting, $s\mathrm{SFR}^{\mathrm{sed}}$ and color coded as indicated) in the *NUVrK* diagram for the entire COSMOS spectroscopic sample with $0.2 \le z \le 1.3$. The density contours (thin black lines) are logarithmically spaced by 0.2 dex. The heavy black lines delineate the region of passive galaxies. The shaded area defines the "intermediate" zone where the SFR estimates disagree between the different methods (UV+IR, NRK and SED fitting) as discussed in Sect. 4.2. We show the attenuation vectors for starburst (SB) and SMC attenuation curves assuming $E(B-V) = 0.4$ and the vector $\mathrm{NRK_{ssfr}}$ perpendicular to the starburst attenuation (see Sect. 4.3, *note the different dynamic ranges in x and y-axis warping the angles).*

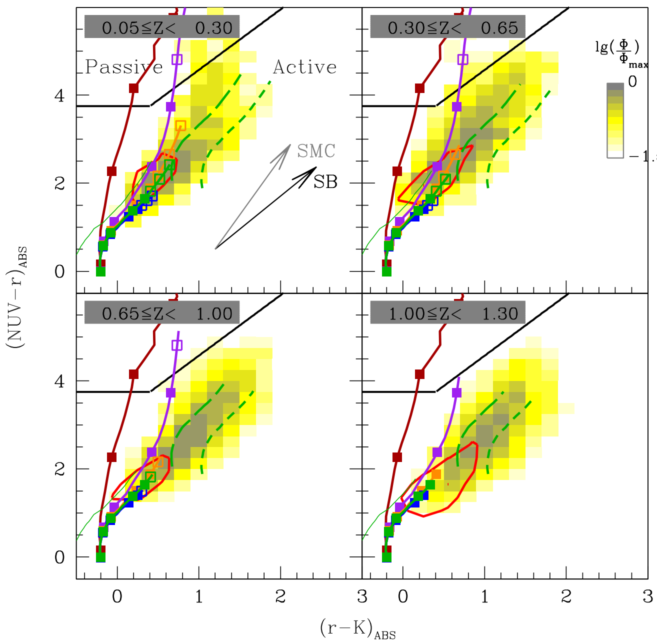


Fig. 3. Observed density distribution of the 24 μm sample in *NUVrK* diagram in four redshift bins, color coded (yellow-to-gray), in a logarithmic scale (step =0.15 dex) and normalized by the maximum density. For comparison we show as a solid red line the 50% contour density for the *K*-selected star-forming population ($K \le 23, 5$). We overplot tracks for BC03 models with different *e*-folding times ($\tau = 0.1$ Gyr, red line; 1 Gyr, purple; 3 Gyr, orange; 5 Gyr, green; 30 Gyr, blue). Symbols mark the model ages at $t = 0.1,\ 0.5,\ 1,\ 3,\ 5$ Gyr (filled squares) and $t = 6.5,\ 9.0,\ 12$ Gyr (open squares). For the model with $\tau = 5$ Gyr (thick green lines), we also show the impact of subsolar metallicity ($Z = 0.2\,Z_\odot$: thin line) and dust attenuation (SMC-like extinction law: long dashed line and Calzetti et al. 2000 law: short dashed line) assuming a reddening excess $E(B-V) = 0.4$. The thick black line in the *top-left* of each panel delineates the region of quiescent galaxies.

and the integrated light becomes dominated by old, red stars. On the other hand, models with longer *e*-folding times show bluer colors at all ages, since they experience a more continuous star formation which replenishes the galaxy with young, hot stars emitting in the UV. This behavior has been widely used to separate the active and passive galaxy populations based on $(\mathrm{NUV}-r)$ vs. stellar mass (or luminosity) diagrams (e.g., Martin et al. 2007; Salim et al. 2005). However, dust attenuation in star-forming galaxies can strongly alter the $(\mathrm{NUV}-r)$ color, producing variations up to several magnitudes [e.g., $\Delta(\mathrm{NUV}-R) \sim 2$ for $E(B-V) \sim 0.4$ and SMC law]. Following Williams et al. (2009) we use a second color $(r-K)$, which does not vary significantly with the underlying stellar population, even for passive galaxies [while $(\mathrm{NUV}-r)$ does], to partially break this degeneracy. As shown in Fig. 3, the models of spectral evolution span a much smaller range of $(r-K)$ color than that observed in the 24 μm sample, unless dust attenuation is included. We can qualitatively reproduce the observed ranges of $(\mathrm{NUV}-r)$ and $(r-K)$ colors by applying a reddening excess $E(B-V) \le 0.4-0.6$, assuming a starburst attenuation law or an SMC-like extinction curve.

3. Infrared excess of star-forming galaxies in the NUVrK diagram

3.1. Definition of the IRX and NRK parameters

We now explore the behavior of the infrared excess in the *NUVrK* diagram. We define the infrared excess as $IRX = L_{\mathrm{IR}}/\mathcal{L}_{\mathrm{NUV}}$, where L_{IR} and $\mathcal{L}_{\mathrm{NUV}}$ are defined in Sect. 2.2.2. This ratio is weakly dependent on the age of the stellar population, dust geometry and nature of the extinction law (Witt & Gordon 2000).

Figure 4 shows on the left, the volume weighted mean IRX ($\langle\mathrm{IRX}\rangle$, color-coded in a logarithmic scale shown in the top right panel), and, on the right, the dispersion around the mean ($\sigma(\mathrm{IRX})$) in the *NUVrK* diagram in four redshift bins. At any redshift, $\langle\mathrm{IRX}\rangle$ increases by ~1.5–2 dex from the blue side (bottom-left) to the red side (top-right) of the diagram, while the dispersion around the mean remains small and constant (≤0.3 dex). We note also the presence of stripes of constant $\langle\mathrm{IRX}\rangle$ values, which we discuss in Sect. 5. The presence of such stripes allows us to describe the variation of $\langle\mathrm{IRX}\rangle$ in the *NUVrK* diagram with a single vector perpendicular to those stripes. This vector, hereafter called ***NRK***, can be defined as a linear combination of rest-frame colors $\boldsymbol{NRK}(\phi) = \sin(\phi) \times (\mathrm{NUV}-r) + \cos(\phi) \times (r-K)$, where ϕ is an adjustable parameter which we require to be perpendicular to the $\langle\mathrm{IRX}\rangle$ stripes. We empirically estimate ϕ in each redshift bin ($\Delta z = 0.1$) by performing a linear least square fit: $IRX = f[NRK(\phi)]$ (the linear approximation is justified in the next section). Since the dispersion $\sigma[IRX(\phi)]$ reaches a minimum when the vector $\boldsymbol{NRK}(\phi)$ is perpendicular to the stripes, we minimize $\sigma[IRX(\phi)]$ to find the best-fit angle, ϕ_{b}. We find $15^\circ \le \phi_{\mathrm{b}} \le 25^\circ$ in all redshift bins, with a median and semi-quartile range $\phi_{\mathrm{b}} = 18^\circ \pm 4^\circ$. Given the narrow distribution of ϕ_{b} and the fact that a change up to $\Delta\phi \pm 7^\circ$ does not affect our

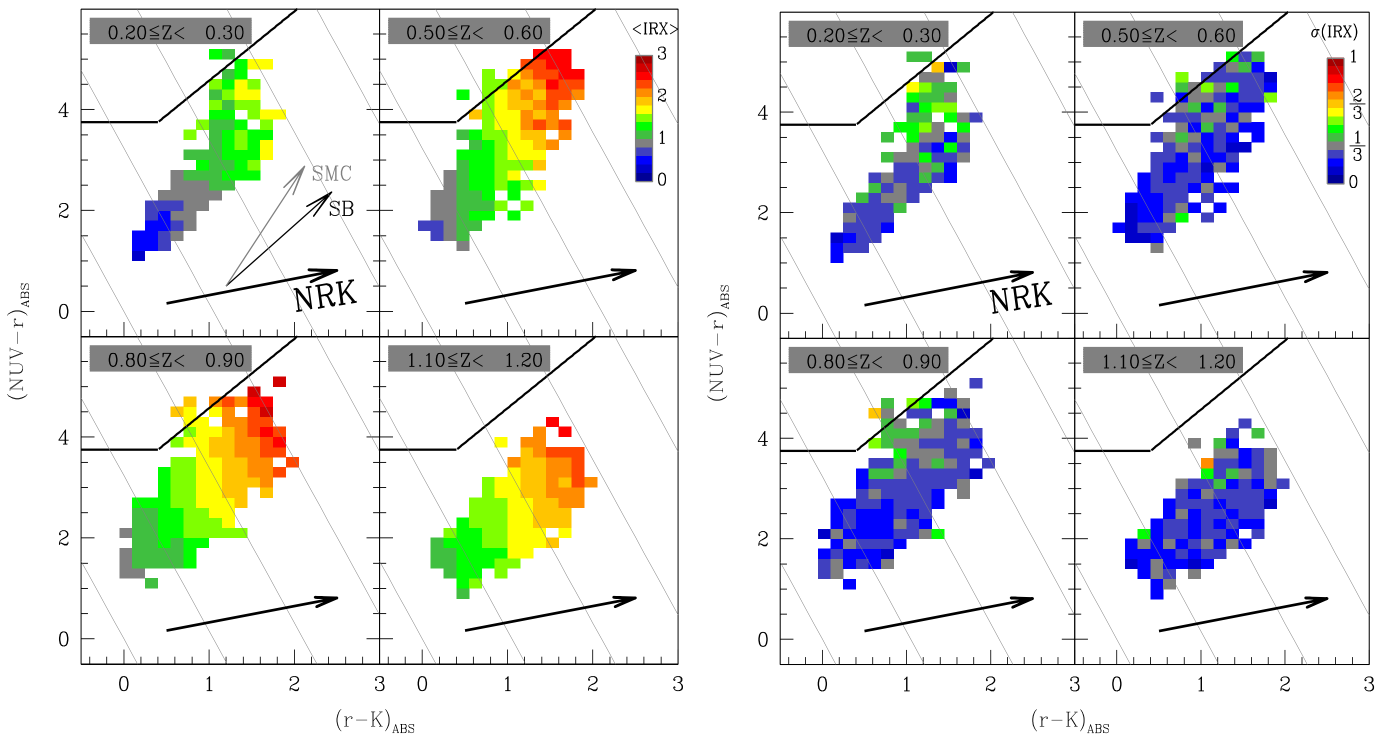


Fig. 4. Infrared excess (IRX) in the (NUV − r) versus (r − K) diagram. *Left figure*: volume-weighted mean IRX (⟨IRX⟩) for the 24 μm-selected sources in 4 redshift bins, color coded in a logarithmic scale (shown in the *top right panel*). We overplot in the *top-left panel* the attenuation vectors for starburst and SMC attenuation curves assuming $E(B-V) = 0.4$. In *each panel*, we overplot the vector ***NRK*** (black arrows) and its perpendicular lines (gray solid lines) corresponding to NRK in the range 0 to 4. (*note the different dynamic ranges in x and y-axis warping the angles*). *Right figure*: dispersion around the mean (σ(*IRX*)), color coded in a logarithmic scale.

results, we adopt a unique definition for the vector ***NRK*** at all redshifts by fixing $\phi_b = 18°$,

$$\boldsymbol{NRK} = 0.31 \times (\mathrm{NUV} - r) + 0.95 \times (r - K). \quad (2)$$

In Fig. 4 we show the vector ***NRK*** (black arrows) and a number of lines perpendicular to it (gray lines) corresponding to constant value of NRK in the range $0 \leq NRK \leq 4$. We note that ***NRK*** has a different orientation in the *NUVrK* diagram with respect to the starburst and SMC dust attenuation vectors, as we discuss in Sect. 5.

3.2. The calibration: ⟨IRX⟩ versus NRK

With the definition of ***NRK***, we can now derive the relationship between ⟨IRX⟩ and NRK. In Fig. 5 we measure the mean IRX values (⟨IRX⟩) and the associated dispersions per bin of NRK for the 24 μm sample (solid black circles) in different redshift bins. At high redshift, a tight correlation is observed with a small dispersion ($\sigma \leq 0.3$) compared to the evolution of ⟨IRX⟩ ($\Delta\langle IRX \rangle \sim 2$ dex). At $z \geq 0.4$, we obtain almost the same results for the spectroscopic sample (gray squares with yellow shaded region). Although this sample is 1/10 of the 24 μm sample, we observe the same dispersions, suggesting that we are dominated by an intrinsic, physical scatter rather than poisson noise.

In the first panel ($0.2 \leq z \leq 0.4$), we also include the lower redshift sample from Johnson et al. (2007a) ($\bar{z} \sim 0.11$, gray squares with yellow shaded region). While the ⟨IRX⟩ vs. NRK relation remains the same as at higher redshift, the dispersion increases in both of the low-z samples. This is related to the increasing fraction of less active galaxies with a higher contribution of evolved stars affecting the IR luminosity and contributing to blur the correlation between IRX and NRK (e.g., Cortese et al. 2008).

Finally we measure the ⟨IRX⟩ vs. NRK relation for three stellar mass bins. We observe no significant difference in the relation derived in each bin and the global one. This, along with the SED reconstruction analysis shown in Appendix C, supports our assumption of neglecting the dependence on stellar mass in the calibration of ⟨IRX⟩ vs. NRK in the mass range considered here.

In the parametrization of ⟨IRX⟩ as a function of NRK and redshift, we assume that the two quantities can be separated:

$$\log[\langle IRX \rangle(z, NRK)] = f(z) + a_N \times NRK \quad (3)$$

where $f(z)$ is a third-order polynomial function describing the redshift evolution, and a_N a constant which describes the evolution with NRK. We bin the data in redshift ($\Delta z = 0.1$) and NRK ($\Delta NRK = 0.5$) and perform a linear least square fit to derive the five free parameters. We obtain for the redshift evolution $f(z) = a_0 + a_1.z + a_2.z^2 + a_3.z^3$ with $a_0 = -0.69 \pm 0.06$; $a_1 = 3.43 \pm 0.33$; $a_2 = -3.49 \pm 0.55$; $a_3 = 1.22 \pm 0.28$, and for the dependence on NRK $a_N = 0.63 \pm 0.01$.

The resulting fits from this calibration are shown in Fig. 5 (solid red lines). We show, in the top panel of Fig. 6, the linear fit of ⟨IRX⟩ as a function of NRK after rescaling all the ⟨IRX⟩ values at $z = 0$ and in the bottom panel, the polynomial fit of ⟨IRX⟩ vs. z after rescaling all the ⟨IRX⟩ values at $NRK = 0$ The small uncertainty in the slope parameter a_N supports our initial choice of a linear function to describe the relation between ⟨IRX⟩

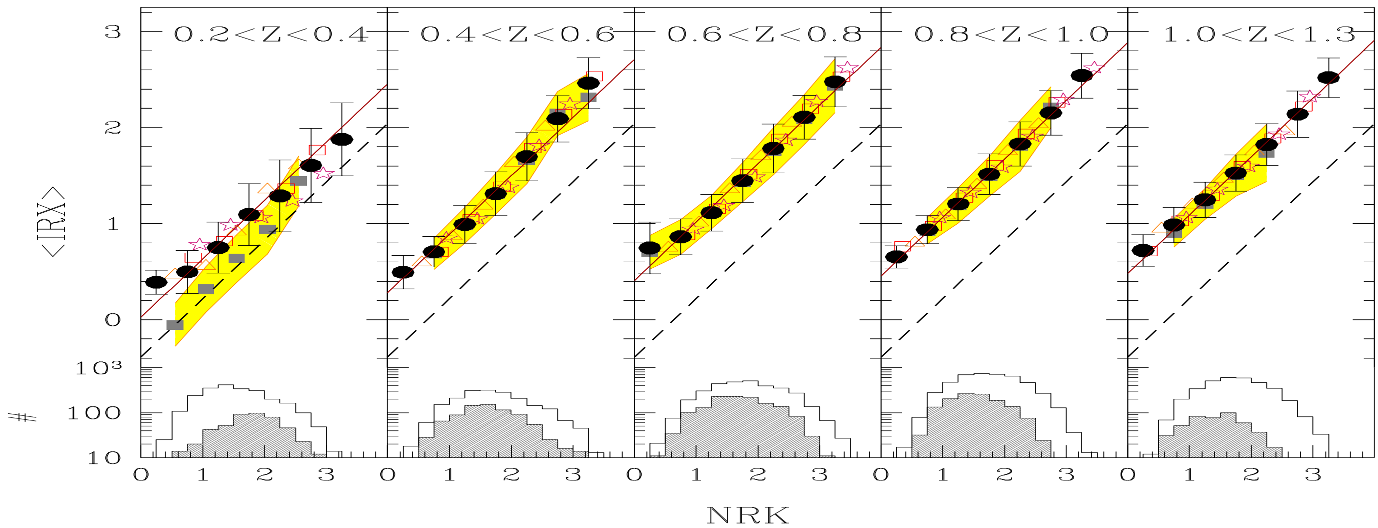


Fig. 5. Volume weighted mean IRX (⟨IRX⟩, on a logarithmic scale) as a function of NRK in five redshift bins. The COSMOS sample based on photometric redshifts is shown as solid black circles. gray squares and yellow shaded area indicate the mean IRX and dispersion for the spectroscopic sample of local galaxies of Johnson et al. (2007a) (*left-most panel*) and the COSMOS spectroscopic sample (*four right-most panels*). The solid red line indicates the predictions of Eq. (3) for ⟨IRX⟩ as a function of NRK at the mean redshift of the bin, while the dashed line refers to $\bar{z} \sim 0.11$ the mean redshift of the Johnson et al. (2007a) sample. We overplot the different mass selected samples: $9.5 \leq \log(M/M_{\odot}) \leq 10.0$ (open triangles); $10.0 \leq \log(M/M_{\odot}) \leq 10.5$ (open squares); $10.5 \leq \log(M/M_{\odot}) \leq 11.5$ (open stars). The distributions of NRK for the total (solid lines) and spectroscopic samples (shaded histograms) are shown in the bottom part of *each panel*.

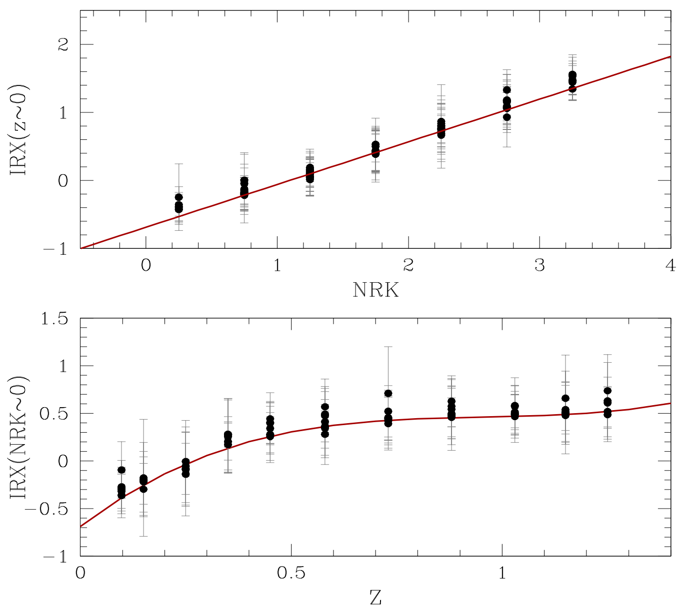


Fig. 6. Illustration of the analytical parametrization of the relation $\log[\langle IRX \rangle] = f(NRK, z)$. *Top panel*: linear fit as a function of NRK after rescaling ⟨IRX⟩ at $z = 0$. *Bottom panel*: polynomial fit as a function of redshift, after rescaling ⟨IRX⟩ at $NRK = 0$.

and NRK. We note that due to adoption of a polynomial function, the redshift evolution is only valid in the range well constrained by the data ($0.1 \leq z \leq 1.3$) and should not be extrapolated outside this range.

A simple interpretation of the redshift evolution of ⟨IRX⟩, at fixed NRK, is the aging of the stellar populations. In fact, Fig. 3 shows that, at a fixed position in the *NUVrK* diagram (or NRK value), a galaxy at higher redshift, which has a younger stellar population, needs a larger dust reddening than a galaxy at lower redshift, which hosts older, intrinsically redder stars. This effect being more pronounced between $0 \leq z \leq 0.5$ where the universe ages by $\Delta T \sim 5$ Gyr, compared to 4 Gyr in between $0.5 \leq z \leq 1.3$, and which is also enhanced by the global decline of the SF activity in galaxies with cosmic time.

4. The infrared luminosity and SFR estimated from *NRK* vector

The relationship between IRX and the vector ***NRK*** allows us to predict the IR luminosity for each galaxy according to its NUV, r, K luminosities and redshift as follows $L_{\rm IR}^{\rm NRK} = \mathcal{L}_{\rm NUV} \times \langle IRX \rangle(z, NRK)$, where NRK and ⟨IRX⟩ are estimated from Eqs. (2) and (3), respectively.

4.1. Comparison with the reference LIR

In Fig. 7 (left and middle panels) we compare the IR luminosity predicted with the NRK method ($L_{\rm IR}^{\rm NRK}$) with our reference IR luminosity derived from the mid/Far-IR bands ($L_{\rm IR}$, see Sect. 2.2.1). The IR luminosities estimated with the two methods for the ~16 500 star-forming galaxies in the COSMOS 24 μm sample (left panel) agrees with almost no bias and a dispersion of ~0.2 dex over the entire luminosity range. Less than 1% of the galaxies shows a difference larger than a factor of 3. The prediction of the IR luminosity for the LIRG population, which dominates the COSMOS sample (the 24 μm population peaks at $L_{\rm IR} \sim 2 \times 10^{11}\ L_{\odot}$), is excellent, considering the small number of parameters involved in the NRK method.

At low redshift, the prediction of the IR luminosity for the SWIRE galaxies (middle panel), which are 10 times less luminous than the 24 μm COSMOS population ($L_{\rm IR} \sim 2 \times 10^{10}\ L_{\odot}$), shows a larger scatter ($\sigma \sim 0.27$ dex). This reflects the larger dispersion observed in Fig. 5 which is likely related to the wider range of galaxy properties at low z. Johnson et al. (2007b) have also modeled the IRX as a function of different rest-frame colors and $D_n(4000)$ break for the SWIRE galaxies (see their Table 2 and Eq. (2)). The fit residuals for their predictions of IRX vary

A&A 558, A67 (2013)

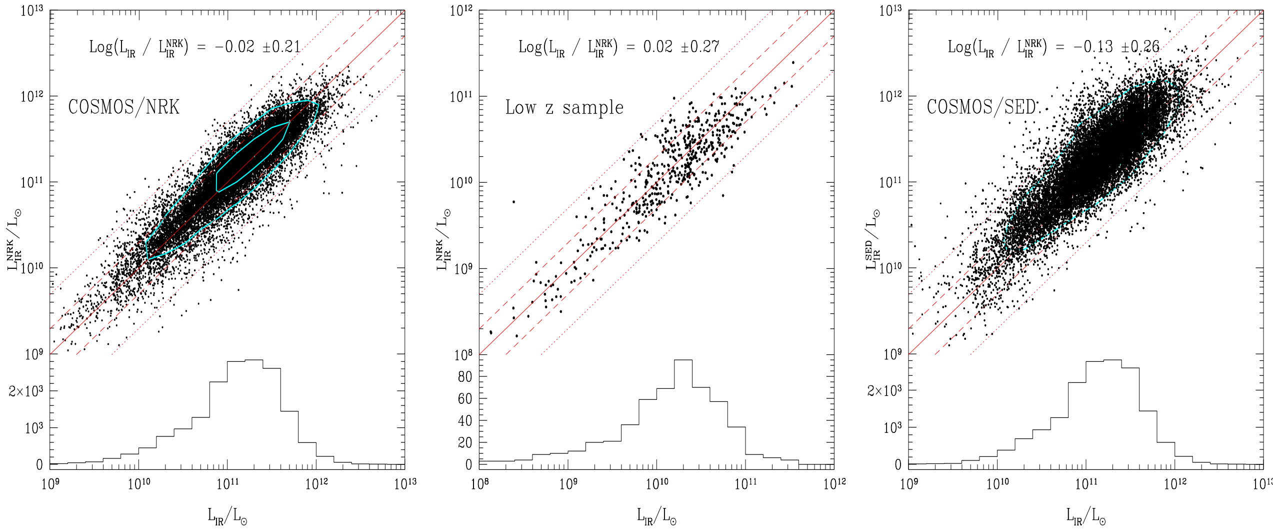


Fig. 7. Comparison of the infrared luminosity L_{IR} estimated from the 24 μm luminosity (COSMOS sample) or 8, 24, 70 and 160 μm luminosities (local sample) with that estimated with NRK ($L_{\mathrm{IR}}^{\mathrm{NRK}}$), for the COSMOS (*left panel*) and SWIRE (*middle panel*) samples. We also compare with the one estimated from the SED fitting ($L_{\mathrm{IR}}^{\mathrm{SED}}$, *right panel*). In each panel, the dashed and dotted lines refer to a factor ratio of 2 and 5 respectively. The distributions of L_{IR} are shown in the bottom part of each panel. The numbers in each panel refer to the logarithm of the mean and dispersion of the luminosity ratio for the star-forming galaxies.

from $\sigma(IRX) \sim 0.27$ to 0.36, depending on the adopted colors. Even with the use of the $D_n(4000)$ break as a dust-free indicator of star formation history, they do not achieve a better accuracy than that obtained with the method presented in this work.

Finally, our method provides a smaller dispersion in the prediction of the IR luminosity than the dust luminosity obtained from the SED fitting (Fig. 7 right panel), where we obtain a dispersion of ~0.26 dex and a bias of $\Delta \sim -0.13$. We note that the L_{IR} derived from SED fitting is computed by integrating all the stellar photons absorbed by dust according to the adopted attenuation law and reddening excess (see Appendix A), while the reference L_{IR} is based on the extrapolation of the 24 μm luminosity. Both measurements suffer from independent source of uncertainties, thus the bias in the IR luminosity may be related to either methods. However, the small scatter (less than a factor 2) and bias show that the SED fitting is a robust and reliable method to estimate the total IR luminosity, when accurate observations on a wide enough wavelength are available (e.g., Salim et al. 2009). This is indeed the case for the COSMOS dataset used here (i.e., 31 passbands from Far-UV to Mid-IR: $0.15 \leq \lambda \leq 4.5$ μm).

4.2. Dependence of the predicted IR luminosity with physical galaxy properties

Despite the use of a single vector, ***NRK***, and the absence of mass dependency, our recipe provides a reasonable estimate of the total IR luminosity L_{IR}. However, galaxies with physical properties that deviate from the bulk of the 24 μm population may be less accurately described in this framework. To test this issue and determine the range of validity of the method, we show in Fig. 8 the residuals (i.e., the difference) between the reference (i.e., $L_{\mathrm{IR}}^{24\,\mu}$) and predicted IR luminosity as a function of redshift (top-left panel), total SFR (top-right), stellar mass (bottom-left) and specific SFR (bottom-right). The total SFR refers to the (UV+IR) SFR (see Eq. (1)), while the stellar mass is obtained with the SED fitting (see Sect. 2.2.2). We perform this comparison for the L_{IR} predicted with both the NRK and SED fitting methods.

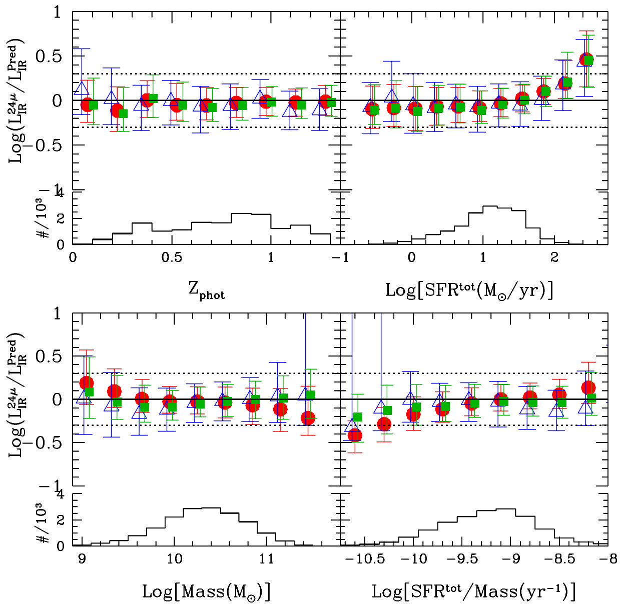


Fig. 8. Median and dispersion of the difference (i.e., residuals) between the IR luminosity based on the 24 μm flux (i.e., $L_{\mathrm{IR}}^{24\mu}$) and that predicted with different methods, as a function of different galaxy physical parameters. In each panel red-filled circles refer to the NRK method, green squares to the NRK sSFR-corrected and open-blue triangles to the SED fitting (open-blue triangles). *Top-left panel* shows the residuals as function of redshift, *top-right* of total SFR, *bottom-left* of stellar mass and *bottom-right* of specific SFR. The number distribution for each physical parameter is shown as an histogram in the lower part of *each panel*.

Top-left panel of Fig. 8 does not reveal any bias with redshift in the IR luminosity predicted with both methods. This is expected by construction for the NRK method, while it confirms the good performance of the SED fitting technique despite the larger scatter than the one obtained with the NRK method.

Top-right panel of Fig. 8 shows that the residuals as a function of the total SFR are relatively stable, with an almost zero bias and a dispersion of ~0.2 dex for $SFR \leq 100\,M_{\odot}\,\mathrm{yr}^{-1}$. At higher SFR, where galaxies approach the ULIRG regime, the residuals start to deviate from zero, indicating that both the NRK and SED fitting methods under-predict the reference IR luminosity. The difference with the NRK predictions is likely to be caused by the rarity of these highly star-forming objects in our redshift range ($0 \leq z \leq 1.3$, see also Le Floc'h et al. 2005), which makes these objects under-represented in the NRK calibration. The disagreement with the prediction of the SED fitting may have a different origin. As shown in Appendix A, the starburst attenuation law is favored at high SFR, however ULIRGs are found to deviate from Meurer's relation (Reddy 2009; Reddy et al. 2012) and exhibit a higher IRX at a given UV-slope. This effect has been observed at higher redshift ($z \geq 1.5$), but it could also be the reason of the under-estimate of L_{IR} by the SED fitting technique at lower redshift. We also note that our *e*-folding SF history models used in the SED fitting may be too simplistic and models including burst episodes would be more appropriated for these actively star-forming galaxies. Finally, we also note that the reference L_{IR}, based on the extrapolation of the 24 μm luminosity, may reach its limit of validity in this regime. Indeed, the conversion of the 24 μm luminosity to total IR luminiosity becomes uncertain for ULIRGs (i.e., for $L_{\mathrm{IR}} \geq 10^{12}\,L_{\odot}$, see Bavouzet et al. 2008; Goto et al. 2011). Also, the merger nature of ULIRGs at low z (Kartaltepe et al. 2010) makes predictions inaccurate in absence of a complete set of far-IR observations.

Bottom-left panel of Fig. 8 does not reveal any bias with stellar mass for the IR luminosity predicted with the SED fitting technique, while the dispersion increases at the extreme sides. On the other hand, the NRK method under-estimates L_{IR} for stellar mass $\log(M/M_{\odot}) \lesssim 9.3$. As shown in Le Floc'h et al. (in prep.) from the analysis of a mass-complete sample obtained with stacking techniques of *Herschel*/SPIRE data at 250/350/500 μm, the NRK calibration presented in this work needs to be modified for galaxies with $M \lesssim 10^{9.5}\,M_{\odot}$. At high stellar mass (i.e., $\log(M/M_{\odot}) \gtrsim 11$) the NRK method over-estimates L_{IR}. This reflects the correlation between stellar mass and sSFR, where quiescent galaxies become the dominant population at high stellar masses (see below).

Bottom-right panel of Fig. 8 shows that the NRK method systematically (as reflected by the small scatter) over-estimates L_{IR} at low sSFR (i.e., $\log(\mathrm{sSFR}/\mathrm{yr}^{-1}) \leq -10$). We investigate this bias in more details in Appendix D and find that the sSFR derived with NRK saturates at such low sSFR, while the reference sSFR keeps decreasing. We therefore propose a simple analytical correction for this bias in order to reconcile the two sSFR estimates, which can then be used to correct $L_{\mathrm{IR}}^{\mathrm{NRK}}$. The results for this $L_{\mathrm{IR}}^{\mathrm{NRK}}$ sSFR-corrected are shown in Fig. 8 as filled green squares. The correction provides a better match to the 24 μm parent distribution across the entire range of stellar mass and sSFR, but at the cost of a larger dispersion for the global sample (i.e., from $\sigma = 0.21$ to $\sigma = 0.24$). It is worth noting that the reference sSFR for quiescent galaxies might also suffer from systematic error. In fact, recent analysis with *Herschel* data have reported evidences for a warmer dust temperature in early- than in late-type galaxies (Skibba et al. 2011; Smith et al. 2012). This will result in an over-estimation of L_{IR} as derived from the 24 μm luminosity (by adopting a too high $L_{\mathrm{IR}}/L_{24\,\mu}$ ratio). Combined with the UV luminosity partially produced by evolved stars, these two effects can lead to an over-estimate of our reference SFR (based on UV+IR). As seen in Fig. 8, the L_{IR} comparison with the SED fitting method shows a large dispersion at low sSFR. The reason is illustrated in Appendix A (Fig. A.1), where a significative fraction of the galaxies with $\log(sSFR_{\mathrm{tot}}/\mathrm{yr}^{-1}) \leq -10$ shows a much lower sSFR with the SED fitting method, leading to a lower L_{dust}.

In conclusion, the NRK and SED fitting methods provide two independent and reliable estimates of L_{IR} over a large range of redshift and galaxy physical parameters for the vast majority (~90%) of the 24 μm star-forming sample. However, differences arise when considering the extreme sides of the population, such as highly star-forming (i.e., $SFR_{\mathrm{tot}} \geq 100\,M_{\odot}\,\mathrm{yr}^{-1}$) or quiescent galaxies (i.e., $\log(sSFR_{\mathrm{tot}}/\mathrm{yr}^{-1}) \leq -10$), which represent $\lesssim$1% and 7% of the whole sample, respectively. While the NRK method suffers from a bias at low sSFR, it is not clear the amplitude of this effect, since there is no robust indicators of SFR in the low activity regime. In the next section, we obtain an independent estimate of this bias for the quiescent population by directly comparing the SFRs estimated with the NRK and SED fitting methods.

4.3. Comparison of SFR estimates with NRK and SED fitting techniques

Once calibrated with a Far-IR sample, the NRK technique can be applied to any sample of galaxies. In this section we compare the SFRs derived with the SED fitting and NRK methods for two samples of galaxies, the 24 μm and a *K*-selected (down to $K \leq 23.5$) samples. We restrict the analysis to galaxies with stellar masses $M \gtrsim 2 \times 10^{9}\,M_{\odot}$ and redshift $0.05 \leq z \leq 1.3$. The SFR^{NRK} is estimated by means of Eqs. (1)–(3). As discussed above, we found that this technique is better suited for active star-forming galaxies than the less active population (low sSFR). As shown in Fig. 2, we exploit the capability of the *NUVrK* diagram to separate galaxies with different sSFRs and we define a new vector $\boldsymbol{NRK}_{\mathrm{ssfr}}$, perpendicular to the starburst attenuation vector ($\boldsymbol{NRK}_{\mathrm{ssfr}} = \cos(54°) \times (\mathrm{NUV} - r) - \sin(54°) \times (r - K)$). In Fig. 2, we show constant values of NRK_{ssfr} (i.e., the norm of the $\boldsymbol{NRK}_{\mathrm{ssfr}}$ vector), corresponding to the range $-2 \leq NRK_{\mathrm{ssfr}} \leq 3$ as gray lines, which is a good proxy to follow the variation of the mean sSFR. The limit adopted to define the passive population corresponds to $NRK_{\mathrm{ssfr}} \gtrsim 1.9$.

The ratios ($SFR^{\mathrm{NRK}}/SFR^{\mathrm{SED}}$) for the 24 μm (top panel) and the K (bottom panel) selected samples are shown in Fig. 9 for different intervals of NRK_{ssfr}. In each panel, we report for each bin of NRK_{ssfr} the percentage of objects in this bin, the percentage of catastrophic objects (estimated as ABS(Δ) > 3σ), the median (Δ) and scatter (σ). For ~85% of the sample, with $NRK_{\mathrm{ssfr}} \leq 1.3$, the two SFRs are in excellent agreement, with small biases, scatters and catastrophic fraction for both the 24 μm and K-selected samples. For $1.3 \leq NRK_{\mathrm{ssfr}} \leq 1.5$ (~6–7% of the two samples), the fraction of catastrophic objects sharply increases to ~10%, the dispersion also increases and the median starts to deviate from zero, with SFR^{SED} being higher than SFR^{NRK}. This trend becomes more severe for $1.5 \leq NRK_{\mathrm{ssfr}} \leq 1.9$ (the remaining ~7 to 9% of the samples). The two SFRs disagree, with a large fraction of catastrophic objects ($\gtrsim$25%), a significant non-zero median and large dispersion, with some slight differences between the two samples. This region with inconsistent SFR estimates is shown in Fig. 2 as a shaded area.

This comparison shows that the NRK method can be successfully applied to a sample of galaxies larger than the 24 μm sample, since it provides SFR estimates in good agreement with the SED fitting for the vast majority (i.e., ~85–90%) of the

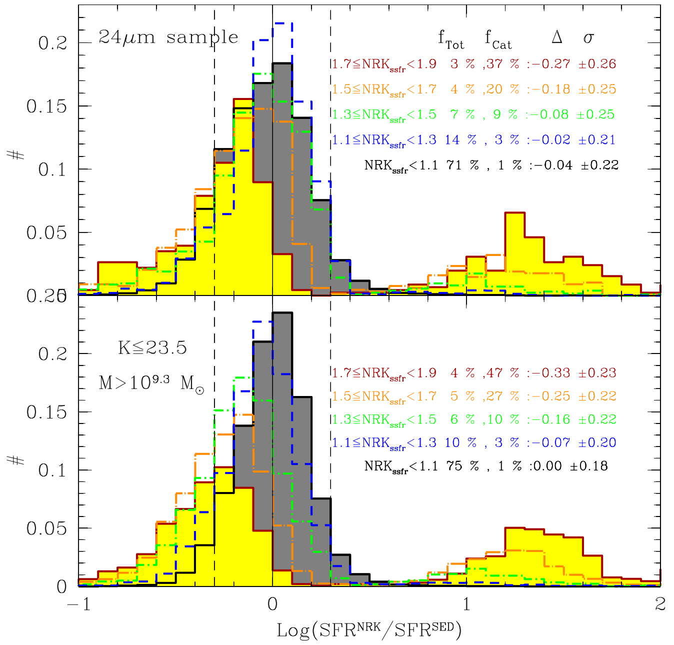


Fig. 9. Distributions of the SFR ratios ($SFR^{\rm NRK}/SFR^{\rm SED}$) in different bins of $\rm NRK_{ssfr}$. The distribution are renormalized to unity. The *top panel* refers to the 24 μm sample while the *bottom panel* includes all galaxies with $K \leq 23.5$ and more massive than $M \sim 2 \times 10^9\ M_{\odot}$. For each interval of $\rm NRK_{ssfr}$, we report the fraction of objects ($f_{\rm Tot}$), catastrophic failures ($f_{\rm Cat}$), the median (Δ) and dispersion (σ).

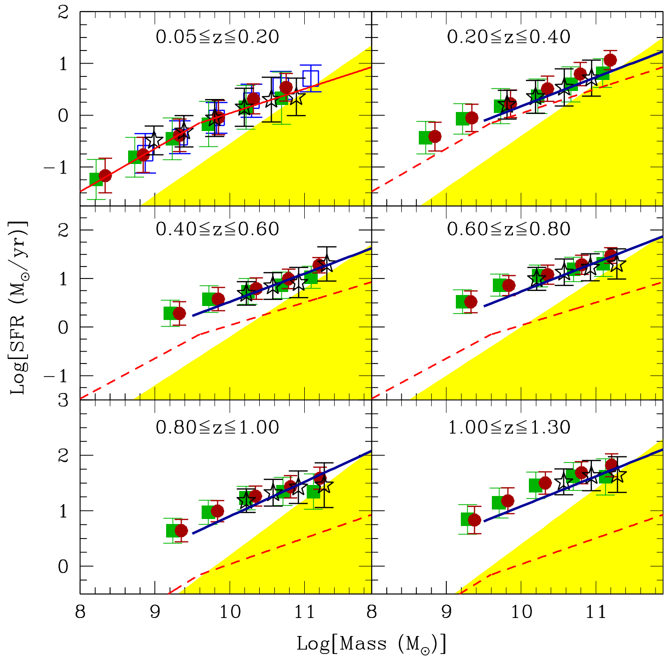


Fig. 10. SFR vs. stellar Mass relation for star-forming galaxies in different redshift bins. The $V_{\max}$ weighted means are shown for the 24 μm sample (SFR based on Eq. (1), open black stars); the COSMOS K-selected sample using the SFR derived from the NRK approaches (original method: filled red circles and with sSFR-correction: filled green squares); the local GALEX-SDSS-SWIRE sample (blue open squares). We compare with the local estimate from Salim et al. (2007) (red lines) and the radio stacking analysis by Karim et al. (2011) (heavy dark-blue lines). The region where the NRK method becomes less reliable is shown as light shaded area.

star-forming population. However, the methods diverge when including low-activity (low-ssFR) galaxies (see also Johnson et al. 2007a; Treyer et al. 2007). To alleviate this limitation, we show that the population with discording results (≤10% of the whole sample) can be easily isolated in the *NUVrK* diagram.

4.4. Application to the SFR vs. stellar mass relationship

A correlation between SFR and stellar mass in star-forming galaxies is observed both at low (Salim et al. 2007) and high redshift (Noeske et al. 2007b; Elbaz et al. 2007). This relation is surprisingly tight, with an intrinsic scatter of ~0.3 dex. We can therefore exploit the presence of this correlation and test the ability of the NRK method to reproduce the slope and amplitude of this relation at different redshifts.

In Fig. 10 we show the $V_{\max}$ weighted mean of SFR in bins of stellar mass for different samples of star-forming galaxies: the COSMOS 24 μm sample, a flux limited K-selected sample ($K \leq$ 23.5) and, in the lowest redshift bin, the local GALEX-SWIRE sample. Symbols indicate SFR estimated with different methods: ($L_{\rm IR} + L_{\rm UV}$) from Eq. (1) (black stars), NRK method (red filled circles and open blue squares), NRK method corrected for the sSFR bias (green-filled squares).

At low-z, the different SFR estimates provide similar results and the SFR-M_* relation is in excellent agreement with the one measured by Salim et al. (2007) for the GALEX-SDSS sample and derived with a SED fitting method (solid red line). At higher redshifts, we compare our finding with the results of Karim et al. (2011), which are based on the 1.4 GHz radio continuum emission of stacked star-forming galaxies in different stellar mass and redshift bins (solid blue lines). We consider their results from the two-parameter fits given in their Table 4. The SFR-M_* relation derived from individual galaxies with both the NRK sSFR bias-uncorrected and corrected methods agrees well with the radio stacking from Karim et al. (2011). The major difference between the NRK and the NRK-sSFR corrected method is in the dispersions around the mean SFRs. The original NRK method shows a small dispersion in the high mass end of the relation, due to an over-estimate of the SFR for massive, evolved galaxies, artificially moving them towards the SFR-M_* sequence. This effect varies with redshift, as described in Appendix D, thus we indicate with the shaded area in Fig. 10 the regions in which the SFR is over-estimated by a factor greater than two. It can be seen that, at all redshifts, the bin corresponding to the most massive galaxies is affected by this bias, hence producing a smaller dispersion on the SFR-M_* relation with respect to that observed with the 24 μm and NRK-sSFR corrected methods. We note that this effect becomes irrelevant at lower masses.

Overall, even if the stellar mass does not enter in the calibration of the NRK method, the SFR estimated with this method for individual galaxies can reproduce the slope and normalisation of the SFR-M_* relation, along with its redshift evolution.

5. Modeling the ⟨IRX⟩ in the *NUVrK* diagram

In the previous section we have shown that galaxies with different ultraviolet-to-infrared luminosity ratios (IRX) are well separated in the *NUVrK* diagram. To confirm the validity of this approach, and to explain the physical origin of the observed trends, we appeal to a library of galaxy SEDs computed with the BC03 spectral evolution model. We follow the approach of Pacifici et al. (2012) and extract a set of star formation and chemical enrichment histories from the semi-analytic post-treatment of De Lucia & Blaizot (2007) of the Millennium cosmological simulation (Springel et al. 2005). The star formation and chemical enrichment histories computed in this way reproduce the mean properties of nearby SDSS galaxies. Thus, they span only limited ranges in sSFR, around $10^{-10}\,\mathrm{yr}^{-1}$, and in the fraction of the current stellar mass formed in the last 2.5 Gyr at $z = 0$. Following Pacifici et al. (2012), to account for the broader range of spectral properties of the galaxies in our sample with respect to the SDSS, we re-draw the evolutionary stage at which a galaxy is looked at in the library of star formation and chemical enrichment histories (we do this uniformly in redshift between 0.2 and 1.5). We also resample the current (i.e., averaged over the last 30 Myr) SFR from a Gaussian distribution centered on $\log(sSFR) = -9.1$, with a dispersion of 0.6. This choice of parameters reproduces the observed global distribution of sSFR (i.e., summed over all redshift bins) and the distribution of galaxies in the *NUVrK* diagram, after accounting for dust attenuation as described in the next section. We adopt the Chabrier (2003) IMF.

5.1. Dust attenuation model

To include the effect of dust attenuation, we adopt the dust prescription of Chevallard et al. (2013). This extends the two-components, angle-average dust model of Charlot & Fall (2000) to include the effect of galaxy inclination and different spatial distributions of dust and stars on the observed SEDs of galaxies. Chevallard et al. (2013) combine the radiative transfer model of Tuffs et al. (2004, hereafter T04) with the BC03 spectral evolution model. To accomplish this, they relate the different geometric components of the T04 model (a thick and thin stellar disks, and a bulge, attenuated by two dust disks) to stars in different age ranges. Here, for the sake of simplicity and to limit the number of adjustable parameters, we describe attenuation in the diffuse ISM using only the thin stellar disk model of Tuffs et al. (2004). This is supported by the finding by Chevallard et al. (2013) that, in a large sample of nearby star-forming galaxies, the thin stellar disk component of the T04 model accounts for ≈80 percent of the attenuation in the diffuse ISM. Also, we note that adding the T04 thick disk component has a weak effect on the results. The dust content of the diffuse ISM in the T04 model is parametrized by means of the B-band central face-on optical depth of the dust disks $\tau_{B\perp}$. This determines, at fixed geometry, the attenuation of starlight by dust at any galaxy inclination θ, which measures the angle between the observer line-of-sight and the normal to the equatorial plane of a galaxy. At fixed $\tau_{B\perp}$, the integration over the solid angle of the attenuation curve $\hat{\tau}_\lambda^{\rm ISM}(\theta)$ in the T04 model yields the angle-average attenuation curve $\langle\hat{\tau}_\lambda^{\rm ISM}\rangle_\theta$ (see Sect. 2 of Chevallard et al. 2013). As in Charlot & Fall (2000), we couple the attenuation in the diffuse ISM described by the T04 model with a component describing the enhanced attenuation of newly born stars ($t < 10$ Myr) in their parent molecular clouds. Following Charlot & Fall (2000), we parametrize this enhanced attenuation by means of the fraction 1-μ of the total attenuation that arises from dust in stellar birth clouds, in the angle-average case.

The attenuation of the radiation emitted by a stellar generation of age t at inclination θ can therefore be written as

$$\hat{\tau}_\lambda^{\rm tot}(\theta,t) = \begin{cases} \hat{\tau}_\lambda^{\rm BC} + \hat{\tau}_\lambda^{\rm ISM}(\theta) & \text{for} \quad t \leqslant 10\,\mathrm{Myr}\,, \\ \hat{\tau}_\lambda^{\rm ISM}(\theta) & \text{for} \quad t > 10\,\mathrm{Myr}\,, \end{cases} \tag{4}$$

where the superscripts "BC" and "ISM" refer to attenuation in the birth clouds (assumed isotropic) and the diffuse ISM, respectively. In this expression the attenuation curve for the diffuse ISM $\hat{\tau}_\lambda^{\rm ISM}(\theta)$ is taken from the T04 thin stellar disk model, and following Wild et al. (2007, see also da Cunha et al. 2008), we compute the attenuation curve in the birth clouds as

$$\hat{\tau}_\lambda^{\rm BC} = \hat{\tau}_V^{\rm BC}\;(\lambda/0.55\;\mu\mathrm{m})^{-1.3}\,, \tag{5}$$

where the V-band optical depth of the birth clouds $\hat{\tau}_V^{\rm BC}$ is related to the angle-average optical depth of the diffuse ISM $\langle\hat{\tau}_V^{\rm ISM}\rangle_\theta$ as

$$\hat{\tau}_V^{\rm BC} = (1-\mu)\,/\mu\,\langle\hat{\tau}_V^{\rm ISM}\rangle_\theta. \tag{6}$$

To compute the ratio of the infrared-to-ultraviolet luminosities *IRX* in this model, we take the IR luminosity to be equal to the fraction of all photons emitted in the range 912 Å $\leq \lambda \leq 3\;\mu$m in any direction that are absorbed by dust (dust is almost transparent at $\lambda > 3\;\mu$m). Assuming that photons at IR wavelengths emerge isotropically from a galaxy, we write the IR luminosity $L_{\rm IR}$ as

$$L_{\rm IR} = \frac{1}{4\pi}\int_\Omega \mathrm{d}\Omega \int_{0.0912}^{3} \mathrm{d}\lambda[1-\exp(-\hat{\tau}_\lambda(\theta))]L_\lambda^0 \tag{7}$$

where L_λ^0 is the unattenuated luminosity emitted by stars (assumed isotropic), Ω is the solid angle, and $\hat{\tau}_\lambda(\theta)$ is the integral of Eq. (4) over the star formation history of the galaxy. We compute the monochromatic ultraviolet luminosity at the frequency ν corresponding to $\lambda = 2300$ Å as $\mathcal{L}_{\rm NUV}(\theta) = \nu L_\nu(\theta)$, where $L_\nu(\theta)$ is given by

$$L_\nu(\theta) = [\exp(-\hat{\tau}_\lambda(\theta))]\,L_\lambda^0\,\frac{\lambda^2}{c}, \tag{8}$$

and L_λ^0 if the luminosity emitted by all stars at $\lambda = 2300$ Å in the direction θ, and c is the speed of light.

5.2. Model library

We use this model to compute a library of 20 000 SEDs of dusty star-forming galaxies, which we divide in bins of constant (NUV − r) and (r − K). We compute the mean sSFR and ⟨IRX⟩ in each bin and explore their distribution in the *NUVrK* diagram. After some experimentation, we find that a Gaussian distribution of $\tau_{B\perp}$ centered at 7, with a dispersion of 3, truncated at the maximum value of available models $\tau_{B\perp} = 8$, and a Gaussian distribution of μ centered at 0.3, with a dispersion of 0.2, and truncated at $\mu = 0$ and $\mu = 1$, allow us to well reproduce the data, as shown in the top-left panel of Fig. 11. We note that a uniform distribution of galaxy inclinations would produce a large tail of highly attenuated galaxies, which is not observed in the data (at $(r-K) > 2.5$). Hence, in Fig. 11 we have adopted the observed distribution of axis ratios, converted to inclinations using the standard formula for an oblate spheroid (e.g., Guthrie 1992)

$$\cos\theta = \sqrt{\frac{(b/a)^2 - q_0^2}{1-q_0^2}}, \tag{9}$$

where q_0 is the intrinsic axis ratio of the ellipsoid representing the galaxy, which we fix to $q_0 = 0.15$.

A comparison between top-left and bottom-right panels of Fig. 11 shows that the models reproduce, at least qualitatively, the distribution of ⟨IRX⟩ in the color–color plane. The reddest

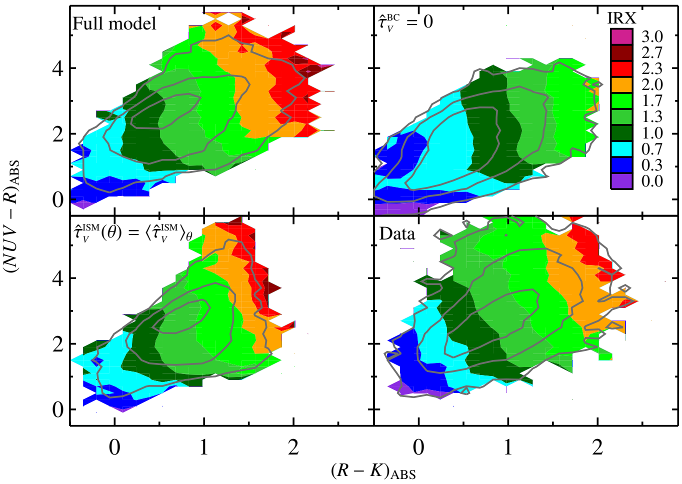


Fig. 11. Values of ⟨IRX⟩ (color coded on a logarithmic scale) in the *NUVrK* diagram. The solid gray lines in each panel indicate the number density contour of the galaxies corresponding to 0.01, 0.1 and 0.5 the maximum density. *Top-left panel*, 20 000 model SEDs computed with the "full model" (see Sect. 5). This includes the dust prescription of Chevallard et al. (2013), which accounts for the effect on dust attenuation of galaxy geometry, inclination and enhanced attenuation of young stars by their birth clouds. Top-right panel, same as *top-left panel*, but neglecting the enhanced attenuation of young stars, i.e., fixing the birth clouds optical depth $\hat{\tau}_V^{\rm BC} = 0$. *Bottom-left panel*, same as *top-left panel*, but neglecting the effect of galaxy inclination, i.e., adopting the angle-averaged attenuation curves $\langle\hat{\tau}_\lambda^{\rm ISM}\rangle_\theta$. *Bottom-right panel*, data (see Sect. 3 and Fig. 4).

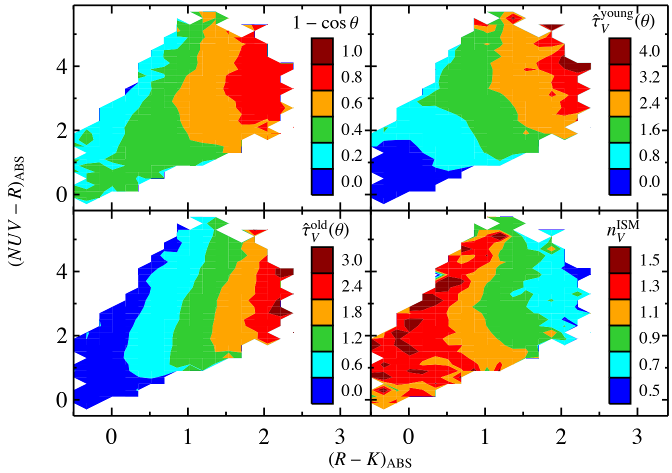


Fig. 12. Mean value in bins of constant (NUV − r) and ($r - K$) of different parameters describing attenuation of starlight from dust for the model SEDs described in Sect. 5. *Top-left panel*: galaxy inclination $1 - \cos\theta$. *Top-right panel*, V-band attenuation optical depth suffered by stars younger than 10^7 yr $\hat{\tau}_V^{\rm young}(\theta)$. *Bottom-left panel*, V-band attenuation optical depth suffered by stars older than 10^7 yr $\hat{\tau}_V^{\rm old}(\theta)$. *Bottom-right panel*, slope of the optical attenuation curve in the diffuse ISM, $n_V^{\rm ISM}(\theta)$, measured from a power law fit to the model attenuation curves in the range $0.4 \le \lambda \le 0.7$ μm.

near-IR colors, $(r - K) > 1.4$, correspond to galaxies seen at large inclination. This is consistent with Fig. B.2, which shows that the galaxies with reddest $(r-K)$ colors have the largest measured ellipticities, i.e., they are more inclined. A large inclination makes the disk appear more opaque, since photons have to cross a larger section of the dust disk before they escape toward the observer.

The location and shape of the *IRX* stripes in the theoretical *NUVrK* diagram depend on several galaxy physical parameters, namely evolutionary stage, current SFR, dust content and distribution. The evolutionary stage and current SFR determine the relative amount of young and old stars in the galaxy, which controls the ratio of unattenuated ultraviolet to optical and near-IR luminosity of the galaxy. The global dust content, $\hat{\tau}_\lambda^{\rm BC} + \hat{\tau}_\lambda^{\rm ISM}(\theta)$, and the distribution of dust between ambient ISM and birth clouds affect the (NUV−r) and the $(r-K)$ colors in different ways. For galaxies with a non-negligible fraction of young stars (i.e., $\log(sSFR) \gtrsim -9$), the (NUV − r) color is mainly driven by the stellar birth clouds optical depth $\hat{\tau}_\lambda^{\rm BC}$, and the $(r - K)$ color by the optical depth of the diffuse ISM $\hat{\tau}_\lambda^{\rm ISM}(\theta)$.

We test the effect of varying the optical depth of stellar birth clouds by computing the same library of 20 000 galaxy SEDs as described above, but fixing $\hat{\tau}_V^{\rm BC} = 0$. Top-right panel of Fig. 11 shows that neglecting birth clouds attenuation prevents us from reproducing the reddest (NUV − r) color observed in the data. The stripes appear almost perpendicular to the $(r - K)$ color driven primarily by the diffuse ISM. Also, this model predicts smaller values of ⟨IRX⟩ at a fixed position in the (NUV − r) vs. $(r-K)$ diagram, since the UV photons emitted by young stars do not suffer enhanced attenuation by the dusty birth cloud environment, which would be re-emitted at IR wavelengths increasing the overall IR luminosity.

We also study the effect of neglecting the dependance of dust attenuation on galaxy inclination. To achieve this, we compute the same library of 20 000 SEDs as above, but we fix the attenuation curve to the angle-averaged curve $\langle\hat{\tau}_\lambda^{\rm ISM}\rangle_\theta$. Bottom-left panel of Fig. 11 shows that this prevents us from reproducing the reddest $(r-K)$ colors of the observed galaxies, which correspond to highly inclined objects (see top-left panel of Figs. 12 and B.2).

We have shown in Fig. 11 that to reproduce qualitatively the observed distribution of galaxies and the value and orientation of the ⟨IRX⟩ stripes in the (NUV − r) vs. $(r - K)$ plane we need a prescription for dust attenuation which includes both a two-component medium (i.e., ISM + birth clouds) and the effect of galaxy inclination. We can now consider the "Full model" shown in the top-left panel of Fig. 11 and study how different dust properties vary in the (NUV−r) vs. $(r-K)$ plane. Figure 12 shows the same library of SEDs as in the top-left panel of Fig. 11. As for Fig. 11, we divide the galaxy SEDs in bins of constant (NUV−r) and $(r - K)$, and compute in each bin the mean value of galaxy inclination $1 - \cos\theta$, V-band attenuation optical depth seen by stars younger [older] than 10^7 yr $\hat{\tau}_V^{\rm young}(\theta)$ [$\hat{\tau}_V^{\rm old}(\theta)$], and slope of the optical attenuation curve in the diffuse ISM $n_V^{\rm ISM}(\theta)$.

The top-left panel of Fig. 12 shows that the galaxy inclination systematically increases as $(r - K)$ increases, with a weaker dependence on the (NUV − r) color. This can be understood in terms of the attenuation optical depth in the diffuse ISM, which increases with increasing galaxy inclination, hence making $(r - K)$ redder. The (NUV − r) color is also influenced by the variation of galaxy inclination, but to a much less extent since it also depends on the birth clouds attenuation optical depth (see Eq. (4)). The variation of the mean galaxy inclination shown in the top-left panel of Fig. 12 is also in qualitative agreement with Fig. B.2, which shows that the mean observed ellipticity of the galaxies in our sample increases from the bottom-left to the top-right side of the (NUV − r) vs. $(r - K)$ plane.

The top-right panel of Fig. 12 shows the variation of the V-band attenuation optical depth suffered by stars younger than 10^7 yr (i.e., $\hat{\tau}_V^{\rm young}(\theta)$, see Eq. (4)). At small $(r - K)$ the stripes of constant $\hat{\tau}_V^{\rm young}(\theta)$ are more parallel to the (NUV − r) color, while they become more and more inclined as (NUV − r) and (r − K) increase. This behavior can be understood in terms of the different fraction of light emitted by young stars attenuated by dust in the birth clouds and in the diffuse ISM. At small (i.e., blue) (r − K), young stars are mostly attenuated by the birth clouds component, which makes the (NUV − r) larger (i.e., redder, by decreasing NUV, at fixed r) without affecting (r − K). As (r − K) increases, the attenuation in the diffuse ISM increases, because galaxies are more inclined or have a larger dust content, and so the fraction of light emitted by young stars attenuated by this component raises too. As a consequence, the (NUV − r) color is determined by attenuation in both components, and the stripes of constant $\hat{\tau}_V^{\rm young}(\theta)$ change orientation.

The variation of the V-band attenuation optical depth suffered by stars older than 10^7 yr (i.e., $\hat{\tau}_V^{\rm old}(\theta)$) shown in the bottom-left panel of Fig. 12 follow that of the (r − K) color, as indicated by the orientation of the stripes of constant $\hat{\tau}_V^{\rm old}(\theta)$ almost perpendicular to (r − K). This is not surprising, since (r − K) traces stars older than 10^7 yr, which are attenuated by dust in diffuse ISM. When moving from left to right on the (r − K) axis, $\hat{\tau}_V^{\rm old}(\theta)$ increases from 0 to 2.5, indicating that the amount of attenuation suffered by stars in galaxies with very large (i.e., red) (r − K) is substantial.

The bottom-right panel of Fig. 12 shows the slope of the optical attenuation curve in the diffuse ISM $n_V^{\rm ISM}(\theta)$, obtained by fitting a power law to the model attenuation curves in the range $0.4 \leq \lambda \leq 0.7$ μm. The slope $n_V^{\rm ISM}(\theta)$ becomes smaller (i.e., the attenuation curve becomes flatter) when moving from the bottom-left to the top-right side of the diagram. This effect, as described in Chevallard et al. (2013, see their Fig 4), is a general prediction of radiative transfer models which consider disk galaxies with a mixed distribution of dust and stars. The variation of the slope of the attenuation curve in the (NUV − r) vs. (r − K) plane can account for the fact that the observed stripes of constant IRX are not perpendicular to the SMC and Calzetti attenuation vectors (see Fig. 4) and that the SED fitting predicts a systematic variation of the slope of the attenuation curve as a function of sSFR (see Fig. A.2).

With this analysis we have shown that we must account for the effect of geometry (i.e., of the spatial distribution of dust and stars) and galaxy inclination to reproduce the attenuation of starlight by the diffuse ISM, which mainly affects the (r − K) color. We have also shown the importance of accounting for the enhanced attenuation of newly born stars by their birth clouds to reproduce the reddest (NUV − r) colors and to match the observed values of ⟨IRX⟩. In the end, the good qualitative agreement between the "fully model" galaxies and the data (i.e., top-left and bottom-right panels of Fig. 11) confirms that the *NUVrK* diagram encodes valuable information about the global energy transfer between starlight and dust and galaxy inclination. We defer to a future work a more detailed and quantitative analysis of the data presented here, which would help us to better constrain the global amount, distribution and redshift evolution of the dust in star-forming galaxies.

6. Conclusion

We present a new method to compute the SFR of individual star-forming galaxies based on their location in the (NUV − r) versus (r − K) color–color diagram. We show that the *NUVrK* diagram provides an efficient way to separate quiescent and star-forming galaxies, an alternative to the UVJ diagram proposed by Williams et al. (2009). For the star-forming galaxies, the UV/optical luminosities in this diagram are highly sensitive to the shape of the dust attenuation laws. On the other hand, the infrared excess, $IRX = L_{\rm IR}/L_{\rm UV}$, as the net budget of the absorbed versus unabsorbed UV light, is weakly dependent on these effects. We combine the two dust diagnostics by analyzing the distribution of the mean infrared excess (i.e., ⟨IRX⟩) in the *NUVrK* diagram for a large sample of star-forming galaxies at redshift $0 \leq z \leq 1.3$ selected from the COSMOS 24 μm and low-z GALEX-SDSS-SWIRE (Johnson et al. 2007a) samples. We observe the presence of stripes with constant ⟨IRX⟩, associated with a small dispersion around the mean, which allows us to describe ⟨IRX⟩ with a unique vector (i.e., ***NRK***, a combination of (NUV − r) and (r − K) colors). We derive a simple relation between ⟨IRX⟩ and NRK, the norm of the vector ***NRK***, and redshift, valid for star-forming galaxies with $M \geq 2\times10^9\ M_\odot$ and 0.05–0.1 $\leq z \leq$ 1.3. This relation allows us to predict the IR luminosity of individual galaxies with an accuracy of ~0.2 dex (up to 0.27 dex for the local sample), which is better than the accuracy obtained with the SED fitting method based on 31 COSMOS pass-bands for the same galaxies.

We perform extensive comparisons of the $L_{\rm IR}$ and SFRs derived with the 24 μm, NRK and SED fitting methods. We find that the three methods provide consistent results for the vast majority of star-forming galaxies (~85–90%). The methods diverge for highly star-forming galaxies ($SFR \geq 100\,M_\odot\,{\rm yr}^{-1}$), which remain a negligible population at $z \leq 1.3$, and for more evolved galaxies ($sSFR \lesssim 10^{-10}\,{\rm yr}^{-1}$). For the latter, we describe a sSFR-, redshift-dependent limit below which the NRK method becomes unreliable and we also show that this population with inconsistent SFR estimates can be easily isolated and discarded in the *NUVrK* diagram.

By using the NRK method, we reconstruct the relationship between SFR and stellar mass for a K-selected sample of star-forming galaxies and find an excellent agreement with previous results over the entire redshift range.

Finally, we investigate the physical origin of the ⟨IRX⟩ stripes in the *NUVrK* diagram by appealing to a library of model SEDs based on the population synthesis code of Bruzual & Charlot (2003). We find that this library of models is able to qualitatively reproduce the location and shape of ⟨IRX⟩ stripes in the *NUVrK* diagram if we adopt a realistic prescription for dust attenuation (Chevallard et al. 2013). We show that to reproduce the observed stripes of ⟨IRX⟩ we must appeal to a two-component (i.e., birth clouds + diffuse ISM) dust model, which must account also for the effect on dust attenuation of galaxy inclination and geometry (i.e., the spatial distribution of dust and stars).

The method discussed in this work offers a simple alternative to assess the total SFR of star-forming galaxies in the absence of Far-IR observations or spectroscopic diagnostics. Because it directly predicts the infrared excess, no assumption on the dust attenuation curves is required to derive the SFR, in contrast to other methods such as the β-slope or SED fitting.

In a companion paper Le Floc'h et al. (in prep.), we extend our analysis toward lower stellar mass and higher redshifts based on the stacking technique of the Far-IR emission using the complete dataset available from *Spitzer*/MIPS at 24 μm to *Herschel*/SPIRE at 250, 350 and 500 μm.

Acknowledgements. We thank the anonymous referee for his useful comments. This work has been partially supported by the CNRS-INSU and the Progamme National de Cosmologie et Galaxies. J.C. acknowledges the support of the European Commission through the Marie Curie Initial Training Network ELIXIR under contract PITN-GA-2008-214227.

References

Arnouts, S., Cristiani, S., Moscardini, L., et al. 1999, MNRAS, 310, 540
Bavouzet, N., Dole, H., Floc'h, E. L., et al. 2008, A&A, 479, 83
Bell, E. F., Papovich, C., Wolf, C., et al. 2005, ApJ, 625, 23
Boquien, M., Calzetti, D., Kennicutt, R., et al. 2009, ApJ, 706, 553
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2011, ApJ, 737, 90
Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, MNRAS, 351, 1151
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Burgarella, D., Le Floc'h, E., Takeuchi, T. T., et al. 2007, MNRAS, 380, 986
Calzetti, D. 1997, AJ, 113, 162
Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682
Capak, P., Aussel, H., Ajiki, M., et al. 2007, ApJS, 172, 99
Chabrier, G. 2003, PASP, 115, 763
Charlot, S., & Fall, S. M. 2000, ApJ, 539, 718
Chary, R., & Elbaz, D. 2001, ApJ, 556, 562
Chevallard, J., Charlot, S., Wandelt, B., & Wild, V. 2013, MNRAS, 432, 2061
Cortese, L., Boselli, A., Franzetti, P., et al. 2008, MNRAS, 386, 1157
Cortese, L., & Hughes, T. M. 2009, MNRAS, 400, 1225
Cowie, L. L., Songaila, A., Hu, E. M., & Cohen, J. G. 1996, AJ, 112, 839
da Cunha, E., Charlot, S., & Elbaz, D. 2008, MNRAS, 388, 1595
Daddi, E., Dickinson, M., Morrison, G., et al. 2007, ApJ, 670, 156
Daddi, E., Dannerbauer, H., Elbaz, D., et al. 2008, ApJ, 673, L21
Daddi, E., Bournaud, F., Walter, F., et al. 2010, ApJ, 713, 686
Dale, D. A., & Helou, G. 2002, ApJ, 576, 159
De Lucia, G., & Blaizot, J. 2007, MNRAS, 375, 2
Donas, J., & Deharveng, J. M. 1984, A&A, 140, 325
Elbaz, D., Daddi, E., Le Borgne, D., et al. 2007, A&A, 468, 33
Elbaz, D., Hwang, H. S., Magnelli, B., et al. 2010, A&A, 518, L29
Erb, D. K., Steidel, C. C., Shapley, A. E., et al. 2006, ApJ, 646, 107
Goto, T., Arnouts, S., Malkan, M., et al. 2011, MNRAS, 414, 1903
Guthrie, B. N. G. 1992, A&AS, 93, 255
Hao, C.-N., Kennicutt, R. C., Johnson, B. D., et al. 2011, ApJ, 741, 124
Heinis, S., Buat, V., Béthermin, M., et al. 2013, MNRAS, 429, 1113
Hoaglin, D. C., Mosteller, F., & Tukey, J. W. 1983, Understanding robust and exploratory data anlysis (New York: Wiley)
Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, A&A, 457, 841
Ilbert, O., Capak, P., Salvato, M., et al. 2009, ApJ, 690, 1236
Ilbert, O., Salvato, M., Le Floc'h, E., et al. 2010, ApJ, 709, 644
Inoue, A. K. 2005, MNRAS, 359, 171
Johnson, B. D., Schiminovich, D., Seibert, M., et al. 2007a, ApJS, 173, 392
Johnson, B. D., Schiminovich, D., Seibert, M., et al. 2007b, ApJS, 173, 377
Karim, A., Schinnerer, E., Martínez-Sansigre, A., et al. 2011, ApJ, 730, 61
Kartaltepe, J. S., Sanders, D. B., Le Floc'h, E., et al. 2010, ApJ, 709, 572
Kennicutt, Jr., R. C. 1983, ApJ, 272, 54
Kennicutt, R. C. 1998, ARA&A, 36, 189
Kong, X., Charlot, S., Brinchmann, J., & Fall, S. M. 2004, MNRAS, 349, 769
Lagache, G., Dole, H., Puget, J.-L., et al. 2004, ApJS, 154, 112
Le Floc'h, E., Papovich, C., Dole, H., et al. 2005, ApJ, 632, 169
Le Floc'h, E., Aussel, H., Ilbert, O., et al. 2009, ApJ, 703, 222
Lee, N., Floc'h, E., & Sanders, D. 2010, ApJ
Leitherer, C., & Heckman, T. M. 1995, ApJS, 96, 9
Lilly, S. J., Le Fevre, O., Hammer, F., & Crampton, D. 1996, ApJ, 460, L1
Lilly, S. J., Le Brun, V., Maier, C., et al. 2009, ApJS, 184, 218
Lonsdale, C. J., Smith, H. E., Rowan-Robinson, M., et al. 2003, PASP, 115, 897
Maraston, C. 2005, MNRAS, 362, 799
Maraston, C., Pforr, J., Renzini, A., et al. 2010, MNRAS, 407, 830
Martin, D. C., Seibert, M., Buat, V., et al. 2005, ApJ, 619, L59
Martin, D. C., Wyder, T. K., Schiminovich, D., et al. 2007, ApJS, 173, 342
McCracken, H. J., Capak, P., Salvato, M., et al. 2010, ApJ, 708, 202
Meurer, G. R., Heckman, T. M., & Calzetti, D. 1999, ApJ, 521, 64
Noeske, K. G., Faber, S. M., Weiner, B. J., et al. 2007a, ApJ, 660, L47
Noeske, K. G., Weiner, B. J., Faber, S. M., et al. 2007b, ApJ, 660, L43
Oke, J. B. 1974, ApJS, 27, 21
Pacifici, C., Charlot, S., Blaizot, J., & Brinchmann, J. 2012, MNRAS, 421, 2002
Pannella, M., Carilli, C. L., Daddi, E., et al. 2009, ApJ, 698, L116
Patel, S. G., Kelson, D. D., Holden, B. P., Franx, M., & Illingworth, G. D. 2011, ApJ, 735, 53
Prevot, M. L., Lequeux, J., Prevot, L., Maurice, E., & Rocca-Volmerange, B. 1984, A&A, 132, 389
Reddy, N. A. 2009, Galaxy Evolution: Emerging insights and future challenges, ASP Conf. Ser., 419, 1
Reddy, N. A., Steidel, C. C., Fadda, D., et al. 2006, ApJ, 644, 792
Reddy, N., Dickinson, M., Elbaz, D., et al. 2012, ApJ, 744, 154
Salim, S., Charlot, S., Rich, R. M., et al. 2005, ApJ, 619, L39
Salim, S., Rich, R. M., Charlot, S., et al. 2007, ApJS, 173, 267
Salim, S., Dickinson, M., Michael Rich, R., et al. 2009, ApJ, 700, 161
Sanders, D. B., Salvato, M., Aussel, H., et al. 2007, ApJS, 172, 86
Scarlata, C., Carollo, C. M., Lilly, S., et al. 2007, ApJS, 172, 406
Schiminovich, D., Ilbert, O., Arnouts, S., et al. 2005, ApJ, 619, L47
Scoville, N., Aussel, H., Brusa, M., et al. 2007, ApJS, 172, 1
Seibert, M., Martin, D. C., Heckman, T. M., et al. 2005, ApJ, 619, L55
Skibba, R. A., Engelbracht, C. W., Dale, D., et al. 2011, ApJ, 738, 89
Smit, R., Bouwens, R. J., Franx, M., et al. 2012, ApJ, 756, 14
Smith, M. W. L., Gomez, H. L., Eales, S. A., et al. 2012, ApJ, 748, 123
Spinoglio, L., Malkan, M. A., Rush, B., Carrasco, L., & Recillas-Cruz, E. 1995, ApJ, 453, 616
Springel, V., White, S., Jenkins, A., et al. 2005, Nature, 435, 629
Steidel, C. C., Giavalisco, M., Pettini, M., Dickinson, M., & Adelberger, K. L. 1996, ApJ, 462, L17
Stern, D., Kirkpatrick, J. D., Allen, L. E., et al. 2007, ApJ, 663, 677
Tacconi, L. J., Genzel, R., Neri, R., et al. 2010, Nature, 463, 781
Takeuchi, T. T., Yoshikawa, K., & Ishii, T. T. 2003, ApJ, 587, L89
Taniguchi, Y., Scoville, N., Murayama, T., et al. 2007, ApJS, 172, 9
Treyer, M., Schiminovich, D., Johnson, B., et al. 2007, ApJS, 173, 256
Treyer, M., Schiminovich, D., Johnson, B. D., et al. 2010, ApJ, 719, 1191
Tuffs, R. J., Popescu, C. C., Völk, H. J., Kylafis, N. D., & Dopita, M. A. 2004, A&A, 419, 821
Wild, V., Kauffmann, G., Heckman, T., et al. 2007, MNRAS, 381, 543
Williams, R. J., Quadri, R. F., Franx, M., van Dokkum, P., & Labbé, I. 2009, ApJ, 691, 1879
Witt, A. N., & Gordon, K. D. 2000, ApJ, 528, 799
Wuyts, S., Förster Schreiber, N. M., Lutz, D., et al. 2011, ApJ, 738, 106
Wuyts, E., Rigby, J. R., Gladders, M. D., et al. 2012, ApJ, 745, 86
Zamojski, M. A., Schiminovich, D., Rich, R. M., et al. 2007, ApJS, 172, 468
Zheng, X. Z., Bell, E. F., Papovich, C., et al. 2007, ApJ, 661, L41

Pages 15 to 18 are available in the electronic edition of the journal at http://www.aanda.org

Appendix A: The SED fitting technique

The broad band SED fitting technique is a simple approach to infer the physical properties of a galaxy, such as stellar mass, SFR, amount of dust , age of stellar populations, by statistically comparing model and observed SED. The constraints on the physical parameters depend on the wavelength range spanned by the data and their quality. The COSMOS field, for which a wealth of multi-wavelength, high signal-to-noise ratio observations exist, is thus well suited for such modeling.

To derive the physical parameters, we adopt a library of SEDs based on the synthetic stellar population code from Bruzual & Charlot (2003, hererafter BC03). We describe the star formation history either with an exponentially declining function, with e-folding time $0.01 \leq \tau \leq 15$ Gyr, or with a constant. We adopt two metallicities, subsolar and solar (i.e., $Z_{\odot}$, $0.2\ Z_{\odot}$), and the IMF of Chabrier (2003), truncated at 0.1 and 100 $M_{\odot}$. Since the maximum redshift of the galaxies in our sample is $z \sim 1.3$, we force the age of galaxies to be larger than 100 Myr, computed from the onset of SF which, in our case, corresponds to the initial burst. We also constrain the age of the galaxies not to exceed the age of the universe at any redshift. We do not adopt rising star formation histories, since these were developed to improve the SED fitting at high redshift ($z \geq 2$, see Maraston et al. 2010). The prescriptions for TP-AGB stars adopted in BC03 produces a lower near-IR luminosity for intermediate age stellar populations with respect to the prescriptions of Maraston (2005). This affects the galaxy mass-to-light ratio, and produces a difference in the stellar mass estimated with Maraston et al. (2010) of ~–0.15 dex. The dust attenuation curve encodes informations about the nature of the dust grains (sizes, chemical compositions) and the spatial distributions of dust and stars. Boquien et al. (2009) have shown the necessity of adopting a range of attenuation laws to reproduce the observed scatter in the IRX vs. β relation. In particular, they show that a gray (i.e., shallow) attenuation curve, such as the starburst curve of Calzetti et al. (2000), and a steeper curve, such the SMC extinction curve of Prevot et al. (1984), are required to span the observed distribution in the IRX versus β diagram of the starbursting and normal star-forming galaxies. Similar conclusions are reached by Ilbert et al. (2009), who find that a range of attenuation curves is required to improve the photometric redshift accuracy via the SED fitting method. For these reasons, we adopt three different attenuation curves: a starburst, SMC-like laws and a curve with a slope in between them ($\lambda^{-0.9}$). We then consider reddening excess in the range $0 \leq E(B-V) \leq 0.6$, which allows us to explore the observed color distribution of our sample (see Fig. 3).

We use LePhare code (Arnouts et al. 1999; Ilbert et al. 2006) to compute the χ^2 for each observed galaxy and the entire model library, with all the photometric passbands from 0.15 μm to 4.5 μm. The physical parameters are derived by computing the median of the marginalized likelihood for each parameter and the errors corresponding to the 68% credible region.

In Fig. A.1, we compare the instantaneous SFR and sSFR derived with the SED fitting with the "total" SFR and sSFR derived from the observed IR and UV luminosities (i.e., Eq. (1)). The mean errorbars (based on 68% errors) for the SED parameters vary between 0.2 to 0.4, as shown by the gray region on the right side of the figures.

For the bulk of the 24 μm population the SFRs are in good agreement over ~3 order of magnitude, with a dispersion lower then a factor of two. The vast majority of sources with large discrepancies is located in the region occupied by passive galaxies (red triangles) or next to it, in the "intermediate" zone (green dots) as defined in Fig. 2. For those galaxies, the SED fitting predicts a low specific SFR (i.e., $sSFR^{\rm SED} \leq 10^{-10.5}\ {\rm yr}^{-1}$). The origin of this difference may have multiple causes: an inadequate description of dust attenuation may cause the SED fitting to reject highly attenuated models and prefers models with low or no dust content and a low specific SFR. Alternatively the adopted definition of the total SFR in Eq. (1) may over-estimate the SFR, since we neglect the contribution of old stars to the

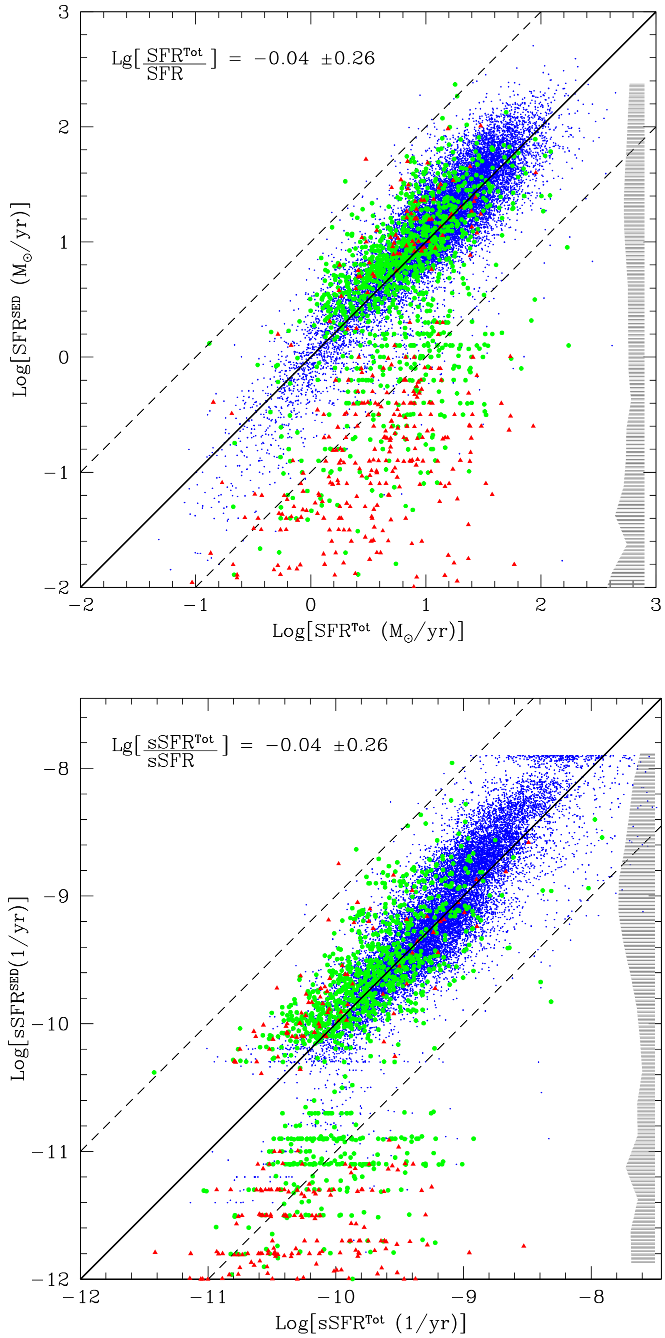


Fig. A.1. Comparison between the SFR (*top panel*) and sSFR (*bottom panel*) derived from the SED fitting and the IR+UV method (i.e., Eq. (1)). The mean errors on the SFR and sSFR estimated are shown as shaded region in the right side of the plots. The passive galaxies and the ones in the "intermediate" region of *NUVrK* diagran are shown as red triangles and green circles respectively. The mean and dispersion of the relation reported in each figure do not include the passive galaxies.

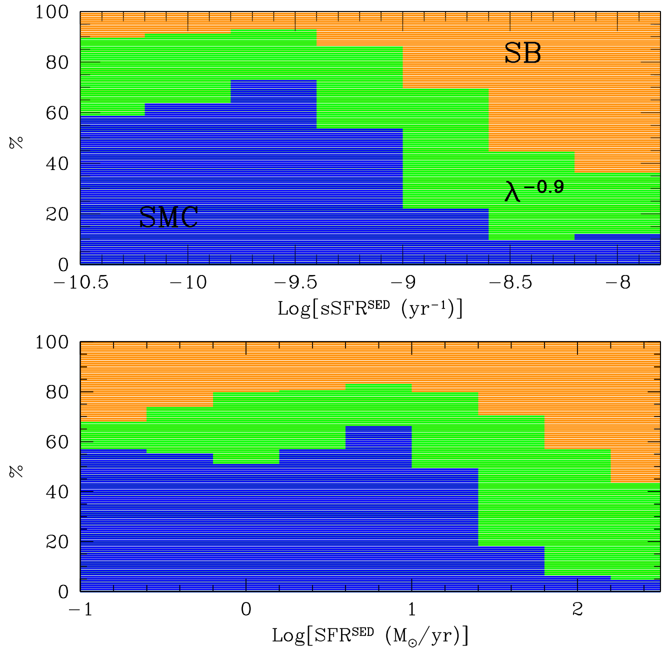


Fig. A.2. Relative contribution of the three attenuation laws (starburst [orange]; SMC [blue] and intermediate, $\propto \lambda^{-0.9}$ [green]) as a function of the specific SFR (*top figure*) and SFR (*bottom figure*) and derived from the SED modeling.

dust heating $((1-\eta)L_{\rm IR})$. This contribution is often considered as 30% ($\eta = 0.3$) for star-forming galaxies (Inoue 2005) but it can be higher for the most evolved galaxies (e.g., Cortese & Hughes 2009). However, even an extreme value of $\eta \sim 0.9$ will not reconcile the estimated SFRs with the two methods. Another possibility is that the extrapolation of the 24 μm flux into the total IR luminosity could fail for galaxies with low specific SFR, if for exemple, a warmer dust temperature is associated to the same 24 μm flux with respect to galaxies in the star-forming main sequence, as recently reported by Skibba et al. (2011); Smith et al. (2012). It is beyond the scoop of this paper to address this issue, since we focus on the star-forming galaxies, but we find that ~7% of the entire 24 μm sample is affected by this mismatch in the SFR estimates.

In Fig. A.2, we show the relative contribution of the different attenuation laws, corresponding to the best-fit templates, as a function of the SFR and specific SFR. As mentioned above, the SMC-like extinction law is favored for galaxies with low SFR and/or low sSFR while the starburst law better fits the active/starbursting galaxies with high SFR and sSFR ($sSFR \geq -9$). Our results agree with Wuyts et al. (2011) with a transition for a steeper attenuation law at $SFR \leq 20\,M_\odot\,{\rm yr}^{-1}$. The most active galaxies are consistent with a mixed distribution of the dust and star resulting in the gray attenuation law (Calzetti et al. 2000), while "normal" star-forming galaxies prefer the SMC-like attenuation consistent with a simple dust screen model. See Sect. 5 and Chevallard et al. (2013) for a purely geometric origin of different attenuation laws.

In conclusion, our analysis shows that SFRs estimated via the SED fitting technique are in good agreement with those derived from the UV+IR contribution. This validates our method to derive the $L_{\rm IR}$ from the 24 μm flux and the use of Eq. (1) as a good measure of the SFR. we have noted that Eq. (1) may become inadequate to describe the SFR in more evolved galaxies, possibly because of the presence of a larger population of old stars or the inadequate conversion of $L_{24\,\mu{\rm m}}$ to $L_{\rm IR}$. While this problem can not be easily solved, in this paper we show that we can isolate in the *NUVrK* diagram the region occupied by galaxies for which we obtain inconsistent SFR estimates with the different methods.

Appendix B: Separation between passive and active

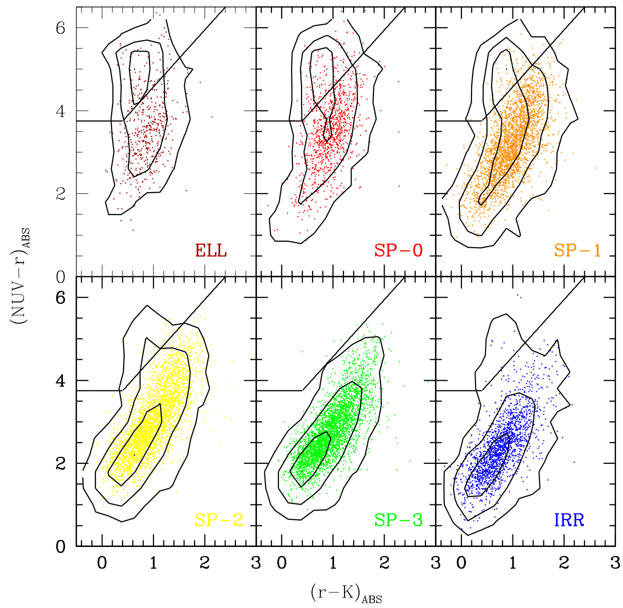


Fig. B.1. Distribution of the morphologically selected samples in the *NUVrK* diagram. The density contours (1/2, 1/10, 1/100 of the peak) refer to the whole morphological sample from Scarlata et al. (2007), while the dots refer to the 24 μm-select subsample.

As the UVJ diagram proposed by Williams et al. (2009), in Sect. 2.3 and in Fig. 2, we define a criterion based on the *NUVrK* diagram to separate the passive and star-forming galaxies. To test the validity of the above color criterion, we show the distribution in the *NUVrK* diagram of galaxies with morphological information provided by the Zurich estimator of structural type (ZEST) catalog (Scarlata et al. 2007). The ZEST classification uses three morphological classes: Early type, Disk and Irregular, with subclasses describing the degree of "irregularity" in the early-type class (i.e., 0 for regular, 1 for irregular), and the contribution of the bulge for disk galaxies (i.e., from 0 for bulge dominated galaxies, to 3 for pure, bulge-less disks). We also consider the ellipticity class for the galaxies classified as disks. This traces the galaxy inclination with an ellipticity of 0 corresponding to a face-on galaxy and up to 3 for an edge-on galaxy. The ZEST catalog includes galaxies down to $I_{\rm AB} \leq 24$ and we restrict the sample to galaxies with good flags. The distributions of the different morphological classes in the *NUVrK* diagram are shown in Fig. B.1 for the whole ZEST sample (as density contours) and for the 24 μm selected subsample (small dots). We detect a clear evolutionary sequence in this diagram, with the Irregular (IRR) and Spiral-disk dominated (Sp-3, and Sp-2) galaxies showing only blue colors, typical of active star-forming galaxies. The Spirals with a growing contribution of bulge (Sp-1 and Sp-0) and the early-types (ELL) show an increasing fraction of their population to lie in the passive region (i.e., top-left

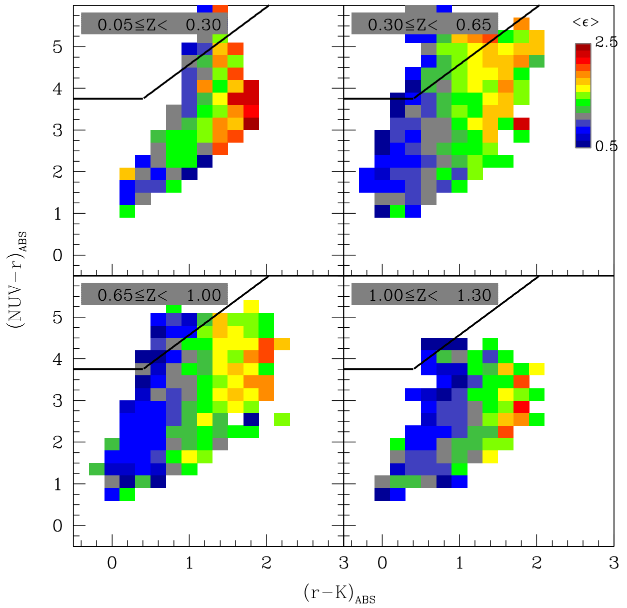


Fig. B.2. Mean value of the ellipticities (color coded) in the *NUVrK* diagram for the morphological sample from Scarlata et al. (2007).

side) of the diagram. The 24 μm sample tends to lie in the blue plums of the early type (ELL) and bulge dominated spiral (Sp-0) samples. The blue plum in the early type class could be due to some residual of star formation activities (the plum is present in the two subclasses based on the regularity criterion: ELL-0 and ELL-1).

The Spiral disk-dominated population (Sp-3, and Sp-2) extend to relatively red colors in (NUV $-$ r) and ($r-K$) (top-right part of the diagram), where high values of IRX are also observed. As discussed by Patel et al. (2011), in the UVJ diagram, the disk inclination can be indeed responsible for this extreme reddening. This is indeed supported by Fig. B.2, where we show the mean values of the ellipticity parameter for the disk population, in different redshift bins (see also the discussion in Sect. 5).

Appendix C: The average spectral energy distributions along the vector NRK

The evolution of ⟨IRX⟩ with ***NRK*** vector should be reflected in the shape of the galaxy SEDs when moving from low to high value of NRK. In Fig. C.1, we have reconstructed the averaged rest-frame SEDs in different bins of NRK, redshift and stellar mass, by using the 31 broad and medium bands available in the COSMOS catalog. They are reconstructed in logarithmic wavelength bins ($\Delta\log(\lambda) = 0.15$). Before computing the average and the dispersion in each wavelength bin, the low resolution spectrum of all the galaxies was first normalized at $\lambda \sim 0.5\,\mu$m (dashed lines). In each panel, we show the number of galaxies used to reconstruct the averaged SEDs and as an indicative value of the IRX, we show the ratio $\log(F_{\rm IR}/F_{\rm NUV})$, where the $F_{\rm IR}$ is the redshift corrected 24 μm flux of the averaged SED ($F_{24\,\mu/(1+z)}$) and $F_{\rm NUV}$ the rest-frame flux at 0.23 μm ($F_{0.23\,\mu}$). As a reference, we overplot the template of a young star-forming galaxy (SB2[4] from Ilbert et al. 2009). For the first bin $0.5 \leq NRK \leq 1$, corresponding to the bluest population in our sample, no attenuation is applied to the template. For the three other bins of NRK, we increase the amount of dust reddening to $E(B-V) = 0.2, 0.4, 0.6$, assuming a starburst attenuation law (solid black lines) and a SMC extinction law (gray dashed lines).

A strong evolution of the SED shape is observed with increasing NRK. We can quantitatively reproduce this evolution by increasing the reddening excess applied to our SB2 star-forming template. On the other hand, redshift and stellar mass appear to play a secondary role in shaping the SED properties. This provides an additional evidence that NRK is sensitive to the global budget between UV and Far-IR emission in galaxies, regardless of other properties. It also supports our choice of neglecting the dependence on the stellar mass in Eq. (3). However, in Le Floc'h et al. (in prep.) we extend our analysis to a complete mass selected sample using the stacking technique at 24, 220, 350 and 500 μm and we show that the ⟨IRX⟩ vs. NRK relation should include a corrective term based on stellar mass when considering galaxies with stellar masses $M_\star \leq 10^{9.3}\,M_\odot$.

[4] We focus only on the UV to Mid-IR part of the template with $\lambda \leq 5\,\mu$m.

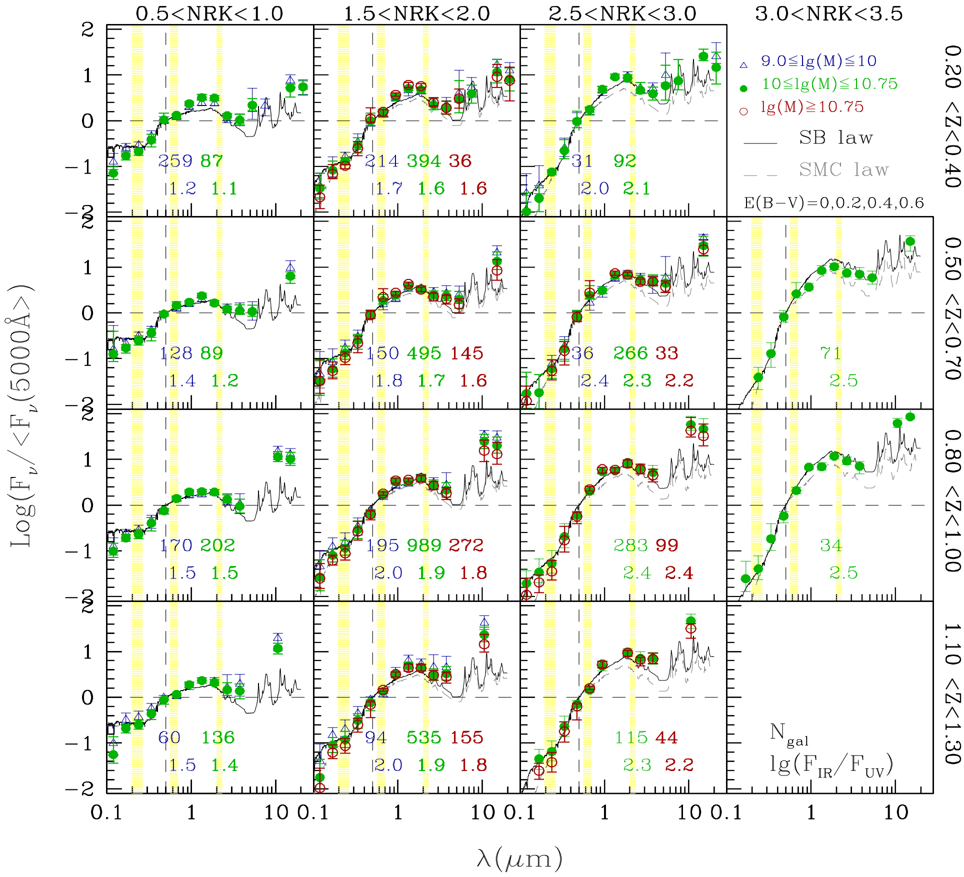


Fig. C.1. Evolution of the average rest-frame SEDs with NRK for the 24 μm-selected sample. Each SED is reconstructed by using all the available COSMOS pass-bands. Galaxies are divided in bins of NRK (columns of the figure), redshift (rows) and stellar mass (blue, green and red symbols in *each panel*). The number of galaxies for each sample is shown in each panel. We also report the logarithm of the mean ratio F_{IR}/F_{UV} as an indicative value of the IRX. In the first bin of NRK (i.e., *first column of the figure*) we overplot the dust-free template of young star-forming galaxy. In the other three bins of NRK (i.e., *three right-most columns*) we overplot the same template, but with increasing reddening excess ($E(B-V)$ = 0.2, 0.4 , 0.6, respectively), assuming the Calzetti et al. (2000) attenuation law (solid, black line) and the SMC extinction law (dashed, gray line). The three vertical bands refer to the bandwidths of the NUV, r and K filters.

Appendix D: Dependence with sSFR

As observed in Fig. 8, the difference between the IR luminosity derived from the 24 μm observations and the NRK method varies with the specific SFR. In particular, we have shown that the NRK method over-estimates L_{IR} for galaxies with low sSFR. In Fig. D.1 we compare, in different redshift bins, the predicted specific SFR ($sSFR^{NRK}$) with the reference sSFR (derived with the 24 μm, $sSFR_{tot}$), with both sSFR estimated with Eq. (1). The $sSFR^{NRK}$ tends to saturate while the true sSFR keeps declining toward galaxies with low star formation activity, and this deviation varies with redshift. We propose an analytical fit which allows us to reproduce this deviation over the entire redshift range:

$$\log\left(sSFR^{NRK^{COR}}\right) = -9.5 + 1.1\left[\log\left(10^{y_0} - 10^{a(z)}\right) - b(z)\right] \quad \text{(D.1)}$$

where, $y_0 = \log(sSFR^{NRK}) + 9.5$, $a(z) = -1.4 + 0.8\,z$ and $b(z) = -0.1\,(1-z)$ and z is the redshift. By mean of this equation, we can define a sSFR threshold below which the NRK method over-estimates the SFR by a given factor. By chosing a factor 2, the redshift evolution of this threshold can be simply described by $sSFR^{NRK}{}_{lim}(z) = -10.6 + 0.8\,z$. These thresholds are shown as shaded yellow regions in Fig. D.1.

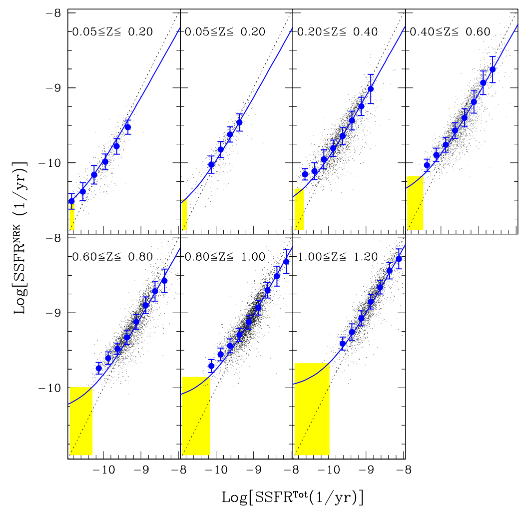


Fig. D.1. The predicted specific SFR ($sSFR^{NRK}$) vs. the reference sSFR ($sSFR_{tot}$) for the star-forming population in the local SWIRE (top left panel) and COSMOS (other panels) samples. The mean and the sigma per bin of $sSFR_{tot}$ are shown as blue symbols and the solid lines show the analytical fit described in the text. The yellow area corresponds to the limit where $SFR^{NRK} > 2 \times SFR_{tot}$.

4.3.3 Article 3 : The VIPERS MLS. II. Diving with massive galaxies in 22 square degrees since z=1.5

A&A 590, A103 (2016)
DOI: 10.1051/0004-6361/201527294

Astronomy & Astrophysics

The VIPERS Multi-Lambda Survey

II. Diving with massive galaxies in 22 square degrees since $z = 1.5$

T. Moutard[1], S. Arnouts[1], O. Ilbert[1], J. Coupon[2], I. Davidzon[1,3], L. Guzzo[4], P. Hudelot[5], H. J. McCracken[5], L. Van Werbaeke[6], G. E. Morrison[7,8], O. Le Fèvre[1], V. Comte[1], M. Bolzonella[3], A. Fritz[9], B. Garilli[9], and M. Scodeggio[9]

[1] Aix Marseille Université, CNRS, LAM – Laboratoire d'Astrophysique de Marseille, 38 rue F. Joliot-Curie, 13388 Marseille, France
e-mail: thibaud.moutard@lam.fr
[2] Astronomical Observatory of the University of Geneva, ch. d'Ecogia16, 1290 Versoix, Switzerland
[3] INAF – Osservatorio Astronomico di Bologna, via Ranzani 1, 40127 Bologna, Italy
[4] INAF – Osservatorio Astronomico di Brera, via E. Bianchi 46, 23807 Merate/via Brera 28, 20122 Milano, Italy
[5] Sorbonne Université, UPMC Univ. Paris 06, et CNRS, UMR 7095, IAP, 98bis Bd Arago, 75014 Paris, France
[6] Department of Physics and Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, V6T 1Z1, BC, Canada
[7] Institute for Astronomy, University of Hawaii, Honolulu, HI, 96822, USA
[8] Canada-France-Hawaii Telescope, Kamuela, HI, 96743, USA
[9] INAF - Istituto di Astrofisica Spaziale e Fisica Cosmica (IASF) Milano, via Bassini 15, 20133 Milano, Italy

Received 2 September 2015 / Accepted 27 March 2016

ABSTRACT

We investigate the evolution of the galaxy stellar mass function and stellar mass density from redshift $z = 0.2$ to $z = 1.5$ of a $K_s < 22$-selected sample with highly reliable photometric redshifts and over an unprecedentedly large area. Our study is based on near-infrared observations carried out with the WIRCam instrument at CFHT over the footprint of the VIPERS spectroscopic survey and benefits from the high-quality optical photometry from the CFHTLS and ultraviolet observations with the GALEX satellite. The accuracy of our photometric redshifts is $\sigma_{\Delta z/(1+z)} < 0.03$ and 0.05 for the bright ($i_{AB} < 22.5$) and faint ($i_{AB} > 22.5$) samples, respectively. The galaxy stellar mass function is measured with ~760 000 galaxies down to $K_s \sim 22$ and over an effective area of ~22.4 deg^2, the latter of which drastically reduces the statistical uncertainties (i.e. Poissonian error and cosmic variance). We point out the importance of carefully controlling the photometric calibration, whose effect becomes quickly dominant when statistical uncertainties are reduced, which will be a major issue for future cosmological surveys with EUCLID or LSST, for instance. By exploring the rest-frame (NUV$-r$) vs. ($r - K_s$) colour-colour diagram with which we separated star-forming and quiescent galaxies, (1) we find that the density of very massive $\log(M_*/M_\odot) > 11.5$ galaxies is largely dominated by quiescent galaxies and increases by a factor 2 from $z \sim 1$ to $z \sim 0.2$, which allows for additional mass assembly through dry mergers. (2) We also confirm the scenario in which star formation activity is impeded above a stellar mass $\log(\mathcal{M}^\star_{SF}/M_\odot) = 10.64 \pm 0.01$. This value is found to be very stable at $0.2 < z < 1.5$. (3) We discuss the existence of a main quenching channel that is followed by massive star-forming galaxies, and we finally (4) characterise another quenching mechanism that is required to explain the clear excess of low-mass quiescent galaxies that is observed at low redshift.

Key words. galaxies: evolution – galaxies: luminosity function, mass function – galaxies: star formation – galaxies: distances and redshifts – galaxies: photometry – galaxies: statistics

1. Introduction

The measurement of the stellar mass function (SMF) is a powerful statistical tool for tracing the stellar mass assembly or galaxy growth over cosmic time. Galaxy formation models rely on the well established ΛCDM cosmological framework that governs the growth of the dark matter structures and the less well understood baryonic physics at play inside the dark matter haloes (gas accretion, minor or major merging, star formation activity, feedback mechanisms, etc.). The shape of the galaxy SMF compared to the expected halo mass function provides valuable information about the physical processes acting at the low- and high-mass ends of the mass function (Silk & Mamon 2012).

A decade ago, early deep extragalactic surveys have revealed that the average stellar mass density decreased gradually (the integrated form of the SMF) from $z \sim 3$ to $z \sim 0$ (e.g. Dickinson et al. 2003; Fontana et al. 2003). This trend is now confirmed up to redshift $z \sim 8$ (Song et al. 2015) and is consistent with a hierarchical build-up of the cosmic structures. Later on, larger surveys have measured the evolution at high redshift of the galaxy bimodality, the well-known separation between star-forming and quiescent galaxies observed in the local Universe (Baldry et al. 2004; Moustakas et al. 2013). They found that the bimodality was already in place at $z \sim 1$ with the quiescent galaxies dominating the massive end of the SMF and the star-forming galaxies dominating its low-mass end (Bundy et al. 2006; Borch et al. 2006). This quiescent population had its main build-up epoch between $z = 2$ and $z = 1$, where the stellar mass density increased by a factor 10 (Cirasuolo et al. 2007; Arnouts et al. 2007), while only a factor 2 increase is observed from $z = 1$ to $z = 0$ (Bell et al. 2004; Faber et al. 2007). According to the hierarchical scenario, such an early formation epoch of the quiescent population was not a problem as long as the stars formed before this in smaller units and galaxies continued to assemble their masses at a later time (through dry merging phases, e.g. De Lucia et al. 2006). This is a natural support of the star formation downsizing picture proposed by Cowie et al. (1996), where the onset of star formation begins earlier for most massive galaxies than for lower mass galaxies (see also Gavazzi & Scodeggio 1996). However, the models predict a continuous increase of stellar mass for these massive galaxies with cosmic time (e.g. De Lucia & Blaizot 2007), which is challenged by the last measurements of the SMF, where the massive end does not show significant evolution from $z = 0$ up to redshift $z \sim 1$ (e.g. Marchesini et al. 2009; Ilbert et al. 2013; Muzzin et al. 2013; Moustakas et al. 2013; Mortlock et al. 2015), suggesting a mass assembly downsizing.

The predominance of quiescent galaxies at the massive end (e.g. Baldry et al. 2012; Moustakas et al. 2013; Ilbert et al. 2013) supports the idea that the star formation activity is preferentially impeded in galaxies above a given stellar mass or a given dark matter halo mass, if we assume a stellar-to-halo mass relationship (e.g. Coupon et al. 2015). A wide variety of quenching mechanisms have been proposed to explain the star formation quenching in massive galaxies, such as major mergers (Barnes 1992), virial shock heating (Kereš et al. 2005), or radio-AGN feedback (Croton et al. 2006; Cattaneo et al. 2006) in massive haloes.

Several studies have emphasised the role played by the environment for the colour-bimodality and star-formation quenching in the local Universe (Hogg et al. 2003; Kauffmann et al. 2004; Baldry et al. 2006; Haines et al. 2007). Mechanisms such as ram-pressure stripping, in which the gas is expelled from the galaxy (Gunn & Gott 1972), or strangulation, in which the cold gas supply is heated and then halted (Larson et al. 1980), can be invoked as environmental quenching mechanisms. We emphasise that strangulation processes can either be linked to environment (e.g. when a galaxy enters the hot gas of a cluster) or to peculiar evolution (e.g. when a the radio-AGN feedback stops the cold gas infall). The latest measurements of the quiescent SMFs reveal an upturn at the low-mass end in the local Universe (Baldry et al. 2012; Moustakas et al. 2013), whose build-up is observed at higher redshift (Drory et al. 2009; Tomczak et al. 2014). This upturn in the low-mass end for quiescent galaxies could be associated to environmental quenching according to Peng et al. (2010, 2012), while Schawinski et al. (2014) suggested a fast process consistent with major merging. Constraining the quenching timescale at different masses might therefore help to highlight the quenching mechanisms.

Until recently, the above conclusions were mostly based on deep galaxy surveys such as GOODS (Giavalisco et al. 2004), VVDS (Le Fèvre et al. 2005), COSMOS (Scoville et al. 2007), and DEEP2 (Newman et al. 2013), which are perfectly suited to provide the global picture of the galaxy stellar mass assembly over a wide range of redshifts. However, given their small angular coverage (they explore a rather small volume below $z < 1$), they can be particularly sensitive to statistical variance (i.e. cosmic variance) at low redshift. This is particularly crucial for the very rare galaxies at the high-mass end of the exponentially declining SMF, and it has been claimed that its apparent lack of evolution may be dominated by observational uncertainties (Fontanot et al. 2009; Marchesini et al. 2009).

A first attempt to constrain the density evolution of the high-mass galaxies at $z < 1$ has been performed by Matsuoka & Kawara (2010). They combined the SDSS southern strip (York et al. 2000) and UKIDSS/LAS survey (Lawrence et al. 2007) over a total area of ~55 deg^2. They observed a mild-to-high increase of the number density of massive galaxies $10^{11-11.5}$ $M_\odot$/$10^{11.5-12.}$ $M_\odot$ with a corresponding drop of the fraction of star-forming galaxies in this stellar mass range from $z \sim 1$ to $z \sim 0$. While subject to large uncertainties in their photometric redshifts, stellar mass estimates, and reliability of the separation into quiescent and star-forming galaxies, this first result suggested that massive galaxies ($M_* > 10^{11}$ $M_\odot$) evolve since $z \sim 1$.

Moustakas et al. (2013) estimated the SMF between $0 < z < 1$ over an area of ~5.5 deg^2 using PRIMUS, a low-resolution prism survey (for galaxies with $i_{\rm AB} \leq 23$; Coil et al. 2011). The wealth of multi-wavelength information from deep ultraviolet (GALEX satellite) to mid-infrared (*Spitzer*/IRAC) photometry allowed them to derive accurate stellar masses and a reliable separation between active and quiescent populations. Their SMF measurements confirmed the modest change in the number density of the massive star-forming galaxies ($M_* \geq 10^{11}$ $M_\odot$), leaving little room for mergers, but observed a significant drop (50%) of the fraction of active star-forming galaxies since $z \sim 1$ that is in contrast with the classical picture, in which the star-forming population remains constant across cosmic time. Another major spectroscopic sample is provided by the VIMOS Public Extragalactic Redshift Survey (VIPERS; Guzzo et al. 2014), whose first ~50 000 galaxies down to $i_{\rm AB} = 22.5$ over an area of 10.3 deg^2 have recently been released (PDR1, Garilli et al. 2014). Using the PDR1 combined with CFHTLS photometry and the same ultraviolet (UV) and near-infrared (NIR) data that we used here, Davidzon et al. (2013) produced the most reliable overall measurement of the high-mass end of the SMF in between $0.5 < z < 1.3$ to date. The VIPERS SMF shows to high precision that the most massive galaxies had already assembled most of their stellar mass at $z \sim 1$, but that a residual evolution is still present. However, as discussed in Davidzon et al. (2013), although these two studies use spectroscopic redshifts, multi-wavelength information, and a large area, they disagree slightly concerning the general amplitude of the SMF. These discrepancies might be due to differences in the stellar mass estimates, for example, or to selection effects that are not fully accounted for. It highlights how subtle effects become crucial and can introduce significant systematic errors when statistical uncertainties are reduced so drastically.

In this paper we exploit the broad photometric coverage assembled over the footprint of VIPERS to build a unique multi-wavelength photometric sample covering more than 22 deg^2 down to $K_s < 22$, as part of the VIPERS-Multi Lambda Survey (VIPERS-MLS; see Moutard et al. 2016). We benefit of the synergy with the VIPERS spectroscopic survey by using the PDR-1 data to compute reliable photometric redshifts, and we derive stellar masses for 760 000 galaxies out to $z = 1.5$; This allows us to obtain a new estimate of the SMF that (a) has greater control over the low-mass slope because of the $i < 23.7$ / $K_s < 22$ depth of our sample for extended sources (more than 1 mag deeper in the i-band than VIPERS); (b) extends over a wider redshift range than VIPERS, from $z = 0.2$ out to $z = 1.5$; (c) is less affected by the cosmic variance because the effective area is doubled with respect to the VIPERS PDR-1 used in Davidzon et al. (2013) (we cover nearly the entire footprint of the final VIPERS survey and avoid the 30% area loss that is due to the detector gaps in VIPERS); (d) suffers a reduced Poisson error because our sample is ten times larger in the common redshift range; and (e) can be studied separately for star-forming and quiescent objects, which means that the quenching channels that characterise the massive galaxies up to $z = 1.5$ can be explored, as well as the low-mass galaxies at low redshift.

The paper is organised as follows. In Sect. 2 we describe our photometric and spectroscopic dataset. The photometric redshifts and galaxy classification are presented in Sect. 3, the stellar mass estimates in Sect. 4. We detail the measurements of the galaxy SMFs and the associated uncertainties in Sect. 5, where we also point out the effect of the photometric absolute calibration in the new generation of large surveys. We present the evolution of the stellar mass function and stellar mass density in Sect. 6. Finally, we discuss our results and their effects on the quenching channels in Sect. 7.

Throughout this paper, we use the standard cosmology ($\Omega_{\rm m}$ = 0.3, Ω_Λ = 0.7 with H_0 = 70 km s^{-1} Mpc^{-1}). Magnitudes are given in the AB system (Oke 1974). The galaxy stellar masses are given in units of solar masses ($M_\odot$) for a Chabrier (2003) initial mass function (Chabrier IMF).

2. Data description

The observations and the data reduction are described in detail in the companion paper (Moutard et al. 2016) and are briefly summarised below.

2.1. Optical CFHTLS photometry

The CFHTLS[1] is an imaging survey performed with the MegaCam[2] camera in five optical bands, u, g, r, i, and z. It covers ∼155 deg^2 over four independent fields with sub-arcsecond seeing (median ∼0.8") and reaches a 80% completeness limit of u ∼ 24.4, g ∼ 24.7, r ∼ 24.0, i/y ∼ 23.7, and z ∼ 22.9 for extended sources in AB system. We emphasise that the y-band refers to the new i-band filter, in accordance with the CFHTLS notation. We have used the y-band response curve in our analysis when appropriate, but we refer to the "i" filter term regardless of whether it was observed with the i or y filter.

In this work we use the W1 (+02$^{\rm h}$18$^{\rm m}$00$^{\rm s}$ −07°00$^{\rm m}$00$^{\rm s}$) and W4 (+22$^{\rm h}$13$^{\rm m}$18s +01°19m00$^{\rm s}$) fields. Two independent photometric catalogues have been released to the community: the 7th and final release (noted T0007[3]) of the CFHTLS produced by Terapix[4], and the release from the CFHT Lensing Survey team (CFHTLenS[5]). Both catalogues are based on the same raw images. The AstrOmatic software suite[6] has been used to generate the mosaic images (SWARP, Bertin et al. 2002) and to extract the photometric catalogues (SExtractor, Bertin & Arnouts 1996). The two releases differ in several points, however.

- In T0007, detection is based on $gri - \chi^2$ images, while the galaxies in CFHTLenS are i-detected.
- A point spread function (PSF) homogenisation is implemented in CFHTLenS (see Hildebrandt et al. 2012) to improve the colour estimates. In practice, the PSF is homogenised across the field of view for each filter and degraded in all filters to the one with the highest PSF.
- A new photometric calibration has been applied to the T0007 release. While the previous releases and the CFHTLenS release rely on Landolt standard stars (see Erben et al. 2013), T0007 is based on the spectrophotometric tertiary standards from the Super Novae Legacy Survey (SNLS; see the procedure described in Regnault et al. 2009). In brief, each tile of the CFHTLS-Wide is re-observed (with short exposures) during stable photometric conditions and bracketed by two observations of the CFHTLS-Deep field containing the SNLS tertiary standards.

[1] http://www.cfht.hawaii.edu/Science/CFHTLS/
[2] http://www.cfht.hawaii.edu/Instruments/Imaging/Megacam/
[3] http://terapix.iap.fr/cplt/T0007/doc/T0007-doc.html
[4] http://terapix.iap.fr/
[5] http://www.cfhtlens.org/
[6] http://www.astromatic.net/

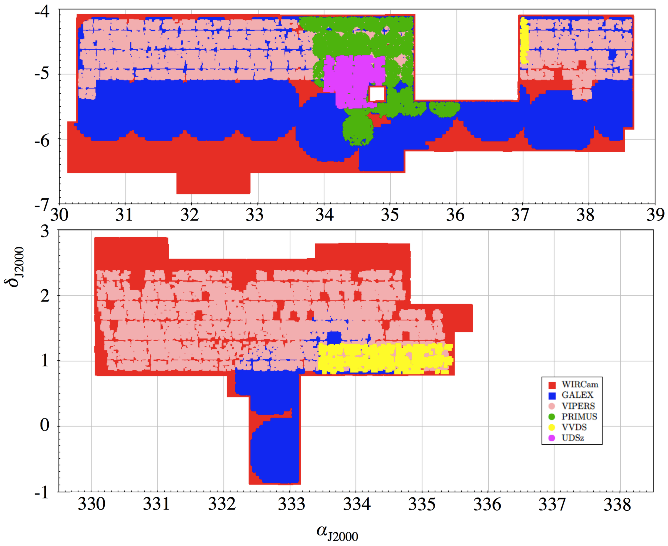


Fig. 1. Footprints of the WIRCam K_s-band (red layout and background) and GALEX NUV/FUV (blue circles) observations in the CFHTLS W1 (*top*) and W4 (*bottom*) fields. The regions covered by VIPERS (pink), PRIMUS (green), VVDS (yellow) and UDSz (magenta) are over-plotted. The SDSS-BOSS redshifts are distributed over the entire survey.

The difference in the calibration scheme of the two releases affects the final photometry. A comparison of the magnitudes for point-like sources between the T0007 and CFHTLenS releases reveals systematic offsets that are significantly larger than the expected uncertainties. These offsets are reported in Table 1 (column Δmag). We emphasise that the differences listed in this table are entirely due to the new calibration scheme established for T0007. The procedure used by the T0007 release allows transferring the percent level accuracy of the SNLS photometric calibration to the entire CFHTLS-Wide survey. For this reason, we use the T0007 catalogue as reference in this paper. However, we also perform the complete analysis with the CFHTLenS catalogue to discuss the effect of such differences in the photometric absolute calibration.

2.2. WIRCam K_s photometry

We conducted a NIR K_s-band follow-up of the VIPERS fields with the WIRCam instrument at CFHT (Puget et al. 2004). The layout of the observations is shown in Fig. 1 (red background). We covered a total area of ∼27 deg^2 with an integration time per pixel of 1050 seconds. The image quality is very homogeneous, with an average seeing of all the individual exposures of $\langle IQ \rangle = 0.6'' \pm 0.09$. The data have been reduced by the Terapix team[7] and the individual images were stacked and resampled on the pixel grid of the CFHTLS-T0007 release (Hudelot et al. 2012). The photometry was performed with `SExtractor` in dual-image mode with a $gri - \chi^2$ image as the detection image and the same settings as those adopted for the T0007 release.

[7] http://terapix.iap.fr/

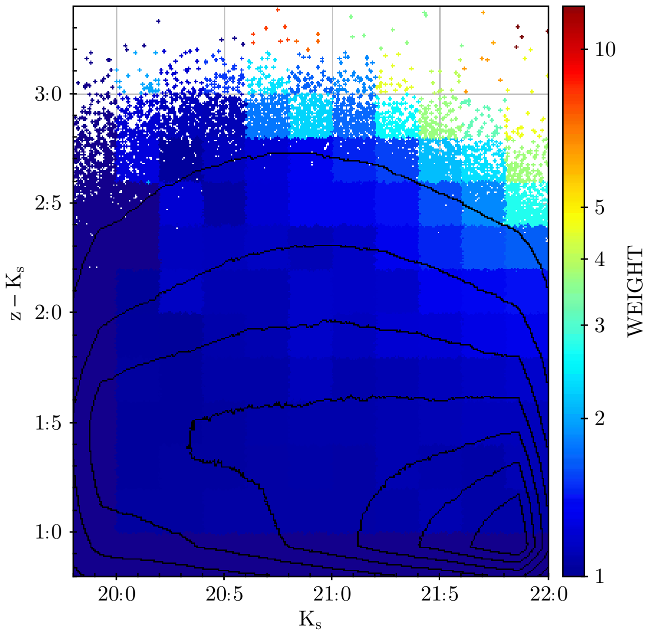


Fig. 2. Colour−magnitude weight map used for our statistical analysis. It takes the missing objects in the $K_s < 22$-limited sample into account. These objects are missed because the *gri*-detection was used to extract the K_s fluxes. Weights are multiplicative. This map is restricted to galaxies with redshift $z \leq 1.5$ (cf. Sect. 3) and the contours outline the galaxy density distribution.

The images reach a depth of $K_s = 22$ at $\sim 3\sigma$. The completeness reaches 80% at $K_s = 22$; this was determined from a comparison with the deepest surveys UKIDSS Ultra-Deep Survey(UDS, Lawrence et al. 2007) and VIDEO (Jarvis et al. 2013) in overlapping regions. Because the primary optical detection is based on the $gri - \chi^2$ image, we may miss the reddest high-redshift galaxies. To account for this possible bias, we measured our source incompleteness as a function of magnitude, K, and colour, $(z - K)$, with all the sources detected in the deep VIDEO survey. We derived a colour-magnitude weight map that we show in Fig. 2 and use in the remaining paper as multiplicative weighting for our statistical analyses. We refer to the companion paper, Moutard et al. (2016), for a complete description of the method that was used to build this weight map.

2.3. GALEX photometry

When available, we made use of the UV deep-imaging photometry from the GALEX satellite (Martin & GALEX Team 2005). We only considered the observations from the Deep Imaging Survey (DIS), which are shown in Fig. 1 as blue circles (∅ ∼ 1.1°). All the GALEX pointings were observed with the near-ultraviolet (NUV) channel with exposure times of $T_{\rm exp} \geq 30$ ks. Far-ultraviolet (FUV) observations are available for ten pointings in the central part of W1.

The large PSF of GALEX (*FWHM* ∼ 5″) means that source confusion becomes a great problem in the deep survey. To extract the UV photometry, we used a dedicated photometric code, `EMphot` (Conseil et al. 2011) that will be described in a separate paper (Vibert et al., in prep.). In brief, EMphot uses the stamps in the *u*-band (here the T0007 release) as priors, and they are then convolved by the GALEX PSF to produce a simulated image. The scaling factors to be applied to each *u*-band prior is obtained by simultaneously maximising the likelihood between the observed and the predicted fluxes for all the sources in tiles of a few square arc-minutes. The uncertainties on the UV fluxes account for the residuals in the simulated or observed image. A typical depth of NUV ∼ 24.5 at $\sim 5\sigma$ is observed over the entire survey. The NUV observations cover part of the WIRCam area with ∼10.8 and 1.9 deg^2 in the W1 and W4 fields, respectively.

2.4. Final photometric catalogue

The catalogue of sources comes from the T0007 release and is based on detection in the $gri - \chi^2$ image. As mentioned above, the same procedure was applied to the K_s images. Following Erben et al. (2013) and Hildebrandt et al. (2012), we used the T0007 isophotal apertures for the photometry to estimate the colours. The apertures are smaller than the Kron-like apertures (Kron 1980), which provides less noisy colours and leads to an improved photometric redshift accuracy (Hildebrandt et al. 2012). We also confirmed this with our large spectroscopic dataset (see below), which is especially relevant for faint sources ($i' > 23.5$).

To derive galaxy physical properties, we need to know the total flux in all wavelengths. Therefore, we rescaled the isophotal flux to the Kron-like flux, $m^f_{\rm total} = m^f_{\rm ISO} + \delta_{\rm m}$, by adopting a unique factor, $\delta_{\rm m}$, for each source to preserve the colours. $\delta_{\rm m}$ is the weighted mean of the individual scaling factors, $\delta^f_{\rm m}$, and is defined as $\delta_{\rm m} = \sum_f \delta^f_{\rm m} w^f / \sum_f w^f$, with $f = g, r, i, K_s$, and w^f its associated errors[8]. Finally, the GALEX photometry, which corresponds to the total flux measurement (i.e. model PSF photometry) was added in the same way as to the optical and NIR magnitudes.

We here limit the catalogue to galaxies brighter than $K_s < 22$. The catalogue includes a total of ∼1.3 millions sources over an area of ∼27.1 deg^2, which drops to one million sources over ∼22.4 deg^2 after applying the masks provided by the CFHTLenS team.

2.5. Spectroscopic sample

Our WIRCam survey has been designed to cover the VIMOS Public Extragalactic Survey (VIPERS; Guzzo et al. 2014) that is carried out with the VIMOS spectrograph and therefore provides many high-quality spectroscopic redshifts. We also added a compilation of the best-quality spectra from the VVDS survey ($I_{\rm AB} \leq 24$, Le Fèvre et al. 2013), the $K < 23$ limited UKIDSS spectroscopic Ultra Deep Survey (UDSz, Bradshaw et al. 2013; McLure et al. 2013), the low-resolution spectra ($\lambda/\Delta\lambda \sim 40$) from the PRIsm MUlti-object Survey (PRIMUS, $i \sim 23$, Coil et al. 2011), and the bright-limited ($i < 19.9$) spectroscopic survey BOSS from the SDSS (Baryon Oscillation Spectroscopic Survey, Dawson et al. 2013). The $K_s < 22$ spectroscopic sample we used is presented in detail in the companion paper (Moutard et al. 2016).

We selected only the most secure spectroscopic redshifts, which means confidence levels above 95% for high-resolution surveys and $\sigma < 0.005$ (8% of outliers with $\delta z/(1 + z) > 5\sigma$) for PRIMUS best redshifts. When they are available, the redshift measurements from VIPERS were used. Otherwise, the measurements from the deepest high-resolution spectra were favoured. In total, we assembled a $K_s < 22$-limited sample

[8] For the CFHTLenS catalogue the scaling factor is computed only in *i* band. Only the final magnitude is available in the public CFHTLenS catalogue of Erben et al. (2013).

of 45, 951 high-quality spectroscopic redshifts to calibrate and measure the accuracy of our photometric redshifts over the unmasked area of the survey (we refer to the companion paper for more details).

3. Photometric redshifts

3.1. Photometric redshift measurement

The photometric redshifts were computed with the spectral energy distribution (SED) fitting code LE PHARE (Arnouts et al. 2002; Ilbert et al. 2006), using the templates of Coupon et al. (2015). The new templates are based on the Ilbert et al. (2009) library of 31 empirical templates from Polletta et al. (2007), complemented by 12 star-forming templates from the Bruzual and Charlot stellar population synthesis models of 2003 (Bruzual & Charlot 2003, hereafter BC03). These templates were optimised to be more representative of the VIPERS spectroscopic sample (for more details we refer to Coupon et al. 2015).

The extinction was added as a free parameter with a reddening excess $E(B-V) < 0.3$ following different laws: Prevot et al. (1984), Calzetti et al. (2000), and a dusty Calzetti curve including a bump at 2175Å. No extinction was allowed for SEDs redder than Sc. The extinction law of Prevot et al. (1984) was used for templates redder than SB3 templates (see Ilbert et al. 2009) and the law of Calzetti et al. (2000) for bluer templates.

Finally, any possible difference between the photometry and the template library was corrected for by LE PHARE according to the method described in Ilbert et al. (2006). In brief, in each band the code tracks a systematic shift between the predicted magnitudes at known redshift and the observed magnitudes. Since our observation area is divided into 47 tiles of $\lesssim 1$ deg^2 with the relative calibration varying from tile to tile, we performed a tile-by-tile colour optimisation. We used the median offset over all the tiles when there were not enough galaxies with spectroscopic redshift in the tile ($N_{\rm gal}^{\rm spec} \leq 100$) available, which was the case in 12 tiles. We stress that the corrections were computed to better fit the colours and are therefore relative. We normalised the median offset on the K_s-band because the NIR fluxes are the same (see Sect. 2). The median relative offsets thus calculated for each photometric band of the T0007 and CFHTLenS catalogues can be found in Table 1, with the associated tile-to-tile deviation estimates (namely, the normalised median absolute deviation, NMAD). The difference between the T0007 and CFHTLenS relative offsets is consistent with the difference Δmag. In other words, we retrieved the shift between the two absolute photometric calibrations through the relative offsets computed with LE PHARE. This safety check confirms that the colour optimisation achieved with LE PHARE absorbs the uncertainties that are linked to the photometric calibration.

3.2. Accuracy and precision of photometric redshifts

The comparison between our photometric redshifts and the corresponding spectroscopic redshifts for our $K_s < 22$ -limited sample is shown in Fig. 3. Using the NMAD to define the scatter[9], we find $\sigma_{\Delta z/(1+z)} \sim 0.05$ for faint ($i > 22.5$) galaxies, while the scatter reaches $\sigma_{\Delta z/(1+z)} \sim 0.03$ for the bright ($i < 22.5$) galaxies. Our photo-z outlier rate[10] is $\eta = 1.2\%$ and $\eta = 9\%$ for corresponding bright and faint samples, respectively (see Fig. 3, top panels, lower right corners).

Although the spectroscopic sample has been assembled to be as representative as possible, it is not as deep as the photometric sample. Aiming to correct this effect, we computed estimators that are weighted with respect to the i-band distribution of the photometric sample. By using these weighted estimators (marked with an orange w in Fig. 3, top panels), we obtained an accuracy $\sigma^w_{\Delta z/(1+z)} \sim 0.03$ for bright ($i < 22.5$) galaxies with an outlier rate of $\eta^w = 1.4\%$, and $\sigma^w_{\Delta z/(1+z)} \sim 0.07$ and $\eta^w = 16.4\%$ for faint ($i > 22.5$) galaxies in our $K_s < 22$ -limited sample.

Even though the T0007 and CFHTLenS calibrations differ, the photometric redshifts obtained in both cases agree well[11] and their accuracies are similar. This is expected from the colour corrections described in Sect. 3.1, which absorb the differences between the two calibrations.

Finally, based on Fig. 3, we can define a range of reliable redshifts up to $z = 1.5$, with $\sigma_{\Delta z/(1+z)}(z) < 0.1$. The highest redshift bin that we consider, namely between $z = 1.1$ and $z = 1.5$, is characterised by the weighted accuracy $\sigma^w_{\Delta z/(1+z)} \sim 0.08$ and weighted outlier rate $\eta^w \sim 20\%$.

3.3. Star and galaxy classification

Being able to separate galaxies and stars is crucial in our sample, especially for the W4 field, which is close to the Galactic plane and therefore highly populated by stars. Garilli et al. (2008) have found that more than 32% of the objects in the VVDS-Wide survey are stars. This is a pure $i < 22.5$-selected spectroscopic survey lying in the CFHTLS W4 field. Aiming to better control the type of the objects that we select as galaxies without compromising the completeness of our sample, we performed a classification based on three different diagnostics. Our classification is presented in detail in the companion paper (Moutard et al. 2016) and is summarised below.

– First, we used the maximum surface brightness versus magnitude (hereafter $\mu_{\rm max} - m_{\rm obs}$) plane where bright point-like sources are well separated from galaxies (see Bardeau et al. 2005; Leauthaud et al. 2007).

Table 1. T0007 – CFHTLenS photometric offsets (Δmag) obtained by comparing point-like sources in the two catalogues and relative corrections obtained with LE PHARE to optimise the photometric redshifts.

Filter	Δmag^*	LE PHARE corrections T0007	LE PHARE corrections CFHTLenS
FUV	–	0.102 ± *0.070*	0.084 ± *0.079*
NUV	–	0.054 ± *0.055*	0.022 ± *0.065*
u	−0.013 ± *0.052*	0.075 ± *0.031*	0.087 ± *0.042*
g	0.071 ± *0.053*	0.028 ± *0.019*	−0.053 ± *0.016*
r	0.038 ± *0.052*	0.022 ± *0.019*	−0.024 ± *0.005*
i	0.066 ± *0.045*	0.013 ± *0.015*	−0.055 ± *0.009*
y	0.048 ± *0.051*	0.008 ± *0.009*	−0.042 ± *0.013*
z	0.148 ± *0.054*	0.087 ± *0.027*	−0.063 ± *0.015*
K_s	–	0.0 ± *0.016*	0.0 ± *0.019*

Notes. Relative corrections are given using the K_s-band as reference (NIR data are identical). $^{(*)}$ $m_{\rm T07} - m_{\rm LenS}$.

[9] $\sigma_z = 1.48\, median(\,|z_{\rm spec} - z_{\rm phot}|/(1+z_{\rm spec}))$.
[10] η is the percentage of galaxies with $\Delta z/(1+z) > 0.15$.
[11] $z^{T07}_{\rm phot} - z^{\rm LenS}_{\rm phot} = -0.008 \pm 0.048$ for $0.2 < z^{T07}_{\rm phot} \leq 1.5$.

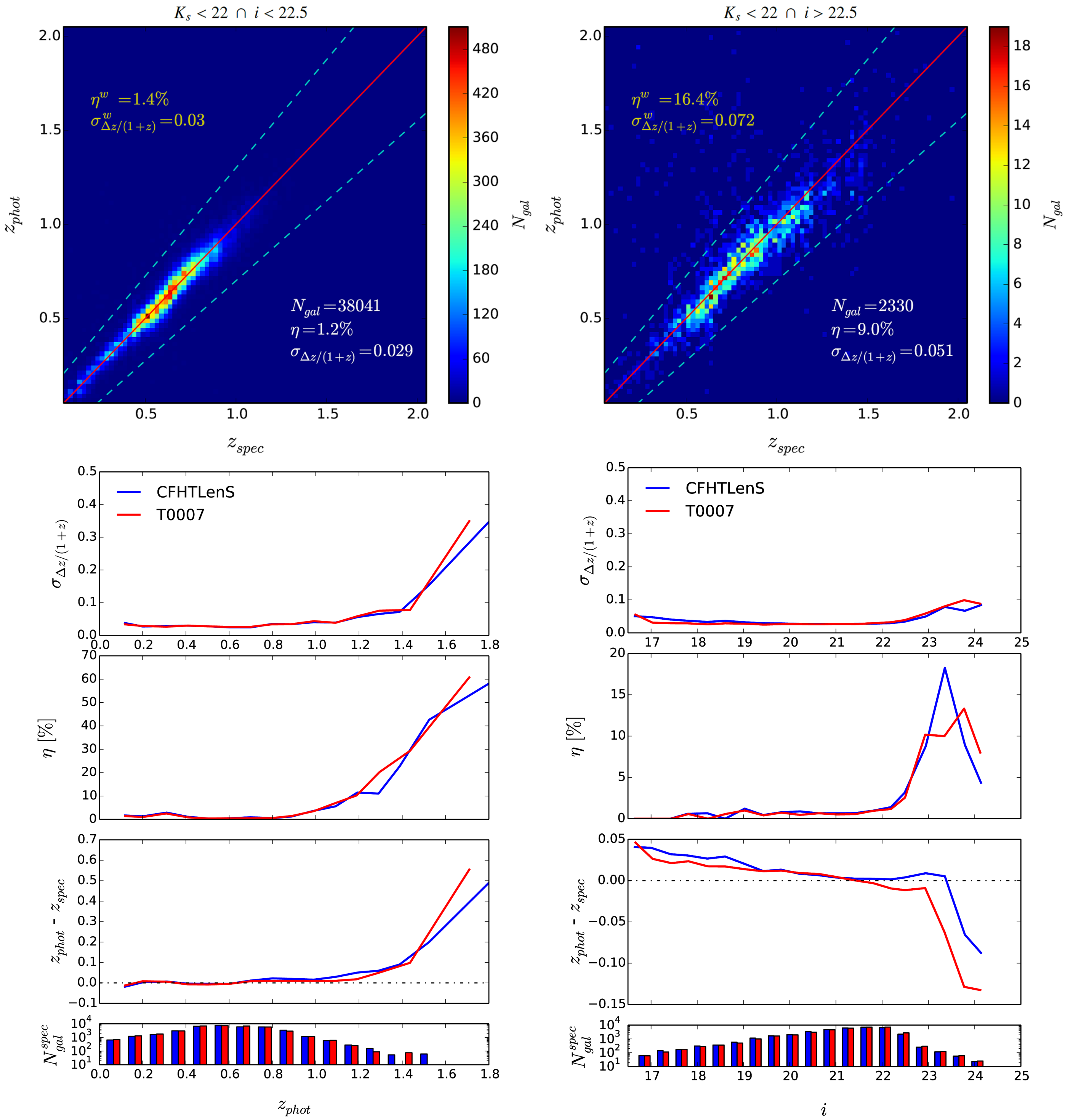


Fig. 3. Photometric redshift accuracy of our $K_s < 22$-limited sample. *Top*: T0007 photometric redshift as a function of spectroscopic redshift for bright ($i < 22.5 \cap K_s < 22$) and faint ($i > 22.5 \cap K_s < 22$) galaxies. The dashed lines delimit the $\sigma_{\Delta z/(1+z)} \leq 0.15$ area, outside which photo-z measurements are considered as outliers. The accuracy estimators written in the upper left corners are weighted with respect to the i-band distribution of our photometric sample (see Sect. 3.2). *Bottom*: dispersion, outlier rate, bias, and spectroscopic redshift number ($N^{\rm spec}_{\rm gal}$) as a function of photometric redshift (*left*) and i-magnitude (*right*), using the T0007 (blue) and CFHTLenS (red) optical photometry.

– Secondly, we compared the reduced χ^2 obtained with galaxy templates described in Sect. 3.1 and a representative stellar library (based on Pickles 1998). An object can be defined as a star when its photometry is better fitted by a stellar spectrum.
– Finally, we used the $g-z/z-K_s$ plane (equivalent to the BzK plane of Daddi et al. 2004) to isolate the stellar sequence and imposed that a star belong to this colour region. This sine qua non condition enabled us to catch faint stars while preventing us from losing faint compact galaxies.

We also identified a sample of QSOs (Type-1 AGNs) as point-like sources lying on the galaxy side of the BzK diagram. Dominated by their nucleus, the emission of these AGN galaxies is currently poorly linked to their stellar mass. However, they represent less than 0.5% of the objects, and we removed them from our sample without compromising its completeness.

All the objects that were not defined as stars or QSOs were considered as galaxies. We verified on a sample of 1241 spectroscopically confirmed stars that we caught 97% of them in this way, while we kept more than 99% of our spectroscopic galaxy sample. With this selection we finally found and removed ~8% and ~19% of objects at $K_s < 22$ for W1 and W4, respectively, outside the masked area.

4. Stellar mass estimation

4.1. Method

Stellar mass, M_*, and the other physical parameters were computed by using the stellar population synthesis models of Bruzual & Charlot (2003) with Le Phare. As in Ilbert et al. (2013), the stellar mass corresponds to the median of the stellar mass probability distribution (PDF_{M_*}) marginalised over all other fitted parameters. Two metallicities were considered ($Z = 0.008$ and $Z = 0.02$ i.e. $Z_\odot$) and the star formation history declines exponentially following $\tau^{-1}e^{-t/\tau}$ with nine possible τ values between 0.1 Gyr and 30 Gyr as in Ilbert et al. (2013).

The importance of the assumed extinction laws for the physical parameter estimation has been stressed in several recent studies, for example, by Ilbert et al. (2010) and Mitchell et al. (2013) for the stellar masses, or by Arnouts et al. (2013) for the star formation rate (SFR). We considered three laws and a maximum dust reddening of $E(B-V) \leq 0.5$: the Prevot et al. (1984), the Calzetti et al. (2000), and an intermediate-extinction curve (see Arnouts et al. 2013, for more details). As in Fontana et al. (2006), we imposed a low extinction for low-SFR galaxies ($E(B-V) \leq 0.15$ if age/$\tau > 4$). The emission-line contribution was taken into account following an empirical relation between UV and line fluxes (Ilbert et al. 2009).

Using a method similar to Pozzetti et al. (2010), we based our estimate of the stellar mass completeness limit, $M_{\rm lim}$, on the distribution of the lowest stellar mass, $M_{\rm min}$, at which a galaxy could have been detected given its redshift. For our sample, which is limited at $K_s < 22$, $M_{\rm min}$ is given by

$$\log(M_{\rm min}) = \log(M_*) + 0.4\,(K_s - 22). \tag{1}$$

We then considered the upper envelope of the $M_{\rm min}$ distribution. In each redshift bin, $M_{\rm lim}$ is defined by the stellar mass at which 90% of the population have $M_* > M_{\rm min}$. We show the resulting stellar mass completeness limits (open circles) as a function of redshift in Fig. 4 over the M_* distribution for our $K_s < 22$-limited sample.

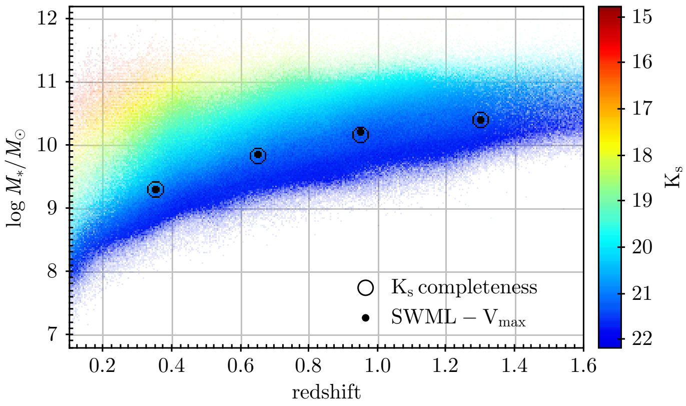


Fig. 4. Stellar mass versus redshift for our $K_s < 22$-limited sample of galaxies. The black open circles represent the stellar mass completeness limit computed with the K_s completeness limit $M_{\rm lim}$ according to Eq. (1). The black dots represent the mass at which the $V_{\rm max}$ and the SWML SMF estimators diverge (see below Sect. 5.1).

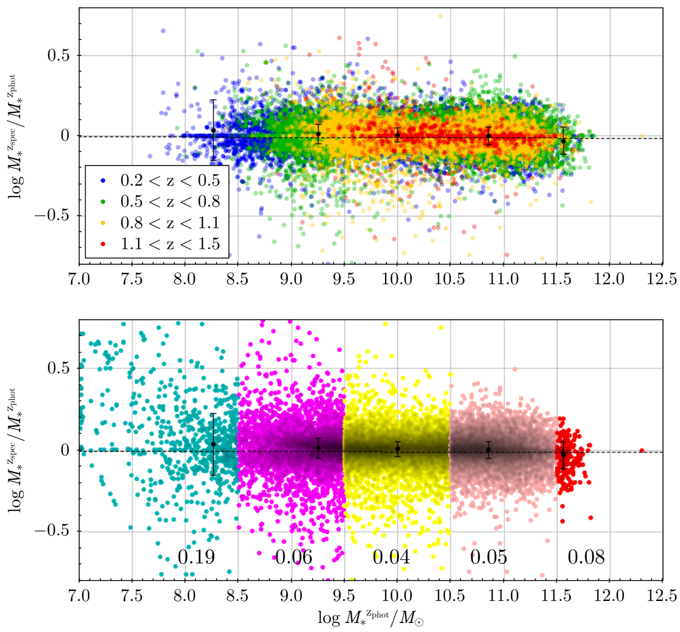


Fig. 5. Redshift contribution to the stellar mass uncertainty as a function of the stellar mass. The uncertainty is computed from the 1σ dispersion of the ratio $M_*(z_{\rm spec})/M_*(z_{\rm phot})$ in the spectroscopic sample after removing photo-z outliers. In the *top panel*, the distribution is shown in four redshift bins, while in the *bottom panel* it is shown in five stellar mass bins. The error bars correspond to the dispersion reported in each bin of M_*, while the dashed line is the linear regression associated with the whole sample.

4.2. Stellar mass error budget

In this section, we quantify the uncertainties associated with the stellar masses, which will be propagated into the error budget of the stellar mass functions. The first to be considered is the uncertainty in the flux measurements. The photon noise is taken into account by the Le Phare code during the χ^2 SED fitting procedure (rescaling and model selection), where it returns the 68% confidence interval enclosed in the probability distribution function marginalised on the stellar mass (PDF_{M_*}).

The second source of error is introduced by the photometric redshift uncertainty, which is not included in the PDF_{M_*}. One way to measure its effect is to compare the stellar masses derived with the photometric and spectroscopic redshifts. We emphasise that this analysis is probably limited by the completeness of our spectroscopic sample. By contrast, it is powerful when used to reflect all the photo-z error contributions (quality of the photometry and representativity of the templates). The difference between the two mass estimates is shown in Fig. 5 as a function of stellar mass $M_*^{z_{\rm phot}}$. In the top panel, we show the difference in four redshift bins between $z = 0.2$ and $z = 1.5$. No dependence with redshift is observed. The linear regression of the whole sample, plotted as a black dashed line, also suggests that it is not mass dependent. The bottom panel shows the $M_*^{z_{\rm phot}}/M_*^{z_{\rm spec}}$ dispersion in five stellar mass bins and reveals a median dispersion of 0.06 dex, with a maximum of 0.19 dex at low mass.

We then defined the resulting mass uncertainty as the sum in quadrature of all contributions:

$$\sigma_{\rm M} = \sqrt{\sigma_{\rm fit}^2 + \sigma_z^2}. \quad (2)$$

However, we have to keep in mind that the stellar mass estimation relies on the numerous assumptions made when we generate our SED templates. For example, Maraston (2005) has shown that a different treatment of the thermally pulsing asymptotic giant branch (TP-AGB) phase in the SSP can lead to a global shift in the stellar mass estimation[12]. Ilbert et al. (2010) showed that the use of the Salpeter (1955) IMF instead of the Chabrier IMF decreases the stellar masses by 0.24 dex. These systematic shifts are therefore not expected to affect the conclusions of our study. Mitchell et al. (2013) also pointed out the potential effect of the assumed dust attenuation on the stellar mass estimation[13]. As presented in the previous section, we considered three different extinction laws. This allows a higher diversity of possible values for dust attenuation, which is expected to limit the bias that may affect our stellar mass estimation.

4.3. Effect of the CFHTLS absolute calibration

As shown in Sect. 2.1, the absolute photometric calibrations of the T0007 and CFHTLenS magnitudes differ by more than 0.05 mag on average. Even if the T0007 were significantly improved in its calibration, we compare the photometric redshifts and the stellar masses computed with both catalogues blindly to quantify the effect of these offsets.

As seen in Sect. 3.2, the colour corrections applied during the photometric redshift computation allows us to obtain very similar photo-z despite the offset between their calibrations. However, these corrections are a combination of the photometry and the SEDs used to calculate photo-z. As described in Sect. 3.1, the templates used for photo-z are different from those used for the masses. Consequently, we did not apply the photo-z colour corrections with the BC03 templates. The differences in the T0007 and CFHTLenS photometries thus directly affect the stellar mass estimation.

Figure 6 presents these differences in the redshift bins $0.2 < z < 0.5$, $0.5 < z < 0.8$, $0.8 < z < 1.1$ and $1.1 < z < 1.5$. The difference between the stellar masses obtained with the T0007 ($M_*^{\rm T07}$) and those obtained with the CFHTLenS ($M_*^{\rm LenS}$) is stellar mass dependent. On average, this systematic difference can reach ± 0.1 dex at low redshift ($z < 0.8$). At higher redshift, we do not observe a systematic difference between the two stellar mass catalogues since we used the same $K_{\rm s}$-band calibration. Even if the object-by-object $M_*^{\rm LenS}$ to $M_*^{\rm T07}$ ratio is characterised by a mean offset that never exceeds 0.2 dex, the comparison of the T0007 and CFHTLenS number counts above the mass completeness limit reveals different shapes around $M_* \sim 10^{11}\,M_\odot$, notably at $z < 0.8$. This suggests a significant effect on the SMF massive end at low redshift, as we discussed below.

5. Measuring the stellar mass functions

To compute the galaxy stellar mass function, we selected a sample of ~760 000 galaxies at $K_{\rm s} \le 22$ over an effective area of 22.38 deg^2. According to what we discussed in Sect. 3.2, we restricted our analysis to the range $0.2 \le z \le 1.5$, where we combined reliable redshifts and large volumes.

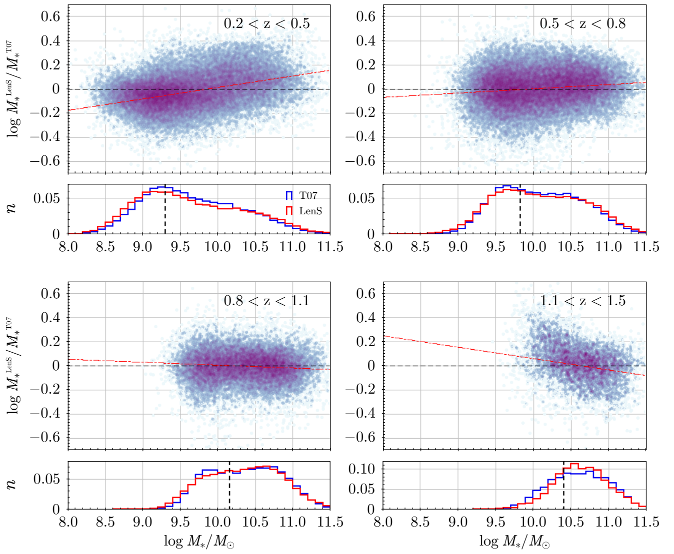


Fig. 6. Differences between the stellar mass obtained with T0007 and CFHTLenS, $M_*^{\rm T07}$ and $M_*^{\rm LenS}$ at different redshifts: the object-by-object $M_*^{\rm LenS}/M_*^{\rm T07}$ ratio versus $M_*^{\rm T07}$ in the *upper panel*, where the red dashed line is the linear regression, and the $M_*^{\rm T07}$ (blue) and $M_*^{\rm LenS}$ (red) normalised number counts in the *lower panel*, where the vertical black dashed line represents the mass completeness limit.

The galaxy stellar mass function was derived with the tool ALF (Ilbert et al. 2005), which provides three non-parametric estimators: $V_{\rm max}$ (Schmidt 1968), SWML (the step-wise maximum likelihood; Efstathiou et al. 1988), and C^+ (Zucca et al. 1997). The $V_{\rm max}$ estimator is most widely used because of its simplicity. The $1/V_{\rm max}$ is the inverse sum of the volume in which each galaxy was observed. The $V_{\rm max}$ is the only estimator that is directly normalised. The SWML (Efstathiou et al. 1988) determines the SMF by maximising the likelihood of observing a given stellar mass – redshift sample. The C^+ method overcomes the assumption of a uniform galaxy distribution, as is the case when using the $V_{\rm max}$[14]. As described in Ilbert et al. (2015), these estimators diverge below a stellar mass limit that should correspond to the limit calculated in Sect. 4. In Fig. 4 we verify that the $V_{\rm max}$ and SWML estimators (black dots) are consistent with our $K_{\rm s}$-based stellar mass completeness limit (black open circles). We used the colour-magnitude weight map shown in Fig. 2 to correct the SMF for the potential incompleteness described in Sect. 2.2. In the remainder of this study, we work with stellar masses $M_* > M_{\rm min}$ where all the non-parametric estimators agree.

5.1. Measurements by type and field

To separate quiescent and star-forming galaxies, we used the rest frame (NUV − r)° versus (r − K)° diagram (hereafter NUVrK) presented by Arnouts et al. (2013), which is based on the method introduced by Williams et al. (2009). Figure 8 presents the galaxy distribution in the NUVrK diagram for several redshift bins. This optical-NIR diagram allows us to properly separate red dusty star-forming galaxies from red quiescent

[12] Pozzetti et al. (2007) and Ilbert et al. (2010) estimated on offset of ~0.14 dex between the stellar masses of BC03 and Maraston (2005).

[13] Mitchell et al. (2013) estimated that the stellar mass can be underestimated by up to 0.6 dex by assuming the Calzetti et al. (2000) for massive galaxies.

[14] For more details about these estimators, we refer to Ilbert et al. (2005) and Johnston (2011).

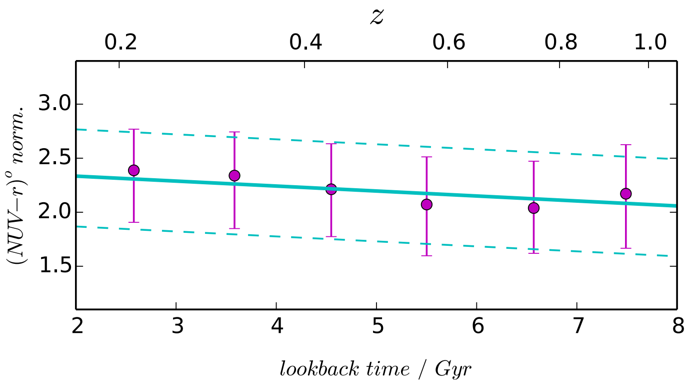


Fig. 7. Cosmic evolution of the (NUV − r)° normalisation. The dots represent the position of the minimum density along the (NUV − r)° axis across cosmic time, while the bars are defined by the extreme values that delimit the NUVrK green valley. The solid line is the linear fit and the dashed lines represent their mean upper and lower envelopes.

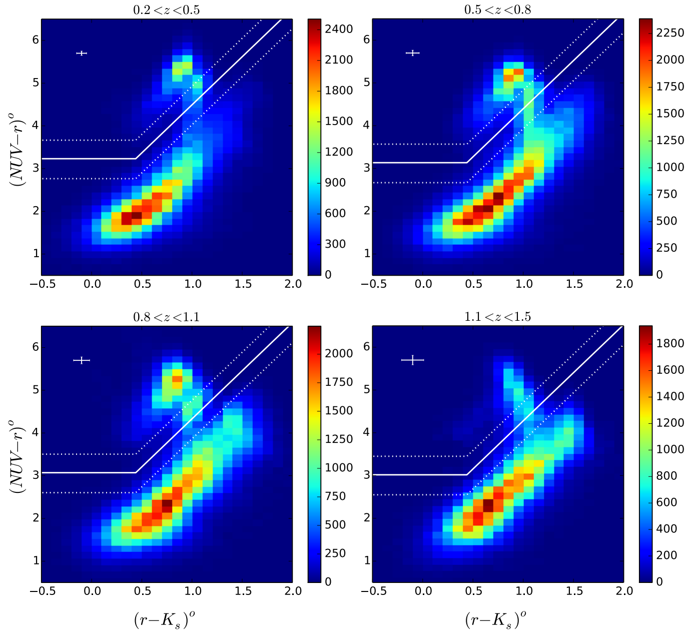


Fig. 8. Star-forming and quiescent galaxy selection in the NUVrK diagram. The colour code shows the galaxy density. The averaged colour uncertainties (based on the observed photometric errors) are shown in the upper left corner of each panel. The binning used for the density map is tuned to match the typical uncertainties at $0.2 < z < 0.5$. The solid line represents the mean selection of quiescent galaxies in a given redshift bin. The dotted lines represent the two extreme selections delimiting the green valley.

ones. Edge-on spirals are clearly identifiable, as is illustrated by the morphological study of the NUVrK diagram at low redshift presented in the companion paper (Moutard et al. 2016).

When we computed the rest-frame colours, we adopted the procedure described in Appendix A.1 of Ilbert et al. (2005) to minimise the dependency of the absolute magnitudes to the template library. An absolute magnitude at λ^0 was derived from the apparent magnitude in the filter passband that was the closest from $\lambda^0 \times (1+z)$ to minimise the k-correction term, except when the apparent magnitude had an error above 0.3 mag, to avoid too noisy colours. The small break[15] in the red clump is artificial and is an effect of the template discretisation, when our procedure used to limit the template dependency fails because of the low signal-to-noise ratio measurements (here due to the intrinsic low rest-frame NUV emission of quiescent galaxies)[16].

As shown in Fig. 8, by following the low-density valley of the NUVrK diagram (the so-called green valley), the selection of quiescent galaxies can be defined with the general form

$$[\,(\mathrm{NUV}-r)^\circ > B_2\,] \cap [\,(\mathrm{NUV}-r)^\circ > A\,(r-K_\mathrm{s})^\circ + B_1\,]. \quad (3)$$

A, B_1, and B_2 are three parameters to be adjusted in each redshift bin, as suggested by Ilbert et al. (2015) and Mortlock et al. (2015), because of the global ageing of the galaxy population.

In the four redshift bins, the slope A of Eq. (3) seems to be constant, with a typical value of $A = 2.25$. By projecting the galaxy distribution in a plane perpendicular to the axis of slope A[17], we clearly distinguish the red and blue clouds as two normal distributions that we fitted by two Gaussians. We define B_1 as the position where the two Gaussians intersect.

In Fig. 7 we show the evolution of B_1 as a function of the look-back time (t_L). By assuming a linear relation between B_1 and cosmic time, we derive $B_1(t_\mathrm{L}) = -0.029\, t_\mathrm{L} + 2.368$ in our highest precision redshift range (0.2 – 1.0). Assuming that B_2 evolves as B_1, we empirically set $B_2(t_\mathrm{L} = 2.5\,Gyr) = 3.3$, and we find $B_2(t_\mathrm{L}) = B_1(t_\mathrm{L}) + 1.004$. We can write our selection of quiescent galaxies as

$$[\,(\mathrm{NUV}-r)^\circ > 3.372 - 0.029\, t_\mathrm{L}\,] \cap [\,(\mathrm{NUV}-r)^\circ > 2.25\,(r-K_\mathrm{s})^\circ + 2.368 - 0.029\, t_\mathrm{L}\,]. \quad (4)$$

All the galaxies that are not selected as quiescent are considered to be star forming. In Fig. 8 the separations between quiescent and star-forming galaxies are shown as white solid line. We also define the green valley as the region around minimum B_1, reaching 10% of the peak of the red Gaussians, as shown by the white dotted lines. We consider these limits as possible systematic uncertainties when discussing the evolution of the quiescent and star-forming SMFs.

Figure 9 presents the global (black), star-forming (blue), and quiescent (red) galaxy SMFs for the two fields separately (W1: dot and W4: cross) in four redshift bins. The sample consists of 481 518 galaxies over 14.43 deg^2 in W1 and 268 010 galaxies over 7.96 deg^2 in W4. The error bars shown in the upper sub-panels reflect only the Poissonian contributions. The SMF comparison between the two fields agrees within the errors. In the lower sub-panels, we plot the stellar mass uncertainty by type, σ_M, defined in Sect. 4.2, as function of the stellar mass. First, σ_M decreases exponentially with stellar mass, as already noted in previous studies (e.g. Grazian et al. 2015). We can then fit the $\sigma_\mathrm{M}(M_*)$ relation with a power law (Fig. 9 sub-panels, dashed lines). Secondly, the size of our galaxy sample allows for very small relative Poissonian errors down to densities of around $\sim 10^{-5}$–10^{-6} Mpc^{-3} even if we split by type and field. The cosmic variance contribution in the budget of the errors that affects our SMF measurement is therefore expected to be small, as discussed in the next section.

[15] It is important to keep in mind that the NUVrK diagram is particularly stretched along the $(r - K_\mathrm{s})^\circ$ axis.

[16] We note that this effect of discretisation from the templates is smoothed if we use a high number of templates, such as with the BC03 library. However, this smoothing is somehow artificial since the NUV part is not better constrained in practice. We verified with BC03 that the SMF of quiescent galaxies is not significantly affected by the set of templates we used to compute absolute magnitudes.

[17] We selected the galaxies in the range $0.4 < (r - K_\mathrm{s})^\circ < 0.9$ to avoid the objects in transition.

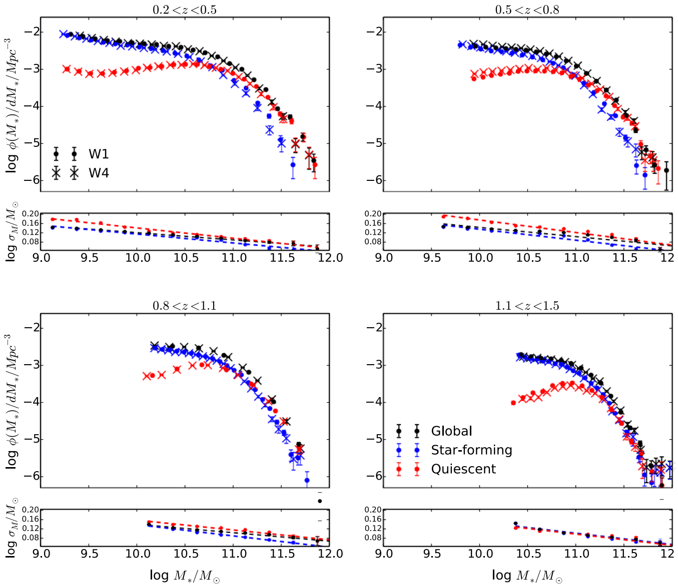


Fig. 9. Galaxy SMF in the fields W1 (dots) and W4 (crosses) for the global (black), star-forming (blue), and quiescent (red) populations in four redshift bins (*upper sub-panels*). The error bars reflect only the Poissonian contribution, while the corresponding mass uncertainties are shown in the *lower sub-panels*. Only SMF points above the stellar mass completeness are plotted.

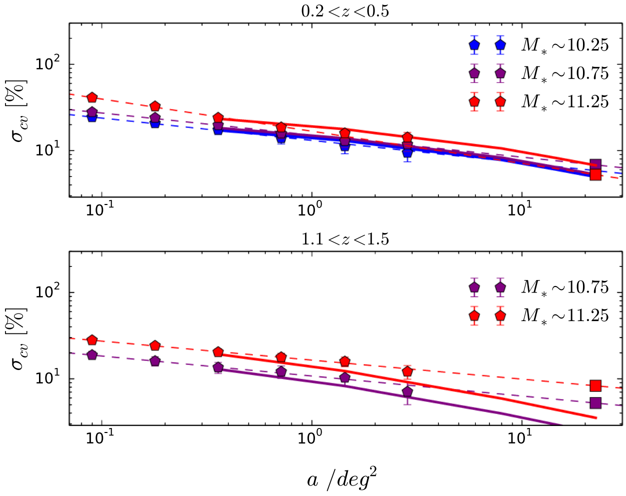


Fig. 10. Cosmic variance as a function of the effective observed area for three stellar mass bins. The dashed lines correspond to the linear fit of the empirical cosmic variance estimates plotted with pentagons. The squares locate the extrapolated cosmic variance estimate for our entire survey. The solid lines show the corresponding theoretical estimates computed according to Moster et al. (2011).

5.2. SMF uncertainties

In this section, we describe the error budget associated to our SMF measurements. All the contributions to the SMF uncertainties are expressed as a function of the stellar mass and redshift. In addition to the stellar mass and Poissonian errors already mentioned, the large-scale density inhomogeneities represent a source of uncertainty. This cosmic variance is known to represent a fractional error of 15–25% at the high-mass end ($M \geq 10^{11}\ M_\odot$) in the COSMOS survey and of around 20–50% in narrower pencil-beam surveys, generally dominating the error budget.

Following the procedure discussed by Coupon et al. (2015), we investigated the contribution of the cosmic variance in our sample by dividing our survey into N patches of equal areas. Since the effective surface can change from one patch to another, every patch was weighted according to its unmasked area. For a given observed area, we computed the number density dispersion N times over $(N-1)$ patches by discarding a different patch every time. We then considered the mean number density dispersion over the N measurements as our internal estimate of the cosmic variance for a given effective area and the dispersion around the mean as an error estimate of the cosmic variance.

In Fig. 10 we plot our cosmic variance estimate $\sigma_{\rm cv}$ in the redshift bins [0.2, 0.5] and [1.1, 1.5], considering three stellar mass bins from $M_* = 10^{10}\ M_\odot$ up to $M_* = 10^{11.5}\ M_\odot$ (with blue, purple and red dots, respectively) and for mean effective areas ranging from $a \simeq 0.1$ to $a \simeq 2.8$ deg^2. The relation of cosmic variance – area is well fitted by a power-law with $\sigma_{\rm cv}(a) = 10^\beta a^\alpha$ (shown as dashed lines). To estimate the cosmic variance that affects our entire survey, we extrapolated the relations up to $a = 22$ deg^2, shown as squares.

For comparison, we also show the cosmic variance predicted for the same redshift and stellar mass bins (triangles) by using the code `getcv` (Moster et al. 2011). Our internal cosmic variance estimate ($\sigma_{\rm cv}$) and the predicted one agree remarkably well up to our observed areas of $a = 2$ deg^2. For larger areas, the two estimates diverge slightly for high-mass ($M_* > 10^{11}\ M_\odot$) galaxies at $z < 0.5$, where we slightly underestimate $\sigma_{\rm cv}$ with respect to the theoretical prediction. We have to stress that the Moster et al. (2011) procedure is optimised for pencil-beam surveys of areas $a < 1$ deg^2. At $z > 0.5$, the theoretical estimators always predict a cosmic variance lower than our own extrapolation. By using our internal estimate, we therefore adopt a conservative approach.

Finally, the last source of error that we need to consider is that of the stellar mass uncertainty defined in Sect. 4.2. To do so, we generated 200 mock catalogues with perturbed stellar masses according to the expected $\sigma_{\rm M}$ (which includes the photometric redshift uncertainties and the photon noise, Eq. (2)) and measured the 1σ dispersion in the density Φ of the reconstructed SMFs that we refer to as $\sigma_{\Phi,\rm M}$.

At the end, the error of the stellar mass function is the quadratic sum of all the contributions discussed above and is defined as

$$\sigma_{\rm tot} = \sqrt{\sigma_{\rm cv}^2 + \sigma_{\rm poi}^2 + \sigma_{\Phi,\rm M}^2}. \quad (5)$$

5.3. Importance of photometric calibration in large surveys

As mentioned in Sect. 4.3, a mean offset of ~0.06 mag in the optical absolute photometric calibration (cf. Δmag in Table 1) can change the stellar mass estimate by 0.1 dex. In the top panel of Fig. 11 we show the difference between the two SMFs measured with the T07 and CFHTLens photometries, $\Delta\Phi_{\bf calib} = \Phi_{\rm LenS} - \Phi_{\rm T07}$. This difference is normalised by the total statistical error discussed in the previous section, $\sigma_{\rm tot}$. In general, $\Delta\Phi_{\bf calib} > \sigma_{\rm tot}$ in our survey (solid lines), which means that the SMF variation induced by the calibration offsets is several times larger than the uncertainty of our SMF. It even reaches 5 $\sigma_{\rm tot}$ at low redshift ($0.2 < z < 0.5$; blue solid line), where the stellar mass is essentially driven by the optical photometry.

In contrast, by considering a subsample of 2 deg^2 (dashed lines), we find that $|\Delta\Phi_{\bf calib}| \lesssim \sigma_{\rm tot}^{\rm 2deg}$ (green shaded area). This means that the SMF variation driven by the calibration offsets is

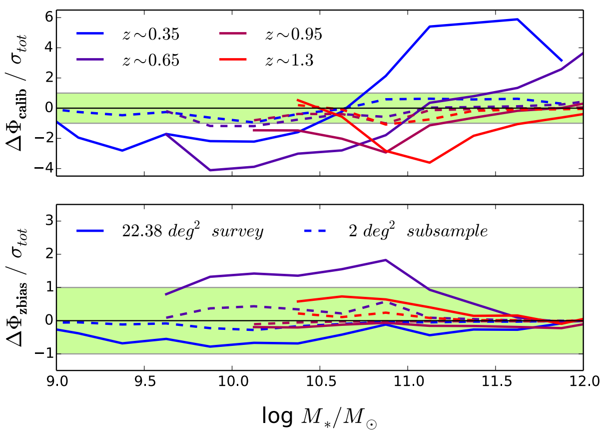


Fig. 11. Ratio between the systematic stellar mass function difference and the total statistical error, $\Delta\Phi/\sigma_{tot}$, as a function of the stellar mass and in four redshift bins. We consider $\Delta\Phi_{calib}$ (*top panel*) and $\Delta\Phi_{zbias}$ (*bottom panel*), the systematics coming from the absolute photometric calibration and from the photometric redshift bias (cf. Sect. 3.1), respectively. The green shaded areas show the region where $|\Delta\Phi| \leq \sigma_{tot}$.

smaller than the other uncertainties affecting the SMF in a 2 deg^2 survey (i.e. the variation is contained within the error bars). In other words, we cannot see the variation that is due to the calibration because it is hidden by other sources of uncertainties (Poissonian and cosmic variance). The systematic differences that are due to the T0007-CFHTLenS photometric offsets can therefore be neglected in a 2 deg^2 survey, while in surveys of 20 deg^2 and more, we reach a regime where the systematic uncertainty that is due to the photometric calibration dominates the error budget.

For comparison, we also investigated another source of systematic uncertainty: the photometric redshift bias. Using the photo-z bias ($z_{bias} = z_{phot} - z_{spec}$) presented in Sect. 3.1 (see Fig. 3, lower panels), we corrected our photometric redshifts. Instead of using a global correction, we applied a photo-z bias correction for different galaxy types[18].

Similarly to $\Delta\Phi_{\mathbf{calib}}$, the difference between the stellar mass functions computed with the *corrected* photometric redshifts and with the original ones ($\Delta\Phi_{zbias}$) is shown in the lower panel of Fig. 11. The effect of the photo-z bias on the SMF measurement is much weaker than the effect of the photometry. The SMF differences induced by the photo-z bias as measured in our sample are largely dominated by the statistical uncertainties in a 2 deg^2 subsample. Moreover, in the entire 22 deg^2 survey, the difference can only be detected at $z \sim 0.65$, while $|\Delta\Phi_{zbias}| < 2\,\sigma_{tot}$.

Given the limited amplitude of its effect on our sample, the photo-z bias can be neglected in our study. By contrast, the SMF variations that are due to the difference in photometry stress the need of carefully controlling the absolute photometric calibration in large surveys. In the present study, the choice of the CFHTLS-T0007 photometry is supported by (1) the SNLS photometric calibration based on a new spectrophotometric standard for high-precision cosmology; and (2) the careful treatment by the Terapix team that enables homogeneously propagating the SNLS photometric calibration over the entire survey.

[18] We also checked this by estimating the correction with half of the spectroscopic sample and improving the photo-zs of the other half.

6. Evolution of the galaxy stellar mass function and density

As shown in Sect. 5.2, the large volume probed by our survey allows us to reduce both the cosmic variance and the Poisson uncertainties. We exploit this large volume to quantify the evolution of the galaxy SMF, especially at the high-mass end, where it is most relevant.

6.1. Evolution of the SMF

6.1.1. Comparison of the global SMF with the literature

In the upper sub-panels of the Fig. 12, we compare our global SMF measurements with the literature. Our results agree well overall with many previous studies, although some differences exist. We discuss these in this section. The error bars corresponding to our measurement (black) reflect the total error σ_{tot} defined in Eq. (5). In the lower sub-panels, we show the contribution of each error to the error budget. We show each contribution normalised by the total error σ_{tot} as a function of the stellar mass. First, we note that the Poissonian error (blue dash-dotted line) represents a minor contribution to the total SMF uncertainty up to the very high-mass end (i.e. above $10^{11.5}$ $M_\odot$). Secondly, the contribution of the cosmic variance (σ_{cv}; red dashed line) is dominant up to stellar masses around the SMF knee ($M_* < 10^{11}$ $M_\odot$). We finally note that the contribution of the stellar mass uncertainty ($\sigma_{\Phi,M}$; cyan solid line) drives the total uncertainty of the SMF high-mass end. We recall that while the Poissonian uncertainty is always taken into account in the literature, the error bars may reflect different contributions to the SMF uncertainty depending on the study considered in Fig. 12 (as specified in the caption).

The comparison with Davidzon et al. (2013) is straightforward since their observations were taken in the same two fields of the CFHTLS survey, covering an effective area of 5.34 and 4.97 deg^2 in W1 and W4, respectively. The authors derived the SMF between $z = 0.5$ and $z = 1.3$ using the VIPERS-PDR1 dataset (~50 000 galaxies), that is, the main spectroscopic sample used to calibrate our redshifts (Sect. 3.2). The work of Davidzon et al. (2013) clearly illustrates the advantages of using spectroscopic redshifts (e.g. the easier removal of stellar interlopers and QSO). However, to estimate the SMF through spectroscopic data, some difficulties need to be solved, such as the statistical weighting to account for the spectroscopic sampling rate (see Garilli et al. 2014, for more details about how these weights are computed in VIPERS). We observe a good agreement between the two SMF estimates, especially at $M_* > 10^{11}$ $M_\odot$.

The statistical uncertainties are very low in both VIPERS and our analysis, and the two surveys are additionally collected almost in the same area. Any difference is likely due to some combination of the photometric redshift uncertainty of our sample, the spectroscopic incompleteness affecting VIPERS, the adopted SED fitting method, or the photometric calibration used in VIPERS (T0005) and in our survey (T0007). However, the only significant discrepancy is observed close to the stellar mass completeness limit of VIPERS, where the measurements of Davidzon et al. (2013) are $\lesssim$0.2 dex lower. This difference at low masses could be due to some incompleteness correction that is due to the i-band selection, while our sample is K_s-band selected.

Moustakas et al. (2013) also measured the SMF by relying on the spectroscopic redshift sample of PRIMUS (~40 000 galaxies between $z = 0.2$ and 1 and cover ~5.5 deg^2

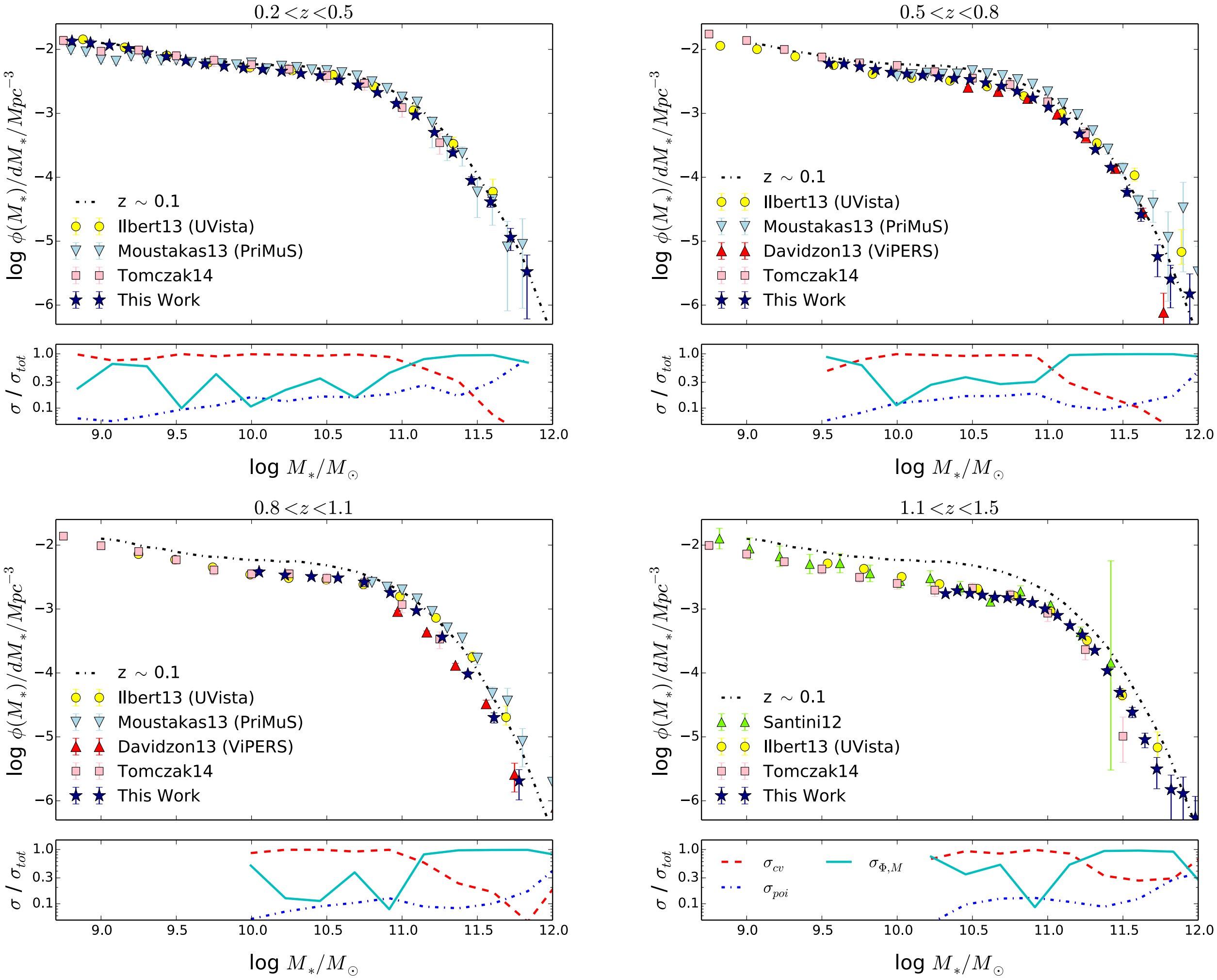


Fig. 12. Galaxy stellar mass functions (SMF) in four redshift bins. *Top sub-panels*: the SMF measured in the present study (black stars) is compared to previous measurement: Tomczak et al. (2014), pink squares; Davidzon et al. (2013), red up triangles; Moustakas et al. (2013), cyan down triangles; Ilbert et al. (2013), yellow circles; and Santini et al. (2012), green up triangles. The error bars plotted on the measures reflect different contributions to the SMF uncertainty, depending on the considered study: only Poissonian for Ilbert et al., Moustakas et al. and Davidzon et al.; Poissonian and stellar mass for Santini et al.; and Poissonian, stellar mass and cosmic variance for Tomczak et al. and the present study. The dashed line shows the SDSS-Galex local measurement of Moustakas et al. (2013). *Lower sub-panels*: the corresponding SMF error contributions normalised by the total SMF uncertainty (see Eq. (5)). The blue dash-dotted line represents the Poissonian contribution. The red dashed line and the cyan solid line represent the cosmic variance and the mass uncertainty contributions, respectively.

over five fields). In general, we observe that the SMF measurements from PRIMUS form the upper limit of the literature. Their SMF estimate is significantly above the others at $0.5 < z < 0.8$. In the range $10^{10.5} < M_* < 10^{11}\,M_\odot$, the difference reaches 0.2 dex, while the authors predict that the cosmic variance should not affect the measurement by more than 10%; a larger offset is observed at $M_* > 10^{11.5}\,M_\odot$, which could be mainly explained by the cosmic variance affecting their measurement, which is estimated to be very strong at high mass[19]. In the next redshift bin ($0.8 < z < 1.1$), the SMF of Moustakas et al. (2013) is also significantly higher than ours. The reason for the discrepancy may be linked to the different recipe (dust models, template libraries, etc.) adopted by Moustakas et al. (2013) in their SED fitting procedure (see also Davidzon et al. 2013, for a discussion about the effect of different SED fitting methods on the SMF). We compared their stellar masses in the XMM-LSS field, which overlaps the W1 field. We found that the PRIMUS masses are higher than ours by 0.17 ± 0.09 dex at $0.2 < z < 0.5$, 0.15 ± 0.08 dex at $0.5 < z < 0.8$ and 0.12 ± 0.1 dex at $0.8 < z < 1$[20]. This could explain part of the observed shift in the SMFs. It is worth noting that the two largest spectroscopic surveys so far, VIPERS and PRIMUS, lead to the largest difference in the SMF measurements. This highlights the great effect of systematic uncertainties in the latest large surveys (see also Coupon et al. 2015).

The comparison of our measurements with deep photometric surveys shows that our results agree well with those of Tomczak et al. (2014) and Ilbert et al. (2013), down to the lowest stellar masses we can explore. Their analysis was based on much deeper data, which confirms the estimate of our lower mass limits. Only in the redshift bin $0.8 < z < 1.1$, we note a

[19] Moustakas et al. (2013) estimated $\sigma_{\rm cv} = 0.1{-}1.4$ for $\log M_* > 11.5\,M_\odot$ at $0.5 < z < 0.8$.

[20] Even by using the same photometry as Moustakas et al. (2013), i.e. including GALEX, CFHTLS, and SWIRE (3.6 and 4.5 μm), similar differences in the stellar masses are observed.

significant difference with Ilbert et al. (2013) in the high-mass end of the SMF. This can be explained by the well-known overdensity in the COSMOS field (Kovač et al. 2010; Bielby et al. 2012)[21].

Finally, we show the local, $z \sim 0.1$, GALEX-SDSS SMF from Moustakas et al. (2013) in all panels of Fig. 12 (as a dash-dotted line). A small but clear and progressive deviation of the SMF with redshift is obvious, in comparison to the local SMF. Far from evident in the previous studies, the trend observed at high mass is confirmed and quantified in Sects. 6.1.3 and 6.2, and is discussed in Sect. 7.1.

6.1.2. Fitting the global, star-forming, and quiescent SMF

To quantify the evolution of the SMF, we adopted the parametrisation proposed by Schechter (1976). As already noted, the total stellar mass function is better fitted with a double Schechter function (Baldry et al. 2008; Pozzetti et al. 2010; Ilbert et al. 2013; Tomczak et al. 2014), defined as

$$\Phi(M_*)\,\mathrm{d}M_* = \mathrm{e}^{-\frac{M_*}{\mathcal{M}^\star}} \left[\Phi_1^\star \left(\frac{M_*}{\mathcal{M}^\star}\right)^{\alpha_1} + \Phi_2^\star \left(\frac{M_*}{\mathcal{M}^\star}\right)^{\alpha_2}\right] \frac{\mathrm{d}M_*}{\mathcal{M}^\star}, \quad (6)$$

where $\mathcal{M}^\star$ is the characteristic stellar mass, $\Phi_1^\star$ and $\Phi_2^\star$ are the normalisation factors, and α_1 and α_2 are the power-law slopes satisfying $\alpha_2 < \alpha_1$.

It has been shown that the massive end of the stellar mass function can be significantly affected by the stellar mass uncertainty (Caputi et al. 2011) through the so-called Eddington bias (Eddington 1913). We corrected the SMF for the Eddington bias during the fitting process by convolving the SMF parametric form by the stellar mass uncertainty σ_M[22] following the procedure described in Ilbert et al. (2013). The authors estimate $\mathrm{d}\sigma_\mathrm{M}$ for each redshift bin, but Grazian et al. (2015) have pointed out the importance of using an estimate of σ_M that depends on the stellar mass in addition to the redshift dependence[23]. We used the $\sigma_\mathrm{M}(M_*, z)$ estimate described in Sect. 5.1 (cf. Fig. 9, sub-panels).

Figure 13 shows the SMF of the global (black stars), star-forming (blue stars), and quiescent (red stars) populations. We included the SMFs measured by Tomczak et al. (2014) and Ilbert et al. (2013), who probed the very low-mass populations for SF and Q galaxies. A simple Schechter function (i.e. $\Phi_2^\star = 0$ in Eq. (6)) seems to be sufficient to fit the star-forming contribution above the stellar mass completeness limit (blue dashed lines). However, as already shown by several studies working with deeper surveys (see e.g. Drory et al. 2009; Ilbert et al. 2013; Tomczak et al. 2014), the SMF of star-forming galaxies reveals an upturn at low mass and is better fitted with a double-Schechter function (Tomczak et al. 2014, Sect. 3.2). Given our stellar mass completeness limit, we can only constrain the low-mass end of the star-forming SMF at $z < 0.5$. Therefore we set the low-mass components of the double-Schechter function to the values found at $0.2 < z < 0.5$, $\alpha_{2\ \mathrm{SF}} = -1.49$ and $\Phi^\star_{2\ \mathrm{SF}} = -3.24$. Our choice is supported the lack of evolution that is observed for α_2 and $\Phi_2^\star$ by Ilbert et al. (2013) and Tomczak et al. (2014). In addition, our values agree quite well with Tomczak et al. (2014), who probed the SMF at lower stellar mass. The resulting double-Schechter function is plotted in Fig. 13 (blue solid line).

For the quiescent galaxies, we clearly need a double-Schechter function to fit the SMF at low redshift (red stars in Fig. 13 upper left panel). The upturn at low mass is slightly more pronounced in our measurement than in the literature, regardless of the position of the quiescent galaxy selection in the NUVrK green valley (cf. Sect. 5.1). In other words, the low-mass slope that we measure does not depend significantly on our selection of quiescent galaxies. We also verified that we find the same shape when we select the quiescent galaxies based on their specific star formation rate (sSFR), using $sSFR < 10^{-11}$ Gyr^{-1} (see Ilbert et al. 2010, for more details on this threshold). We find the upturn position around $M_* \simeq 10^{9.5}\ M_\odot$, in good agreement with previous measurements, that is, $M_* \simeq 10^{9.2}\ M_\odot$, $M_* \simeq 10^{9.4}\ M_\odot$ and $M_* \simeq 10^{9.6}\ M_\odot$ for Tomczak et al. (2014), Ilbert et al. (2013) and Drory et al. (2009), respectively. Even though several deep surveys show that the low-mass upturn of the quiescent SMF is still present at $z > 0.5$, using a single-Schechter function is sufficient given our survey stellar mass limit. The discrepancies between our star-forming and quiescent SMF and the literature are mainly explained by the different criteria used to separate quiescent and star-forming galaxies. If we include the galaxies lying in the green valley in the quiescent sample (i.e. by considering the upper or lower envelopes of the quiescent or star-forming SMF), our measurements of the SMF agree with those of Tomczak et al. (2014) and Ilbert et al. (2013) at $z < 1.1$. At higher redshift, including the green valley in the quiescent locus of the NUVrK diagram is not enough to reconcile the estimates.

We cannot exclude the possibility that we may have missed some fainter red galaxies as a result of the gri-detection described in Sect. 2.4. However, this effect should be limited since we corrected for this incompleteness according to the weight colour map shown in Fig. 2, as previously explained. To add an independent validation of our procedure of correcting for this incompleteness, we used the CFHTLS-Deep/WIRcam Deep Survey (WIRDS; Bielby et al. 2012), which overlaps our CFHTLS-Wide/$K_\mathrm{s} < 22$ survey. We estimated the completeness of the gri selection as a function of redshift and stellar masses separately for quiescent and star-forming galaxies. Below $z < 1.1$, we did not find any completeness problems, regardless of galaxy type or stellar mass range. At $1.1 < z < 1.5$, the quiescent sample is >85% complete after applying our weighting scheme, and applying a correction based on the WIRDS sample would shift the density by less than 0.1 dex, which is well inside our uncertainties. It appears that star-forming galaxies can also be affected by incompleteness around $M_* \sim 10^{10.9}\ M_\odot$ (probably because of dust extinction in massive galaxies at high redshift)[24]. However, comparison with the literature suggests that our SF sample does not significantly suffer from this incompleteness (Fig. 13).

Moreover, we have to highlight that our total SMF agrees with Tomczak et al. (2014) at $M_* > M_\mathrm{lim}$, while our SMF estimate for SF galaxies is continuously higher (by 0.02 dex at $M_* = 10^{10.5}\ M_\odot$ and 0.08 dex at $M_* = 10^{10.75}\ M_\odot$). This SMF difference for SF galaxies would allow a transfer (between the SF and Q populations) that is sufficient to reconcile the our SMF estimate for quiescent galaxies with the estimate of these

[21] The SMFs at $0.8 < z < 1.1$ are consistent with each other if σ_cv computed by Ilbert et al. (2013) is included in the error budget of their SMF ($\sigma_\mathrm{cv} = 0.1$–0.25 for $\log M_*/M_\odot = 11$–12).

[22] Only statistical uncertainties (Poisson and cosmic variance) are considered during the fitting process, while the mass uncertainty is already taken into account in the convolution with the SMF.

[23] By considering the mass dependency of σ_M, we find that the deconvolution has a weaker effect than if we use the mean estimate of σ_M at a given redshift.

[24] A similar trend for extremely dusty star-forming galaxies is visible in Fig. 8 of Ilbert et al. (2010).

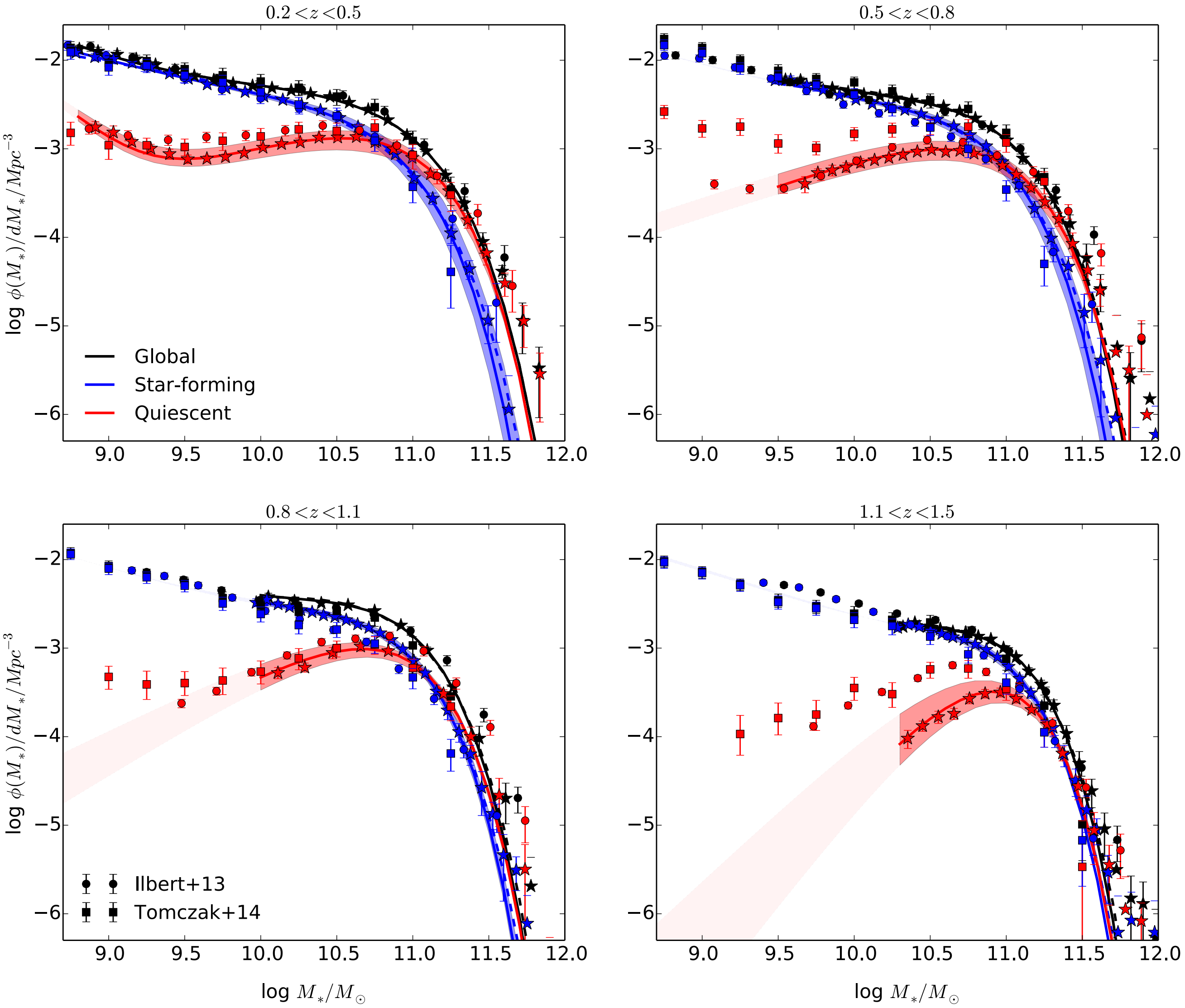


Fig. 13. Stellar mass function for all (black), star-forming (blue), and quiescent (red) galaxies in four redshift bins. The solid lines show the best parametric form of our SMF measurements (stars), while the shaded areas represent the systematic uncertainty due to the SF/Q separation (cf. Sect. 5.1). The dashed lines show the parametric forms obtained if a single-Schechter function is assumed to fit the SF population. The measurements of Tomczak et al. (2014, squares) and Ilbert et al. (2013, circles) are plotted for comparison.

authors. This stresses the sensitivity of the SMF to the Q/SF selection[25].

Since the low-mass end of the global SMF is strongly dominated by the star-forming population at $z > 0.5$, we assumed the same parametrisation of α_2^* and $\Phi_2^\star$ (i.e. $\alpha_2^* = \alpha_{2\ \mathrm{SF}}^*$ and $\Phi_2^\star = \Phi_{2\ \mathrm{SF}}^\star$). We derived two parametric forms of the global SMF, depending on whether the double or the simple Schechter form of the star-forming SMF is considered, as shown in Fig. 13 (with dashed and solid black lines, respectively). The corresponding best-fit parameters are reported in Tables 2 and 3, respectively[26].

[25] We recall that Ilbert et al. (2013) and Tomczak et al. (2014) used a constant selection of quiescent galaxies at $z < 1.5$, while we used a time-dependent selection (cf. Sect. 5.1, Eq. (4)).

[26] All the parameters are given after correction for the Eddington bias (cf. Sect. 6.1.2).

6.1.3. Quantifying the SMF evolution

In Fig. 14 we plot the evolution of the SMF for all (left panel), star-forming (middle panel), and quiescent galaxies (right panel). Each redshift bin is coded with a different colour. As in Fig. 13, the shaded areas show the systematic uncertainty induced by the star-forming or quiescent classification in the NUVrK diagram, while the solid lines represent the parametric form of reference. The arrows show the position of the corresponding characteristic mass $\mathcal{M}^\star$.

As mentioned above, the galaxy population at low masses is strongly dominated by its star-forming component, and the global SMF evolution is then mainly driven by the star-forming population. We note that the evolution of the global SMF that is characterised by a ~0.2 dex increase of the $\mathcal{M}^\star$ (see the arrows in Fig. 14 left panel). However, there is almost no evolution of the star-forming population (Fig. 14 middle panel): the characteristic mass is nearly constant, with $\log(\mathcal{M}_{\mathrm{SF}}^\star/M_\odot) = 10.66_{-0.03}^{+0.02}$ in

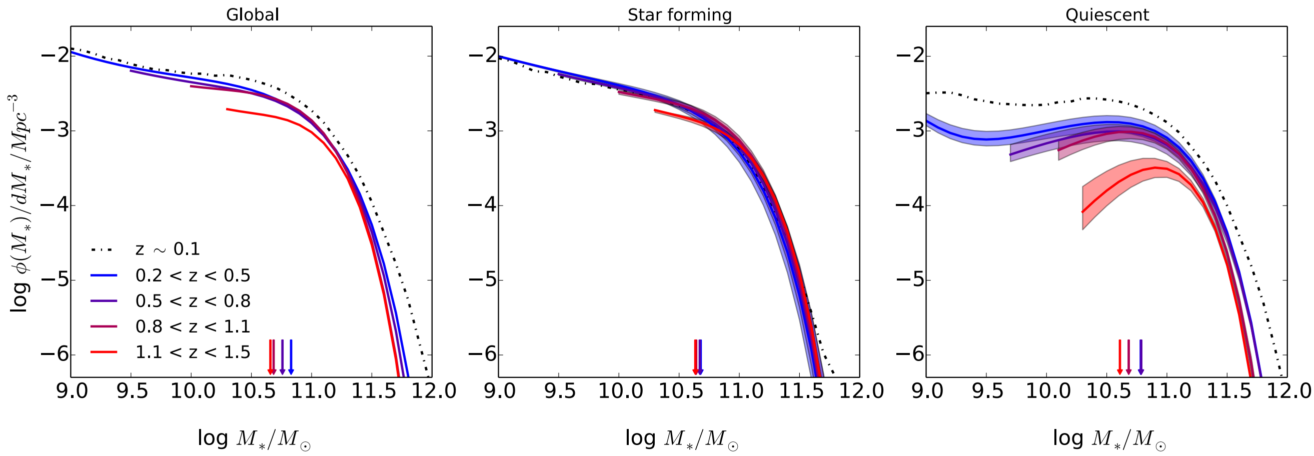


Fig. 14. Evolution of the SMF for the global (*left*), star-forming (*middle*), and quiescent (*right*) populations. The solid lines represent the best SMF parametric form at each redshift, while the arrows show the corresponding $\mathcal{M}^\star$ parameter positions. *Left and middle panels*: the dashed lines show the best fit with a single-Schechter function. *Middle and right panels*: the shaded areas represent the systematic uncertainties that are due to the separation into star-forming or quiescent galaxies depending on whether we insert the galaxies in transition (cf. Sect. 5.1).

Table 2. Best-fit parameters of the SMF parametric form for the total, quiescent, and star-forming populations.

Redshift	$N_{\rm gal}$	$\log(M_{\rm lim})^a$	$\log(\mathcal{M}^\star)^a$	$\log(\Phi_1^\star)^b$	α_1	$\log(\Phi_2^\star)^b$	α_2	$\log(\rho_*)^c$
Quiescent								
$0.2 < z < 0.5$	29078	8.75	$10.78^{+0.02}_{-0.02}$	$-2.86^{+0.02}_{-0.02}$	$-0.44^{+0.05}_{-0.04}$	$-5.88^{+0.21}_{-0.42}$	$-2.43^{+0.20}_{-0.21}$	$7.88^{+0.03}_{-0.03}$
$0.5 < z < 0.8$	38708	9.50	$10.79^{+0.01}_{-0.01}$	$-2.97^{+0.01}_{-0.01}$	$-0.38^{+0.03}_{-0.03}$			$7.76^{+0.02}_{-0.02}$
$0.8 < z < 1.1$	43421	9.97	$10.68^{+0.02}_{-0.02}$	$-2.94^{+0.02}_{-0.02}$	$-0.03^{+0.10}_{-0.10}$			$7.73^{+0.03}_{-0.03}$
$1.1 < z < 1.5$	15567	10.28	$10.61^{+0.02}_{-0.02}$	$-3.60^{+0.03}_{-0.04}$	$1.04^{+0.15}_{-0.14}$			$7.31^{+0.03}_{-0.03}$
Star-forming								
$0.2 < z < 0.5$	143500	8.75	$10.68^{+0.04}_{-0.04}$	$-2.89^{+0.09}_{-0.11}$	$-0.82^{+0.30}_{-0.23}$	$-3.24^{+0.22}_{-0.48}$	$-1.49^{+0.09}_{-0.18}$	$7.98^{+0.03}_{-0.03}$
$0.5 < z < 0.8$	155173	9.50	$10.67^{+0.01}_{-0.01}$	$-2.85^{+0.02}_{-0.03}$	$-0.64^{+0.03}_{-0.03}$	-3.24	-1.49	$8.00^{+0.01}_{-0.01}$
$0.8 < z < 1.1$	114331	9.97	$10.64^{+0.01}_{-0.01}$	$-2.78^{+0.02}_{-0.02}$	$-0.36^{+0.05}_{-0.05}$	-3.24	-1.49	$8.01^{+0.01}_{-0.01}$
$1.1 < z < 1.5$	73600	10.28	$10.63^{+0.02}_{-0.02}$	$-2.97^{+0.02}_{-0.02}$	$0.02^{+0.06}_{-0.06}$	-3.24	-1.49	$7.92^{+0.01}_{-0.01}$
Total								
$0.2 < z < 0.5$	166 658	8.75	$10.83^{+0.02}_{-0.03}$	$-2.63^{+0.03}_{-0.03}$	$-0.95^{+0.10}_{-0.08}$	$-4.01^{+0.28}_{-1.14}$	$-1.82^{+0.18}_{-0.22}$	$8.23^{+0.02}_{-0.02}$
$0.5 < z < 0.8$	185 245	9.50	$10.76^{+0.01}_{-0.01}$	$-2.66^{+0.02}_{-0.02}$	$-0.57^{+0.03}_{-0.03}$	-3.24	-1.49	$8.20^{+0.01}_{-0.01}$
$0.8 < z < 1.1$	153 881	9.97	$10.68^{+0.02}_{-0.02}$	$-2.57^{+0.03}_{-0.03}$	$-0.33^{+0.08}_{-0.08}$	-3.24	-1.49	$8.19^{+0.02}_{-0.03}$
$1.1 < z < 1.5$	85 722	10.28	$10.66^{+0.02}_{-0.02}$	$-2.88^{+0.01}_{-0.01}$	$0.19^{+0.07}_{-0.07}$	-3.24	-1.49	$8.01^{+0.02}_{-0.02}$

Notes. [(a)] $M_\odot$. [(b)] dM_*^{-1} Mpc^{-3}. [(c)] $M_\odot$ Mpc^{-3}.

the redshift range $0.2 < z < 1.5$, while the evolution of the low-mass slope remains very stable, as discussed previously. This confirms that the probability of finding a star-forming galaxy declines exponentially above a certain stellar mass $M_* > \mathcal{M}^\star_{\rm SF}$, which is constant with time. This stresses that the star formation seems to be impeded beyond this stellar mass independent of the redshift up to $z = 1.5$. This is one of the cornerstones of the empirical description proposed by Peng et al. (2010), in which the evolution of high-mass galaxy is dominated by internal quenching mechanisms (named mass quenching by the authors). Peng et al. (2010) suggested that the efficiency of mass quenching is proportional to SFR/$\mathcal{M}^\star$ to keep the SMF of star-forming galaxies constant with redshift.

The right panel of Fig. 14 shows that the main contribution to the evolution of the total SMF is due to the quiescent population build-up. In addition to galaxies that are quenched by mass quenching (around $\mathcal{M}^\star_{\rm SF}$), the SMF evolution of quiescent galaxies reveals an increase of low-mass galaxies with time, as shown in Ilbert et al. (2010). In particular, the SMF upturn built at $z < 0.5$ suggests that the star formation of $M_* < 10^{9-9.5}$ $M_\odot$ galaxies is efficiently quenched, at least at low redshift. Ascribed by Peng et al. (2010) to environmental quenching, the build-up

Table 3. Best-fit parameters of the SMF parametric form for the total and star-forming populations if a single-Schechter function is assumed to fit the SMF of star-forming galaxies.

Star-forming								
Redshift	N_{gal}	$\log(M_{lim})^a$	$\log(\mathcal{M}^\star)^a$	$\log(\Phi_1^\star)^b$	α_1	$\log(\Phi_2^\star)^b$	α_2	$\log(\rho_*)^c$
$0.2 < z < 0.5$	143 500	8.75	$10.79^{+0.01}_{-0.01}$	$-2.89^{+0.02}_{-0.02}$	$-1.29^{+0.01}_{-0.01}$			$7.98^{+0.02}_{-0.02}$
$0.5 < z < 0.8$	155 173	9.50	$10.78^{+0.01}_{-0.01}$	$-2.83^{+0.02}_{-0.02}$	$-1.18^{+0.02}_{-0.02}$			$7.99^{+0.01}_{-0.01}$
$0.8 < z < 1.1$	114 331	9.97	$10.72^{+0.01}_{-0.01}$	$-2.70^{+0.02}_{-0.02}$	$-0.88^{+0.04}_{-0.04}$			$7.99^{+0.02}_{-0.02}$
$1.1 < z < 1.5$	73 600	10.28	$10.73^{+0.02}_{-0.02}$	$-2.83^{+0.02}_{-0.02}$	$-0.71^{+0.07}_{-0.04}$			$7.85^{+0.02}_{-0.02}$
Total								
Redshift	N_{gal}	$\log(M_{lim})^a$	$\log(\mathcal{M}^\star)^a$	$\log(\Phi_1^\star)^b$	α_1	$\log(\Phi_2^\star)^b$	α_2	$\log(\rho_*)^c$
$0.2 < z < 0.5$	166 658	8.75	$10.83^{+0.02}_{-0.03}$	$-2.63^{+0.03}_{-0.03}$	$-0.95^{+0.10}_{-0.08}$	$-4.01^{+0.28}_{-1.14}$	$-1.82^{+0.18}_{-0.22}$	$8.23^{+0.02}_{-0.02}$
$0.5 < z < 0.8$	185 245	9.50	$10.79^{+0.02}_{-0.02}$	$-2.99^{+0.05}_{-0.06}$	$-0.40^{+0.07}_{-0.07}$	-2.83	-1.18	$8.19^{+0.01}_{-0.01}$
$0.8 < z < 1.1$	153 881	9.97	$10.73^{+0.03}_{-0.04}$	$-2.99^{+0.09}_{-0.11}$	$-0.33^{+0.08}_{-0.08}$	-2.70	-0.88	$8.17^{+0.02}_{-0.02}$
$1.1 < z < 1.5$	85 722	10.28	$10.68^{+0.10}_{-0.05}$	$-3.40^{+0.08}_{-0.32}$	$0.64^{+0.27}_{-0.73}$	-2.83	-0.71	$7.96^{+0.02}_{-0.02}$

Notes. [(a)] $M_\odot$. [(b)] $\mathrm{d}M_*^{-1}\ \mathrm{Mpc}^{-3}$. [(c)] $M_\odot\ \mathrm{Mpc}^{-3}$.

of the low-mass quiescent population is discussed in Sect. 7.2. The increase of the very high-mass population that we observe in the quiescent sample (and consequently also in the total SMF) is discussed in Sect. 7.1.

6.2. Evolution of the number densities and stellar mass densities

We derived the galaxy number and stellar mass densities, n_* and ρ_*, respectively, by integrating the stellar mass function

$$n_* = \int_{M_1}^{M_2} \Phi(M_*)\,\mathrm{d}M_* \tag{7}$$

and

$$\rho_* = \int_{M_1}^{M_2} \Phi(M_*)\,M_*\,\mathrm{d}M_*. \tag{8}$$

We adopted the parametric form of the SMF corrected for the Eddington bias. We derived the number densities above the stellar mass completeness limit. The stellar mass density was calculated by integrating the SMF over the stellar mass range $9 < \log(M_*/M_\odot) < 13$, as in Tomczak et al. (2014). We recall that at $z > 0.5$, the stellar mass density relies partially on the extrapolation of the SMF to the lower stellar mass limit.

In Fig. 15 we plot the cosmic evolution of the number densities, n_*, in the stellar mass bins $10.5 < \log(M_*/M_\odot) < 11$ (left), $11 < \log(M_*/M_\odot) < 11.5$ (middle), and $11.5 < \log(M_*/M_\odot) < 12$ (right), between redshifts $z = 0.2$ and $z = 1.5$. For every mass bin, we show the densities for the global, star-forming, and quiescent galaxy populations that we compare with the measurements from Moustakas et al. (2013, triangles) and Matsuoka & Kawara (2010, pentagons) when available. For the global population in our sample, we distinguish two types of evolution. In the two lowest stellar mass bins ($10^{10.5} < M_*/M_\odot < 10^{11.5}$), we observe a two-phase evolution, with an increase of ~25–50% from $z \sim 1.3$ down to $z \sim 1$, followed by a plateau down to $z \sim 0.2$. For the most massive population ($M_* > 10^{11.5} M_\odot$), we observe a continuous increase by slightly less than a factor two from $z \sim 1.5$ to $z \sim 0.2$. A similar, but weaker, trend is seen in VIPERS because of the narrower redshift range. Our results are directly comparable with Matsuoka & Kawara (2010) for $M_* > 10^{11}\ M_\odot$. These authors also took the Eddington bias in their density estimates into account (the estimates are based on simulations). They also emphasised that their measurements at $z < 0.5$ are strongly biased because of their less reliable photo-zs. Within these limits, our n_* evolution measurements for the entire population agree well with their results. The trend observed with PRIMUS is also similar for the lowest mass bins, $M_* < 10^{11.5}\ M_\odot$, although they have systematically higher densities (~40% and ~25% for $M_* \sim 10^{10.75}\ M_\odot$ and $M_* \sim 10^{11.25}\ M_\odot$, respectively), as expected from the higher normalisation of their SMFs (cf. Fig. 12). In addition, it is important to recall that they did not take Eddington bias into account, which can enhance the differences, especially at $M_* > 10^{11}\ M_\odot$.

For the evolution by galaxy type, we observe a two-phase evolution of $M_* < 10^{11.5}\ M_\odot$ quiescent galaxies, while star-forming galaxies experience a constant evolution, if not a decreasing evolution. At low mass, $M_* < 10^{11}\ M_\odot$ (left panel), the density of quiescent galaxies increases with redshift and equals the star-forming density in the lowest redshift bin, at $z \sim 0.3$. For the intermediate masses, $10^{11} < M_*/M_\odot < 10^{11.5}$ (middle panel), the quiescent population becomes dominant at higher redshift, $z \sim 0.9$. In the highest stellar mass bin ($M_* > 10^{11.5}\ M_\odot$, right panel), the quiescent population always outnumbers the star-forming one by representing already 50–60% of the global population at $z \sim 1.3$ and more than 80% at $z \sim 0.3$ (i.e. n_* multiplied by 2.5). From $z \sim 1$ to $z \sim 0.2$, the number density of the massive star-forming galaxies has diminished by a factor of 1.5 and 2 in the two highest mass bins, respectively.

The number densities computed in VIPERS are not plotted since the stellar mass bins used by Davidzon et al. (2013) are different from ours. However, the authors observed the same general trends, though their uncertainties prevent them from distinguishing the two-phase evolutions observed in our survey (Davidzon et al. 2013, Fig. 6). We also generally agree with Moustakas et al. (2013) for star-forming galaxies, as our studies observe a decreasing n_* between $z = 1$ and $z = 0.3$ for

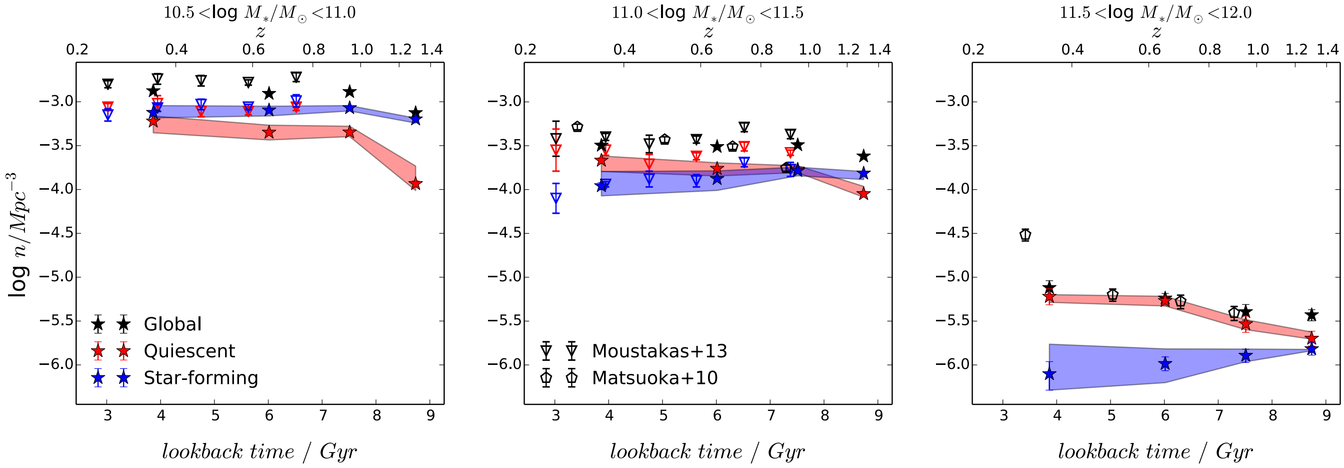


Fig. 15. Evolution of the number densities in three bins of M_*, for the global (black), SF (blue), and Q (red) populations. The corresponding shaded area shows the systematic uncertainty that is due to the SF/Q selection around our reference measurement (stars). The measurements of Moustakas et al. (2013, triangles) and Matsuoka & Kawara (2010, pentagons) are plotted for comparison.

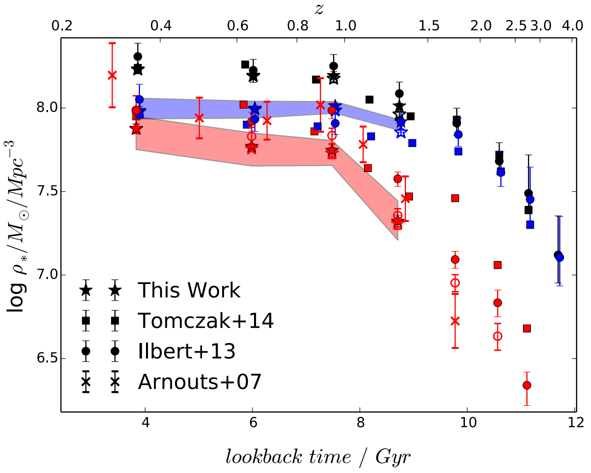


Fig. 16. Evolution of the cosmic stellar mass density for all (black), star-forming (blue), and quiescent (red) galaxies. The shaded areas show the corresponding systematic uncertainties that are due to the SF/Q selection. The open stars represent the measurement that we obtain by assuming a single-Schechter function to fit the star-forming galaxies. Measurements of Tomczak et al. (2014, squares), Ilbert et al. (2013, circles), and Arnouts et al. (2007, quiescent only, red crosses) are shown for comparison. The filled and open red circles represent the quiescent measurements of Ilbert et al. (2013), using a selection of quiescent galaxies based on the NUV-r/r-J plan and the sSFR respectively. The quiescent measurement of Arnouts et al. (2007) is based on the K-band luminosity density, and the selection uses the SED-fitting. For the sake of clarity, the star-forming or quiescent measurements are plotted with of shift of +0.03 or /−0.03 Gyr.

$M_* > 10^{10.5}$ $M_\odot$[27]. The continuous increase of the corresponding quiescent population is also detected by Moustakas et al. (2013) between $z = 1$ and $z = 0.1$ when they measured the weighted linear fits of $n_*(z)$.

Figure 16 presents the cosmic evolution of the stellar mass density ρ_* for all (black), star-forming (blue), and quiescent (red) galaxies. We compare our results (filled stars) with previous studies. We also plot the stellar mass density obtained by assuming a different slope of the star-forming SMF low-mass end (open stars; cf. Sect. 6.1), but it does not change the results significantly. In good agreement with Ilbert et al. (2013, circles)[28] and Tomczak et al. (2014, squares), our measurement of the global evolution of ρ_* reveals two phases: a >50% increase from $z \sim 1.3$ down to $z \sim 1$, and a continuous 12–20% increase from $z \sim 1$ down to $z \sim 0.3$.

As mentioned in Sect. 6.1, our selection of quiescent galaxies is more compatible with the selections of Ilbert et al. (2013) and Tomczak et al. (2014) when we consider that galaxies lying in the green valley are classified as quiescent. This corresponds to the upper red and lower blue envelopes of ρ_* in Fig. 16. Still, our measurement for quiescent galaxies is smaller than previous measurements by up to 25%. We do not find this difference when we consider the global stellar mass density. The importance of the Q/SF selection is reinforced by the fact that the agreement is better with Ilbert et al. (2013), when they use the log $sSFR = -11$ selection[29] (Fig. 16, open red circles). Our measurement is also consistent with the ρ_* measured by Arnouts et al. (2007, red crosses) for quiescent galaxies, which are selected thanks to SED-fitting (we do not plot their star-forming ρ_*[30]).

As previously suggested, the evolution of the stellar mass density of star-forming galaxies seems to be quite stable at $z < 1.5$. At the same time, a rapid increase of the stellar mass contained in quiescent galaxies is observed, increased by a factor > 2.5 from $z \sim 1.3$ down to $z \sim 1$. At lower redshift, we detect a small and continuous $\gtrsim 30\%$ increase of ρ_* from $z \sim 1$ down to $z \sim 0.3$, which reflects the progressive quenching of less massive galaxies.

[27] Our highest stellar mass bin is not explored in Moustakas et al. (2013), who limited their analysis to $M_* < 10^{11.5}$ $M_\odot$).

[28] With respect to our results, the slightly higher values measured in COSMOS are expected, given the $8 < \log(M_*/M_\odot) < 13$ integration range adopted by Ilbert et al. (2013).

[29] See Ilbert et al. (2010) concerning this threshold.

[30] The ρ_* measurement of Arnouts et al. (2007) is based on the K-band luminosity. Ilbert et al. (2010, Appendix D) showed that the mass-to-light ratio derived by Arnouts et al. (2007) for star-forming galaxies is not appropriate at low and intermediate masses.

7. Discussion

7.1. High-mass end evolution

As highlighted above, our sample can be used to investigate the evolution of massive ($M_* > 10^{10.5}$ $M_\odot$) and rare ($M_* > 10^{11.5}$ $M_\odot$) galaxies, thanks to the large volume of our survey. Most importantly, we are interested in the evolution of these objects across cosmic time, in particular to understand which mechanisms determine their evolution from star-forming to quiescent galaxies. Several studies (e.g. Kauffmann et al. 2003; Bundy et al. 2006; Davidzon et al. 2013) have characterised galaxy quenching with the so-called transition mass, which is the stellar mass at which the quiescent and star-forming populations are equal in a given redshift bin. In the same spirit, we define the transition redshift, $z_{\rm tr}$, at which the quiescent population becomes dominant. As shown in Fig. 15, the transition redshift is found to be $z_{\rm tr} \gtrsim 1.4$, $z_{\rm tr} \sim 0.9$ and $z_{\rm tr} \sim 0.2$, for $M_* \sim 10^{11.75}$ $M_\odot$, $M_* \sim 10^{11.25}$ $M_\odot$, and $M_* \sim 10^{10.75}$ $M_\odot$ galaxies, respectively: globally, the more massive a galaxy, the earlier its star formation is stopped. This is qualitatively consistent with the redshift evolution of the transition mass (e.g. see Davidzon et al. 2013). As already mentioned, several physical mechanisms could explain this trend within a hierarchical context (e.g. De Lucia & Blaizot 2007; Neistein & Dekel 2008; Weinmann et al. 2012). For instance, based on the stellar-halo mass relation from Coupon et al. (2015), the star-forming galaxies with stellar masses of $M^*_{\rm SF}(\sim 10^{10.64}$ $M_\odot)$ should reside in dark matter halos of masses of around $M_{\rm h} \sim 10^{12.4}$ $M_\odot$. This value agrees well with the halo mass threshold invoked by Cattaneo et al. (2006), corresponding to halo' quenching, but we cannot exclude that some radio-AGN quenching could also explain why massive galaxies cease forming stars and/or are not fuelled anymore by fresh infalling gas (Croton et al. 2006).

We find that the number density of the most massive ($M_* > 10^{11.5}$ $M_\odot$) galaxies almost doubled from $z \sim 1$ to $z \sim 0.3$ (Fig. 15). This corresponds to the <0.25 dex increase of the SMF high-mass end that is seen between $z \sim 1$ and $z \sim 0.3$ (Fig. 14). Because the high-mass end is dominated by quiescent galaxies at $z < 1$, the increase of the $M_* > 10^{11.5}$ $M_\odot$ population cannot be explained by incidental star formation (Arnouts et al. 2007). If we assume that, in general, these very high-mass galaxies do not experience significant star formation, they can still assemble stellar mass through mergers at $z < 1$, in particular through dry merging.

7.2. Taming of galaxies

In Sect. 6.1 we have shown that the characteristic stellar mass of the star-forming SMF does not vary significantly between redshifts $z = 0.2$ and $z = 1.5$. As described in Sect. 5.1, we performed three selections of the SF galaxies, and the values of $\mathcal{M}^\star_{\rm SF}$ differed slightly from one selection to another. In Fig. 17 we plot $\mathcal{M}^\star_{\rm SF}$ as a function of the redshift and the SF galaxy selection in the NUVrK diagram. First, we find that $\mathcal{M}^\star_{\rm SF}$ is between $10^{10.6}$ and $10^{10.8}$ $M_\odot$ at $0.2 < z < 1.5$, regardless of the SF selection in the NUVrK diagram. More precisely, we find

- $\log \mathcal{M}^\star_{\rm SF}/M_\odot = 10.69^{+0.04}_{-0.05}$ if the galaxies in transition are included in the selection of SF galaxies (upper dotted lines in Fig. 8);
- $\log \mathcal{M}^\star_{\rm SF}/M_\odot = 10.66^{+0.02}_{-0.03}$ for our intermediate selection; and
- $\log \mathcal{M}^\star_{\rm SF}/M_\odot = 10.64^{+0.01}_{-0.01}$ for the most conservative selection.

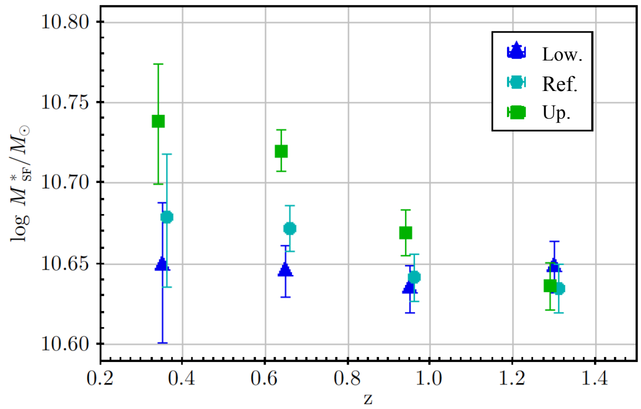


Fig. 17. Redshift evolution of $\mathcal{M}^\star_{\rm SF}$, corresponding to the three selections of SF galaxies in the NUVrK diagram defined in Sect. 5.1: the reference selection (for which the limit lies in the middle of the green valley; cyan circles), its lower limit (when galaxies in transition are excluded; blue triangles), and the upper limit (if the green valley is included in the SF locus; green squares).

Therefore, the evolution of $\mathcal{M}^\star_{\rm SF}$ is consistent with being constant if the galaxies transitioning in the green valley are excluded from the selection of SF galaxies. The invariance with respect to redshift of $\mathcal{M}^\star_{\rm SF}$ for the most conservative selection strongly supports a mass-quenching process occurring around a constant stellar mass, which makes this selection suitable for investigating the galaxies that are about to quench.

7.2.1. Tracking galaxies in the green valley

To identify a potential quenching channel for $\mathcal{M}^\star_{\rm SF}$ galaxies, we isolate and characterise the green valley galaxies in Fig. 18, where each panel shows a different redshift bin. The contours represent the density of quiescent and star-forming galaxies, when the galaxies in transition are excluded (i.e. using the strictest selection of Q/SF galaxies). The colour code expresses the stellar mass. In the lower panels, we show the rest-frame $(r - K_{\rm s})^\circ$ distribution of the transitioning galaxies (i.e. the galaxies lying in the NUVrK green valley). As explained in Sect. 5.1, the NUVrK diagram is very efficient in separating dusty star-forming galaxies from quiescent ones (see Fig. 16 of the companion paper), which allows us to properly define transitioning galaxies in the green valley. We observe that

1. the $(r - K_{\rm s})^\circ$ distribution of galaxies in transition is narrow and does not evolve with redshift (>80% of these galaxies have $0.76 < (r - K_{\rm s})^\circ < 1.23$); and that
2. the typical stellar mass of galaxies in transition is around $\mathcal{M}^\star_{\rm SF}$ (>60% of these galaxies have $10^{10.5} < M_*/M_\odot < 10^{11}$).

Therefore, we isolated the quenching channel of the $\mathcal{M}^\star_{\rm SF}$-galaxies with the colour criterion $(r - K_{\rm s})^\circ > 0.76$ in the NUVrK green valley (green dashed lines in Fig. 18, sub-panels).

We also detect a clear plume of young quiescent galaxies in Fig. 18, with $(r - K_{\rm s})^\circ < 0.76$ (i.e. bluer than observed for galaxies following the $\mathcal{M}^\star_{\rm SF}$ channel) at $z < 0.5$. It is well established that rest-frame optical-NIR colours are sensitive to both dust attenuation and age of the stellar populations (see e.g. Whitaker et al. 2012). Under the assumption that, on average, the $(r - K_{\rm s})^\circ$ colour of quiescent galaxies cannot become bluer with time, the young part of the quiescent population should have used another quenching channel. According to the limit that we defined to isolate the $\mathcal{M}^\star_{\rm SF}$ quenching channel (green dashed line

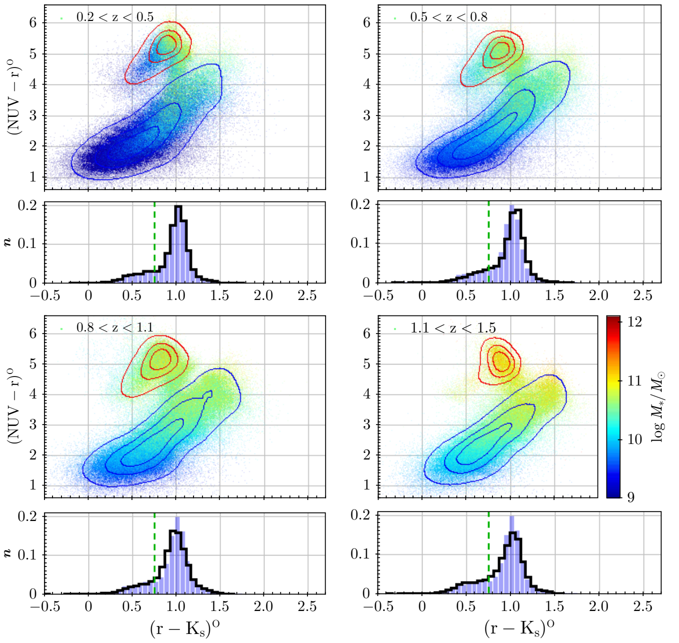


Fig. 18. NUVrK galaxy distribution ouside and inside the green valley, shown in four redshift bins. *Top sub-panels*: NUVrK diagram as a function of the galaxy stellar mass. The red and blue contours show the equal density of the quiescent and star-forming populations, respectively, after excluding the transitioning galaxies (i.e. the galaxies lying in the green valley defined in Fig. 8). *Bottom sub-panels*: normalised number counts along the $(r - K_s)^\circ$ colour in the green valley (black solid line). The distribution at $0.2 < z < 0.5$ is repeated in each panel for comparison (blue shaded area). The vertical green dashed line shows the limit of the $\mathcal{M}^\star_{\rm SF}$-quenching channel, as discussed in Sect. 7.2.1.

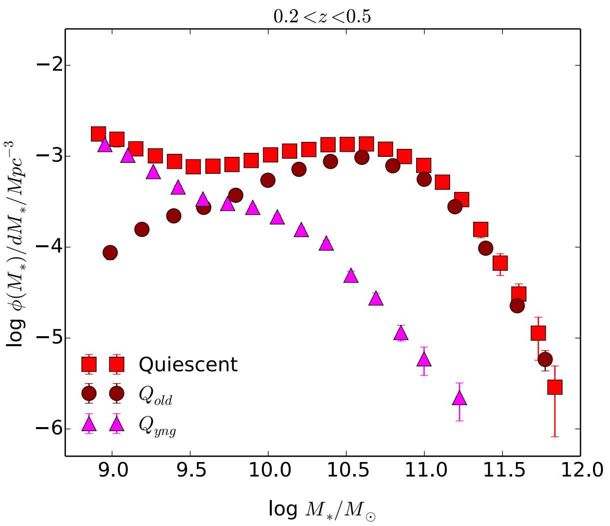


Fig. 19. Deconstruction of the quiescent SMF at $0.2 < z < 0.5$. The red squares represent the measurement for the whole quiescent population, while the magenta triangles and the darkred circles show the SMF for the young ($Q_{\rm yng}$) [$(r - K_s)^\circ < 0.76$] and old ($Q_{\rm old}$) [$(r - K_s)^\circ > 0.76$] quiescent populations, respectively.

in Fig. 18), we separated the young quiescent ($Q_{\rm yng}$) and old quiescent ($Q_{\rm old}$) galaxies with $(r - K_s)^\circ = 0.76$. Figure 18 also reveals that $Q_{\rm yng}$ galaxies are characterised by relative low masses ($M_* \lesssim 10^{9.5}\ M_\odot$), which seems to match the low-mass upturn of the quiescent SMF (see Fig. 13) at $z < 0.5$. In Fig. 19 we compute the SMF for $Q_{\rm yng}$ (magenta triangles) and $Q_{\rm old}$ (dark red circles) galaxies at $0.2 < z < 0.5$. The $Q_{\rm yng}$ galaxies dominate at low mass, and they are responsible for the low-mass upturn in the quiescent SMF. At the same time, the SMF of $Q_{\rm old}$ galaxies peaks at $\mathcal{M}^\star_{\rm SF}$, which clearly supports the idea that the building of the quiescent SMF is led through two quenching channels that can be distinguished with a cut in the NUVrK diagram at $(r - K_s)^\circ = 0.76$. The timescale might then be a key element for characterising the mechanisms that are involved in each channel.

7.2.2. Quenching timescales

In Sect. 7.2.1 we have identified two possible channels in which galaxies are transitioning to build the quiescent population. We now investigate the nature of these channels through their characteristic timescales.

The restframe UV is sensitive to timescales of 10^{-2}–10^{-1} Gyr, and the scarcity of young/low-mass galaxies in the green valley allows us to expect that some quenching processes occur on timescales of the same order or shorter. To better constrain the timescale of the quenching that affects the star formation of low-mass and $\mathcal{M}^\star_{\rm SF}$ galaxies, we explored the behaviour of simple scenarios of star formation history (SFHs) within the NUVrK diagram in a similar way as the approach adopted by Schawinski et al. (2014). We performed this analysis at $0.2 < z < 0.5$, where both old and young quiescent galaxies are well identified. The use of simple e-folding SFHs implies that we assumed that galaxies can only become redder with time. This is motivated by the fact that the fraction of quiescent galaxies has continuously increased between $z \sim 3$ and $z \sim 0.2$ (e.g. Ilbert et al. 2010; Muzzin et al. 2013; Mortlock et al. 2015) and by assuming that most green valley galaxies are transitioning for the first time (Martin et al. 2007). Doing so, we neglect the green valley galaxies produced by rejuvenation processes, as observed in the local Universe (e.g Salim & Rich 2010; Thomas et al. 2010) and recently predicted at higher redshift in the EAGLE simulations (Trayford et al. 2016). However, in these simulations, the rejuvenation is responsible for a small fraction of the green valley galaxies.

Figure 20 presents the resulting tracks in the NUVrK diagram for SFHs constructed in the same way: a continuous star formation up to the quenching time at t_Q, followed by an exponentially declining star formation characterised by τ. To mimic the average properties of our $K_s < 22$ sample at $0.2 < z < 0.5$, the example is plotted for one metallicity ($Z = 0.008$), one extinction law (Calzetti et al. 2000), one value of the dust attenuation ($E(B - V) = 0.2$), and with a stellar age of at least 1 Gyr. The stellar age is colour coded, and only the ages allowed by the given redshift bin are plotted. In the left panel of Fig. 20 the SFHs are characterised by $t_Q = 1$ Gyr, with $\tau = 0.1$, 0.25, 1, 2, and 2.5 Gyr. The tracks are constructed in a very simple way, and the evolution assumes a constant dust attenuation based on its average value for the bluest SF galaxies. The arrows show the shift that is due to a 0.1 increase of $E(B - V)$. It is expected that quiescent galaxies are less affected by dust, which would tend to make the tracks steeper in the NUVrK green valley. Keeping this effect in mind, we see as a first result that the presence of $Q_{\rm yng}$ galaxies is expected if any quenching process occurs early ($t_Q \sim 1$ Gyr) with a typical timescale of $\tau \lesssim 0.25$ Gyr (triangles and squares in the left panel of Fig. 20). As a second result, $\tau = 1$ Gyr (inverted triangles in Fig. 20 left panel) seems to be a lower limit for the quenching timescale that is compatible with the channel drawn by $\mathcal{M}^\star_{\rm SF}$ galaxies. The galaxies with a quenching $\tau > 2$ Gyr do not reach the quiescent cloud.

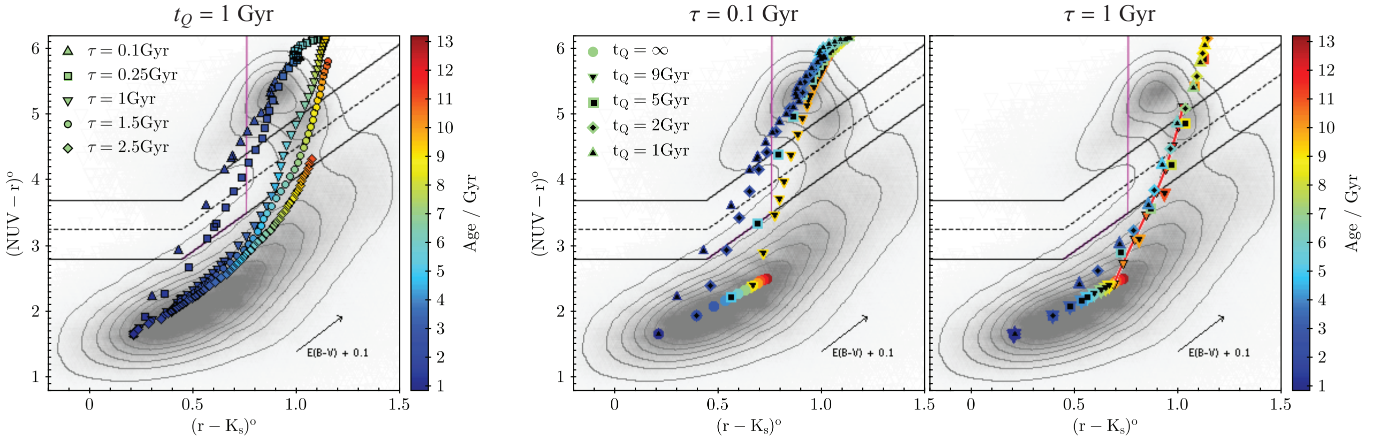


Fig. 20. Predicted BC03 tracks in the NUVrK diagram at $0.2 < z < 0.5$ for $Z = 0.008$ (Calzetti et al. 2000) and $E(B-V) = 0.2$. The arrow shows the shift expected for $E(B-V) + 0.1$. Analogously to Fig. 8, the black solid and dashed lines correspond to the limits of the green valley and its middle, respectively, while we report the $(r-K_s)°$-limit of the $\mathcal{M}^\star$-quenching channel with a vertical magenta solid line. The grey contours outline the galaxy density distribution. Each marker is coloured with respect to the corresponding stellar age (in Gyr). *Left panel*: only one quenching time is considered: t_Q = 1 Gyr, with τ = 0.1 Gyr (triangles), τ = 0.25 Gyr (squares), τ = 1 Gyr (inverted triangles), τ = 1.5 Gyr (circles), and τ = 2.5 Gyr (diamonds). *Right panels*: two quenching timescales are considered: τ = 0.1 Gyr (*middle panel*) and τ = 1 Gyr (*right panel*), for t_Q = 1 Gyr (triangles), 2 Gyr (diamonds), 5 Gyr (squares), and 9 Gyr (inverted triangles). The filled circles show the track for a continuous star formation without quenching. The red solid line linking the black edge triangles shows the track for t_Q = 9 Gyr and τ = 0.5 Gyr.

In the middle and right panels of Fig. 20, we also investigate the effect of the quenching epoch. We fixed τ = 1 Gyr and τ = 0.1 Gyr for several values of t_Q between 1 and 9 Gyr. Any $t_Q > 9$ Gyr will produce the same result as t_Q = 9 Gyr since the NUVrK colours of SF galaxies saturate at ages >9 Gyr, as shown by the predicted track with a continuous star formation (circles). All the models with τ = 1 Gyr are able to explain the galaxy presence within the $\mathcal{M}^\star$ channel. We could also imagine that a shorter timescale combined with a late quenching time can reproduce the observed $\mathcal{M}^\star$ channel. However, if we consider an SFH with τ = 0.1 Gyr after 9 Gyr on the SF main sequence (middle panel, inverted triangles), the track seems to move away from the channel that is drawn by $\mathcal{M}^\star$-galaxies in the NUVrK diagram. To produce a track that is compatible with this channel, we need to consider a quenching timescale $\tau \gtrsim 0.5$ Gyr (red solid line), regardless of the considered SHF. We recall that we have considered the shortest timescales compatible with the $\mathcal{M}^\star_{\rm SF}$-quenching channel, and we could pick out SFHs that agree better. Namely, SFHs characterised by t_Q = 1 Gyr and τ = 1.5 Gyr, t_Q = 5 Gyr and τ = 1 Gyr, or t_Q = 9 Gyr and $0.5 < \tau < 1$ Gyr could also explain the presence of this channel. This suggests a quenching timescale range of $0.5 < \tau < 2$ Gyr for $\mathcal{M}^\star$-galaxies, which corresponds to a quenching duration of between ~1 and 3.5 Gyr[31]. Therefore, the physical mechanism explaining the building of the quiescent SMF around $\mathcal{M}^\star_{\rm SF}$ at $z < 1$ seems to be a slow process. Such a mass dependent mechanism is compatible with a strangulation picture where the star formation quenching occurs on several Gyr, moving slowly away from the SF main sequence in the NUVrK diagram, while the gas supply is progressively halted (Schawinski et al. 2014; Peng et al. 2015).

Figure 20 shows that the plume formed by $Q_{\rm yng}$ galaxies in the NUVrK plan is explained by a ~0.1 Gyr-quenching process occurring during the first ~5 Gyr of the galaxy life (squares, diamonds, and triangles in the middle panel of Fig. 20). The absence of these low-mass galaxies lying in the green valley can be first explained by the rapidity of their quenching. Indeed, a galaxy quenching with τ = 0.1 Gyr (triangles in the left and middle panels of Fig. 20) is expected to cross the green valley (delimited by the black solid lines) in ~0.4 Gyr, while a galaxy with $\tau \sim$ 0.5–2 Gyr spends ~1–3.5 Gyr there, on average. Nevertheless, the potential reservoir of SF $M_* < 9.5$ galaxies that can quench is about ten times larger than for galaxies around $\mathcal{M}^\star_{\rm SF}$ (cf. Fig. 13). We could then expect to see more low-mass galaxies in transition. By adopting a conservative approach, we can assume that the ratio between the two quenching timescales is ~10 (0.1 Gyr for $M_* < 9.5$ galaxies, 1 Gyr around $\mathcal{M}^\star_{\rm SF}$). The corresponding quenching rate should consequently be about ten times lower for the low-mass galaxies that are the progenitors of the $Q_{\rm yng}$ galaxies than for the $\mathcal{M}^\star_{\rm SF}$ galaxies. The resulting flux of quenching galaxies (i.e. quenching rate × SF reservoir) that cross the green valley is then expected to be of the same order of magnitude, both at low and high mass, except when only a fraction of the low-mass galaxies is likely to be affected by the quenching. The SF satellite galaxies, which are estimated to be ≳3 times less abundant than field galaxies (Yang et al. 2009; Peng et al. 2012), are therefore good candidates for this low-mass quenching mechanism. Moreover, its typical timescale is compatible with the scenario suggested by Schawinski et al. (2014) for the rapid formation of young early-type galaxies. In this picture, the quiescent low-mass galaxies are formed through dramatic events such as major mergers and not through ram-pressure stripping or strangulation, by explaining both the almost instantaneous star formation shutdown and the morphological transformation.

8. Summary

We analysed the evolution of the stellar mass function in an unprecedentedly large (>22 deg^2) NIR selected ($K_s < 22$) survey. This allowed us to provide reliable constraints on the evolution of massive galaxies and to investigate quenching processes below redshift $z \sim 1.2$. Covering the VIPERS spectroscopic survey, we computed highly reliable photometric redshifts, with usual

[31] These values agree with the estimate of Fritz et al. (2014) in VIPERS, who found that massive ($\log(M_*/M_\odot) > 11$) galaxies are expected to turn quiescent in ~1.5 Gyr at $0.7 < z < 1.3$, and more slowly at $z < 0.7$ (i.e. with longer quenching durations).

estimates of the precision $\sigma_{\Delta z/(1+z)} < 0.03$ and $\sigma_{\Delta z/(1+z)} < 0.05$ for bright ($i < 22.5$) and faint ($i > 22.5$) galaxies, respectively.

Paying particular attention to several sources of uncertainties (photometry, star-galaxy separation, photometric redshift, dust extinction treatment, and classification into quiescent and star-forming galaxies), we computed the SMF between redshifts $z = 0.2$ and $z = 1.5$. The unique size of our sample enabled us to drastically reduce the statistical uncertainties affecting the SMFs and stellar mass densities with respect to other current surveys over the stellar mass range we consider: the Poissonian error and cosmic variance are reduced by factors of ∼3.3 and ∼2, respectively, compared to a 2 deg^2-survey. Combined with a careful treatment of the Eddington bias that is due to the stellar mass uncertainty, we produced an unprecedentedly precise measurement of the massive end of the SMF at $z < 1.5$. In particular, we stress the importance of constraining all sources of systematic uncertainties, which quickly become the dominant sources of error in large-scale surveys such as those planned with Euclid or LSST.

Using the (NUV$-r$) versus ($r-K$) rest-frame colour diagram to classify star-forming and quiescent galaxies in our sample, we measured the evolution of the SMFs of the two populations and investigated the possible quenching processes that could explain the build-up of the quiescent population. Our main conclusions are summarised below.

1) We provided clear evidence that the number density of the most massive ($M_* > 10^{11.5}\ M_\odot$) galaxies increases by a factor ∼2 from $z \sim 1$ to $z \sim 0.3$, which was first highlighted by Matsuoka & Kawara (2010). This population is largely dominated by the quiescent population since $z \sim 1$, allowing for the possibility of galaxy mass assembly through dry-mergers in very massive galaxies.
2) The characteristic mass of the SF population was found to be very stable in the redshift range $0.2 < z < 1.5$, with $\log(\mathcal{M}^\star_{\rm SF}/M_\odot) = 10.64 \pm 0.01$. This confirms that the star formation is impeded above a certain stellar mass (Ilbert et al. 2010; Peng et al. 2010).
3) Using the NUVrK diagram as a tracer of the galaxy evolution, we identified one main quenching channel between the star-forming and quiescent sequences at $0.2 < z < 1.5$, which is followed by galaxies with stellar masses around $\mathcal{M}^\star_{\rm SF}$. This channel is characterised by a colour $(r - K_{\rm s})^\circ > 0.76$, typical of evolved massive star-forming galaxies, which should feed the majority of the quiescent population. We also identified a young quiescent population with $(r - K_{\rm s})^\circ < 0.76$, whose galaxies likely followed another path to reach the quiescent sequence. We showed that this blue quiescent population, dominated by low-mass galaxies, is responsible for the upturn of the quiescent SMF at low redshift.
4) Assuming simple e-folding SFHs (galaxies can only become redder with time), we found that the $\mathcal{M}^\star_{\rm SF}$ channel is explained by long quenching timescales, with $0.5 < \tau \lesssim 2$ Gyr. Galaxies in this channel are expected to turn quiescent after ∼1−3.5 Gyr on average. This is compatible with strangulation processes occurring when the gas cooling or the cold gas inflows are impeded, allowing the galaxy to progressively consume its remaining gas reservoir (Peng et al. 2015). Conversely, the quenching of low-mass galaxies that is visible at low redshift is characterised by short timescales with $\tau \sim 0.1$ Gyr. This quenching that halts star formation in ∼0.4 Gyr can be consistent with major merging (Schawinski et al. 2014) and may preferentially affect satellite galaxies.

Acknowledgements. We gratefully acknowledge the anonymous referee, whose advice helped much in improving the clarity of the paper. We wish to thank J.-C. Cuillandre, H. Aussel, S. de la Torre and B. C. Lemaux for helpful discussions. We would also like to thank J. Moustakas for providing the PRIMUS stellar mass estimates. This research is in part supported by the Centre National d'Études Spatiales (CNES) and the Centre National de la Recherche Scientifique (CNRS) of France, and the ANR Spin(e) project (ANR-13-BS05-0005, http://cosmicorigin.org). L.G. acknowledges support of the European Research Council through the Darklight ERC Advanced Research Grant (# 291521). This paper is based on observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/DAPNIA, and with WIRCam, a joint project of CFHT, Taiwan, Korea, Canada and France, at the Canada-France-Hawaii Telescope (CFHT). The CFHT is operated by the National Research Council (NRC) of Canada, the Institut National des Sciences de l'Univers of the Centre National de la Recherche Scientifique (CNRS) of France, and the University of Hawaii. This work is based in part on data products produced at Terapix available at the Canadian Astronomy Data Centre as part of the Canada-France-Hawaii Telescope Legacy Survey, a collaborative project of NRC and CNRS. We thank the Terapix team for the reduction of all the WIRCAM images and the preparation of the catalogues matching with the T0007 CFHTLS data release. This paper is based on observations made with the Galaxy Evolution Explorer (GALEX). GALEX is a NASA Small Explorer, whose mission was developed in cooperation with the Centre National d'Études Spatiales (CNES) of France and the Korean Ministry of Science and Technology. GALEX is operated for NASA by the California Institute of Technology under NASA contract NAS5-98034. This paper uses data from the VIMOS Public Extragalactic Redshift Survey (VIPERS). VIPERS has been performed using the ESO Very Large Telescope, under the "Large Programme" 182.A-0886. The participating institutions and funding agencies are listed at http://vipers.inaf.it. This paper uses data from the VIMOS VLT Deep Survey (VVDS) obtained at the ESO Very Large Telescope under programs 070.A-9007 and 177.A-0837, and made available at the CESAM data center, Laboratoire d'Astrophysique de Marseille, France. Funding for PRIMUS is provided by NSF (AST-0607701, AST-0908246, AST-0908442, AST-0908354) and NASA (Spitzer-1356708, 08-ADP08-0019, NNX09AC95G). Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the US Department of Energy Office of Science. The Participating Institutions of the SDSS-III Collaboration are listed at http://www.sdss3.org/.

References

Arnouts, S., Moscardini, L., Vanzella, E., et al. 2002, MNRAS, 329, 355
Arnouts, S., Walcher, C. J., Le Fèvre, O., et al. 2007, A&A, 476, 137
Arnouts, S., Le Floc'h, E., Chevallard, J., et al. 2013, A&A, 558, A67
Baldry, I. K., Glazebrook, K., Brinkmann, J., et al. 2004, ApJ, 600, 681
Baldry, I. K., Balogh, M. L., Bower, R. G., et al. 2006, MNRAS, 373, 469
Baldry, I. K., Glazebrook, K., & Driver, S. P. 2008, MNRAS, 388, 945
Baldry, I. K., Driver, S. P., Loveday, J., et al. 2012, MNRAS, 421, 621
Bardeau, S., Kneib, J.-P., Czoske, O., et al. 2005, A&A, 434, 433
Barnes, J. E. 1992, ApJ, 393, 484
Bell, E. F., Wolf, C., Meisenheimer, K., et al. 2004, ApJ, 608, 752
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Bertin, E., Mellier, Y., Radovich, M., et al. 2002, in Astronomical Data Analysis Software and Systems XI, eds. D. A. Bohder, D. Durand, & T. H. Handley, ASP Conf. Ser., 281, 228
Bielby, R. M., Hill, M. D., Metcalfe, N., & Shanks, T. 2012, MNRAS, 419, 1315
Borch, A., Meisenheimer, K., Bell, E. F., et al. 2006, A&A, 453, 869
Bradshaw, E. J., Almaini, O., Hartley, W. G., et al. 2013, MNRAS, 433, 194
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Bundy, K., Ellis, R. S., Conselice, C. J., et al. 2006, ApJ, 651, 120
Calzetti, D., Armus, L., Bohlin, R., et al. 2000, ApJ, 533, 682
Caputi, K. I., Cirasuolo, M., Dunlop, J. S., et al. 2011, MNRAS, 413, 162
Cattaneo, A., Dekel, A., Devriendt, J., Guiderdoni, B., & Blaizot, J. 2006, MNRAS, 370, 1651
Chabrier, G. 2003, PASP, 115, 763
Cirasuolo, M., McLure, R. J., Dunlop, J. S., et al. 2007, MNRAS, 380, 585
Coil, A. L., Blanton, M. R., Burles, S. M., et al. 2011, ApJ, 741, 8
Conseil, S., Vibert, D., Amouts, S., et al. 2011, in Astronomical Data Analysis Software and Systems XX, eds. I. N. Evans, A. Accomazzi, D. J. Mink, & A. H. Rots, ASP Conf. Ser., 442, 107
Coupon, J., Arnouts, S., van Waerbeke, L., et al. 2015, MNRAS, 449, 1352
Cowie, L. L., Songaila, A., Hu, E. M., & Cohen, J. G. 1996, AJ, 112, 839
Croton, D. J., Springel, V., White, S. D. M., et al. 2006, MNRAS, 365, 11
Daddi, E., Cimatti, A., Renzini, A., et al. 2004, ApJ, 617, 746
Davidzon, I., Bolzonella, M., Coupon, J., et al. 2013, A&A, 558, A23
Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, AJ, 145, 10

De Lucia, G., & Blaizot, J. 2007, MNRAS, 375, 2
De Lucia, G., Springel, V., White, S. D. M., Croton, D., & Kauffmann, G. 2006, MNRAS, 366, 499
Dickinson, M., Papovich, C., Ferguson, H. C., & Budavári, T. 2003, ApJ, 587, 25
Drory, N., Bundy, K., Leauthaud, A., et al. 2009, ApJ, 707, 1595
Eddington, A. S. 1913, MNRAS, 73, 359
Efstathiou, G., Ellis, R. S., & Peterson, B. A. 1988, MNRAS, 232, 431
Erben, T., Hildebrandt, H., Miller, L., et al. 2013, MNRAS, 433, 2545
Faber, S. M., Willmer, C. N. A., Wolf, C., et al. 2007, ApJ, 665, 265
Fontana, A., Donnarumma, I., Vanzella, E., et al. 2003, ApJ, 594, L9
Fontana, A., Salimbeni, S., Grazian, A., et al. 2006, A&A, 459, 745
Fontanot, F., De Lucia, G., Monaco, P., Somerville, R. S., & Santini, P. 2009, MNRAS, 397, 1776
Fritz, A., Scodeggio, M., Ilbert, O., et al. 2014, 563, A92
Garilli, B., Le Fèvre, O., Guzzo, L., et al. 2008, A&A, 486, 683
Garilli, B., Guzzo, L., Scodeggio, M., et al. 2014, A&A, 562, A23
Gavazzi, G., & Scodeggio, M. 1996, A&A, 312, L29
Giavalisco, M., Ferguson, H. C., Koekemoer, A. M., et al. 2004, ApJ, 600, L93
Grazian, A., Fontana, A., Santini, P., et al. 2015, A&A, 575, A96
Gunn, J. E., & Gott, III, J. R. 1972, ApJ, 176, 1
Guzzo, L., Scodeggio, M., Garilli, B., et al. 2014, A&A, 566, A108
Haines, C. P., Gargiulo, A., La Barbera, F., et al. 2007, MNRAS, 381, 7
Hildebrandt, H., Erben, T., Kuijken, K., et al. 2012, MNRAS, 421, 2355
Hogg, D. W., Blanton, M. R., Eisenstein, D. J., et al. 2003, ApJ, 585, L5
Hudelot, P., Cuillandre, J.-C., Withington, K., et al. 2012, VizieR Online Data Catalog: II 317
Ilbert, O., Tresse, L., Zucca, E., et al. 2005, A&A, 439, 863
Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, A&A, 457, 841
Ilbert, O., Capak, P., Salvato, M., et al. 2009, ApJ, 690, 1236
Ilbert, O., Salvato, M., Le Floc'h, E., et al. 2010, ApJ, 709, 644
Ilbert, O., McCracken, H. J., Le Fèvre, O., et al. 2013, A&A, 556, A55
Ilbert, O., Arnouts, S., Le Floc'h, E., et al. 2015, A&A, 579, A2
Jarvis, M. J., Bonfield, D. G., Bruce, V. A., et al. 2013, MNRAS, 428, 1281
Johnston, R. 2011, A&ARv, 19, 41
Kauffmann, G., Heckman, T. M., White, S. D. M., et al. 2003, MNRAS, 341, 33
Kauffmann, G., White, S. D. M., Heckman, T. M., et al. 2004, MNRAS, 353, 713
Kereš, D., Katz, N., Weinberg, D. H., & Davé, R. 2005, MNRAS, 363, 2
Kovač, K., Lilly, S., Cucciati, O., et al. 2010, ApJ, 708, 505
Kron, R. G. 1980, ApJS, 43, 305
Larson, R. B., Tinsley, B. M., & Caldwell, C. N. 1980, ApJ, 237, 692
Lawrence, A., Warren, S. J., Almaini, O., et al. 2007, MNRAS, 379, 1599
Le Fèvre, O., Vettolani, G., Garilli, B., et al. 2005, A&A, 439, 845
Le Fèvre, O., Cassata, P., Cucciati, O., et al. 2013, A&A, 559, A14
Leauthaud, A., Massey, R., Kneib, J.-P., et al. 2007, ApJS, 172, 219
Maraston, C. 2005, MNRAS, 362, 799
Marchesini, D., van Dokkum, P. G., Förster Schreiber, N. M., et al. 2009, ApJ, 701, 1765
Martin, C., & GALEX Team. 2005, in Maps of the Cosmos, eds. M. Colless, L. Staveley-Smith, & R. A. Stathakis, IAU Symp., 216, 221
Martin, D. C., Wyder, T. K., Schiminovich, D., et al. 2007, ApJSS, 173, 342
Matsuoka, Y., & Kawara, K. 2010, MNRAS, 405, 100
McLure, R. J., Pearce, H. J., Dunlop, J. S., et al. 2013, MNRAS, 428, 1088
Mitchell, P. D., Lacey, C. G., Baugh, C. M., & Cole, S. 2013, MNRAS, 435, 87
Mortlock, A., Conselice, C. J., Hartley, W. G., et al. 2015, MNRAS, 447, 2
Moster, B., Somerville, R., Newman, J., & Rix, H.-W. 2011, ApJ, 731, 113
Moustakas, J., Coil, A. L., Aird, J., et al. 2013, ApJ, 767, 50
Moutard, T., Arnouts, S., Ilbert, O., et al. 2016, A&A, 590, A102
Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013, ApJ, 777, 18
Neistein, E., & Dekel, A. 2008, MNRAS, 388, 1792
Newman, J. A., Cooper, M. C., Davis, M., et al. 2013, ApJS, 208, 5
Oke, J. 1974, ApJS, 27, 21
Peng, Y.-J., Lilly, S., Kovač, K., et al. 2010, ApJ, 721, 193
Peng, Y.-J., Lilly, S. J., Renzini, A., & Carollo, M. 2012, ApJ, 757, 4
Peng, Y., Maiolino, R., & Cochrane, R. 2015, Nature, 521, 192
Pickles, A. J. 1998, PASP, 110, 863
Polletta, M., Tajer, M., Maraschi, L., et al. 2007, ApJ, 663, 81
Pozzetti, L., Bolzonella, M., Lamareille, F., et al. 2007, A&A, 474, 443
Pozzetti, L., Bolzonella, M., Zucca, E., et al. 2010, A&A, 523, A13
Prevot, M., Lequeux, J., Prevot, L., Maurice, E., & Rocca-Volmerange, B. 1984, A&A, 132, 389
Puget, P., Stadler, E., Doyon, R., et al. 2004, in Ground-based Instrumentation for Astronomy, eds. A. F. M. Moorwood, & M. Iye, SPIE Conf. Ser., 5492, 978
Regnault, N., Conley, A., Guy, J., et al. 2009, A&A, 506, 999
Salim, S., & Rich, R. M. 2010, ApJ, 714, L290
Salpeter, E. 1955, ApJ, 121, 161
Santini, P., Fontana, A., Grazian, A., et al. 2012, A&A, 538, A33
Schawinski, K., Urry, C. M., Simmons, B. D., et al. 2014, MNRAS, 440, 889
Schechter, P. 1976, ApJ, 203, 297
Schmidt, M. 1968, ApJ, 151, 393
Scoville, N., Aussel, H., Brusa, M., et al. 2007, ApJS, 172, 1
Silk, J., & Mamon, G. A. 2012, RA&A, 12, 917
Song, M., Finkelstein, S. L., Ashby, M. L. N., et al. 2015, ApJ, submitted [arXiv:1507.05636]
Thomas, D., Maraston, C., Schawinski, K., Sarzi, M., & Silk, J. 2010, MNRAS, 404, 1775
Tomczak, A. R., Quadri, R. F., Tran, K.-V. H., et al. 2014, ApJ, 783, 85
Trayford, J. W., Theuns, T., Bower, R. G., et al. 2016, MNRAS, submitted [arXiv:1601.07907]
Weinmann, S. M., Pasquali, A., Oppenheimer, B. D., et al. 2012, MNRAS, 426, 2797
Whitaker, K. E., Kriek, M., van Dokkum, P. G., et al. 2012, ApJ, 745, 179
Williams, R. J., Quadri, R. F., Franx, M., van Dokkum, P., & Labbé, I. 2009, ApJ, 691, 1879
Yang, X., Mo, H. J., & van den Bosch, F. C. 2009, ApJ, 695, 900
York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, AJ, 120, 1579
Zucca, E., Zamorani, G., Vettolani, G., et al. 1997, A&A, 326, 477

4.3.4 Article 4 : Measuring and modelling the redshift evolution of clustering : the Hubble Deep Field North

Mon. Not. R. Astron. Soc. **310**, 540–556 (1999)

Measuring and modelling the redshift evolution of clustering: the *Hubble Deep Field North*

S. Arnouts,[1,2]★ S. Cristiani,[1,3] L. Moscardini,[1] S. Matarrese,[4] F. Lucchin,[1] A. Fontana[5] and E. Giallongo[5]

[1]*Dipartimento di Astronomia, Università di Padova, vicolo dell'Osservatorio 5, I-35122 Padova, Italy*
[2]*Institut d'Astrophysique de Paris, CNRS, 98bis Bd. Arago, F-75014 Paris, France*
[3]*Space Telescope-European Coordinating Facility, E.S.O., Karl-Schwarzschild Str. 2, D-85748 Garching, Germany*
[4]*Dipartimento di Fisica G. Galilei, Università di Padova, via Marzolo 8, I-35131 Padova, Italy*
[5]*Osservatorio Astronomico di Roma, via dell'Osservatorio, I-00040 Monteporzio, Italy*

Accepted 1999 July 15. Received 1999 July 9; in original form 1999 February 23

ABSTRACT

The evolution of galaxy clustering from $z = 0$ to $z \simeq 4.5$ is analysed using the angular correlation function and the photometric redshift distribution of galaxies brighter than $I_{AB} \leqslant 28.5$ in the *Hubble Deep Field North*. The reliability of the photometric redshift estimates is discussed on the basis of the available spectroscopic redshifts, comparing different codes and investigating the effects of photometric errors. The redshift bins in which the clustering properties are measured are then optimized to take into account the uncertainties of the photometric redshifts. The results show that the comoving correlation length r_0 has a small decrease in the range $0 \lesssim z \lesssim 1$ followed by an increase at higher z. We compare these results with the theoretical predictions of a variety of cosmological models belonging to the general class of Cold Dark Matter scenarios, including Einstein–de Sitter models, an open model and a flat model with non-zero cosmological constant. Comparison with the expected mass clustering evolution indicates that the observed high-redshift galaxies are biased tracers of the dark matter with an effective bias b strongly increasing with redshift. Assuming an Einstein–de Sitter universe, we obtain $b \simeq 2.5$ at $z \simeq 2$ and $b \simeq 5$ at $z \simeq 4$. These results support theoretical scenarios of biased galaxy formation in which the galaxies observed at high redshift are preferentially located in more massive haloes. Moreover, they suggest that the usual parameterization of the clustering evolution as $\xi(r, z) = \xi(r, 0)(1 + z)^{-(3+\epsilon)}$ is not a good description for any value of ϵ. Comparison of the clustering amplitudes that we measured at $z \simeq 3$ with those reported by Adelberger et al. and Giavalisco et al., based on a different selection, suggests that the clustering depends on the abundance of the objects: more abundant objects are less clustered, as expected in the paradigm of hierarchical galaxy formation. The strong clustering and high bias measured at $z \simeq 3$ are consistent with the expected density of massive haloes predicted in the frame of the various cosmologies considered here. At $z \simeq 4$, the strong clustering observed in the *Hubble Deep Field* requires a significant fraction of massive haloes to be already formed by that epoch. This feature could be a discriminant test for the cosmological parameters if confirmed by future observations.

Key words: galaxies: clusters: general – galaxies: photometry – cosmology: observations – cosmology: theory – large-scale structure of Universe.

1 INTRODUCTION

Clustering properties represent a fundamental clue about the formation and evolution of galaxies. Several large spectroscopic surveys have measured the correlation function of galaxies in the local universe, studying its dependence on morphological type or absolute magnitude (Santiago & da Costa 1990; Park et al. 1994; Loveday et al. 1995; Benoist et al. 1996; Tucker et al. 1997). Higher values of the correlation length r_0 are observed for elliptical galaxies (or galaxies with brighter absolute magnitude),

★ E-mail: sarnouts@eso.org

while lower values are obtained for late type galaxies (or galaxies with fainter absolute magnitude). This difference in the clustering strength suggests that the various galaxy populations are not related in a straightforward way to the distribution of the matter. To account for these observations, one has to consider as a first approach that galaxies are biased tracers of the matter distribution as $\xi_{\rm gal}(r) = b^2(\mathcal{M})\xi_{\rm m}(r)$ (Kaiser 1984), where $\xi_{\rm gal}(r)$ refers to the spatial correlation function of the galaxies, $\xi_{\rm m}(r)$ refers to the spatial correlation function of the mass and $b(\mathcal{M})$ represents the bias associated with different galaxy populations. Here $\mathcal{M}$ describes the intrinsic properties of the objects (like mass, luminosity, etc.).

Deep spectroscopic surveys have made it possible to reach higher redshifts and study the evolution of galaxy clustering. For example the Canada–France Redshift Survey (CFRS) (Le Fèvre et al. 1996) samples the universe up to $z \simeq 1$ while the K-selected galaxy catalogue by Carlberg et al. (1997) reaches $z \simeq 1.5$. From these data it has been possible to find a clear signal for evolution in the clustering strength: the correlation length is three times smaller at high redshift ($z \simeq 1$) than its local value. In addition, Carlberg et al. (1997) have found segregation effects between the red and blue samples similar to those observed locally. A common approach is to assume that the galaxy sample traces the underlying mass density fluctuation

$$b(\mathcal{M}, z) = 1, \qquad \text{or at least} \qquad b(\mathcal{M}, z) = \text{constant},$$

and fits the clustering evolution of the mass with a parametric form

$$\xi(r, z) = \xi(r, 0)(1 + z)^{-(3+\epsilon)}$$

(Peebles 1980), where ϵ describes the evolution of the mass distribution due to gravitational instability. Such an assumption makes it straightforward to discriminate between different cosmological models. From N-body simulations, Colín, Carlberg & Couchman (1997) found faster evolution in the Einstein–de Sitter (hereafter EdS) universe than in an open universe with matter density parameter $\Omega_{0\rm m} = 0.2$ ($\epsilon \simeq 0.8$ and $\epsilon = 0.2$, respectively). Carlberg et al. (1997) obtained from their data a small value of ϵ that would be quite difficult to reconcile with an EdS universe, while Le Fèvre et al. (1996) found a value $0 \leqslant \epsilon \leqslant 2$, still consistent with any fashionable cosmological model. However, using the galaxy clustering evolution directly to derive the relevant properties of the mass is a questionable practice, due to the bias acting as a complicating factor. Different samples select a mixture of galaxy masses and the effective bias, which is expected in current hierarchical galaxy formation theories to depend on redshift and mass, i.e. $b(\mathcal{M}, z)$, plays a key role in the observed evolution of clustering. Exciting progress in this field has been achieved with the recent discovery of a large number of galaxies at $z \simeq 3$ (Lyman-Break Galaxies, hereafter LBGs) using the U-dropout technique (Steidel et al. 1996). For the first time, the high-z universe is probed via a population of quite 'normal' galaxies in contrast with the previous surveys dominated by QSOs or radio galaxies. The LBG samples offer the opportunity to estimate in a narrow time-scale ($2.6 \leqslant z \leqslant 3.4$) number densities, luminosities, colours, sizes, morphologies, star formation rates (SFR), chemical abundances, dynamics and clustering of these primordial galaxies. By using different catalogues and statistical techniques, Giavalisco et al. (1998, hereafter G98) and Adelberger et al. (1998, hereafter A98) have measured the correlation length r_0 of this population. The values they found are at least comparable to that of present-day spiral galaxies ($r_0 = 2$–$4\,h^{-1}$ Mpc when an EdS universe is assumed). Such a strong clustering at $z \simeq 3$ is inconsistent with clustering evolution modelled in terms of the ϵ parameter for any value of ϵ (G98). By comparing the correlation amplitudes with the predictions for the mass correlation, G98 and A98 obtained (for an EdS universe) a linear bias $b \simeq 4.5$ and $b \simeq 6$, respectively. These results suggest that the LBGs formed preferentially in massive dark-matter haloes.

An alternative way to extend the present information over a larger range of redshift is to use the photometric measurements of redshifts in deep multicolour surveys. This technique, based on the comparison between theoretical (and/or observed) spectra and the observed colours in various bands, makes it possible to derive a redshift estimate for galaxies that are one or two magnitudes fainter than the deepest limit for spectroscopic surveys (even with 10-m-class telescopes).

An optimal combination of deep observations and the photometric redshift technique has been attained with the *Hubble Deep Field North* (*HDF North*). Photometric redshifts have been used to search for high-redshift galaxies (Lanzetta, Yahil & Fernández-Soto 1996) and investigate the evolution of their luminosity function and SFR (Sawicki, Lin & Yee 1997; Madau et al. 1996; Gwyn & Hartwick 1996; Franceschini et al. 1998), their morphology (Abraham et al. 1996; van den Bergh et al. 1996; Fasano et al. 1998) and clustering properties (Connolly, Szalay & Brunner 1998; Miralles & Pelló 1998; Magliocchetti & Maddox 1999; Roukema et al. 1999). A critical issue is the statistical uncertainty of the photometric redshifts, which strongly depends on the number of bands following at the various redshifts – the main features of a galaxy spectral energy distribution (hereafter SED), in particular the 4000 Å break and the 912 Å Lyman break.

The aim of this paper is to measure the galaxy clustering evolution in the full redshift range $0 \leqslant z \leqslant 4.5$, using the photometric redshifts of a galaxy sample with $I_{\rm AB} \leqslant 28.5$ in the *HDF North* (including infrared data, i.e. Fernández-Soto, Lanzetta & Yahil 1999, hereafter FLY99) and carry out an extended comparison of the results with the theoretical predictions of different current galaxy formation scenarios based on variants of the Cold Dark Matter model. This comparison will be performed using the techniques introduced in Matarrese et al. (1997) and Moscardini et al. (1998), which allow a detailed modelling of the evolution of galaxy clustering, accounting both for the nonlinear dynamics of the dark matter distribution and for the redshift evolution of the galaxy-to-mass bias factor.

Our sample probes a population fainter than the spectroscopic LBGs and an inter-comparison of their clustering properties will be useful to address the differences in the nature of the two populations. However, the photometric redshift approach should be used with some caution when reaching such faint limits. In fact, uncertainties and systematic errors are expected to be larger than those estimated in the comparison of photometric and spectroscopic redshifts, which is typically limited to $I_{\rm AB} \leqslant 26$. This problem is particularly relevant for the analysis of the angular correlation function since in this statistic all galaxies at a given redshift contribute with the same weight. This is different, for example, from what happens when these objects are used to estimate the star formation rate history, where brighter objects, with smaller uncertainties in the redshift determination, have more weight. For these reasons we try to provide a rough estimate of the errors in the redshift estimates at faint magnitudes, by comparing the results of different photometric redshift techniques and by using Monte Carlo simulations. This in turn provides the necessary information to define optimal redshift bin sizes (i.e. minimizing the effects of the redshift uncertainties) for the clustering analysis.

The plan of the paper is as follows. In Section 2, we present the photometric database and we describe the photometric redshift technique. In Section 3, we investigate the reliability of the photometric redshift estimates. In Section 4, we present the results for the angular correlation function computed in different redshift ranges. Section 5 is devoted to a comparison of these results with the theoretical predictions of different cosmological models belonging to the general class of the Cold Dark Matter scenario. Finally, discussion and conclusions are presented in Section 6.

2 THE PHOTOMETRIC REDSHIFT MEASUREMENT

2.1 The photometric database

As a basis for the present work, we have used the photometric catalogue produced by FLY99 on the *HDF North* using the source extraction code SExtractor (Bertin & Arnouts 1996). In addition to the four optical WFPC2 bands (Williams et al. 1996), infrared observations in the J, H and Ks bands (Dickinson et al., in preparation) are incorporated.

A particularly valuable feature of the FLY99 catalogue is that the optical images are used to model spatial profiles that are fitted to the infrared images in order to measure optimal infrared fluxes and uncertainties. In this way, for the large majority of the objects, an estimate of the infrared flux is available down to the fainter magnitudes. This is a definite advantage for the derivation of photometric redshifts.

The analysis described below has been applied to the $F300W$, $F450W$, $F606W$, $F814W$, J, H and Ks magnitudes of 1023 objects down to $I_{AB} \simeq 28.5$ (here we note that the magnitude I_{AB} refers directly to the photometric catalogue given in FLY99 and not to their best fit I_{AB} reported in their photometric redshift catalogue).

2.2 The photometric redshift technique

Various authors have explored a number of different approaches to estimating redshifts of galaxies from deep broad-band photometric databases. Empirical relations between magnitudes and/or colours and redshifts have been calibrated using spectroscopic samples (Connolly et al. 1995; Wang, Bahcall & Turner 1998). Other techniques are based on the comparison of the observed colours of galaxies with those expected from template SEDs, either observed (Lanzetta et al. 1996; FLY99) or theoretical (Giallongo et al. 1998) or a combination of the two (Sawicki et al. 1997, hereafter SLY97). Bayesian estimation has also been used (Benítez 1998).

2.2.1 The synthetic spectral libraries

The type of approach followed in the present work is based on the comparison of observed colours with theoretical SEDs and has been described in Giallongo et al. (1998). Here we summarize its main ingredients.

(i) The SEDs are derived from the GISSEL library (Bruzual & Charlot, in preparation). The spectral synthesis models are governed by a number of free parameters listed in Table 1. The star formation rate for a galaxy with a given age is governed by the assumed e-folding star formation time-scale τ. Several values of τ and galaxy ages are necessary to reproduce the different observed spectral types. We also have to assume a shape for the initial mass function (IMF). As shown in Giallongo et al. (1998), the photometric redshift estimate is not significantly changed by using different IMFs. Here we restricted our analysis to a Salpeter IMF.

(ii) In addition to the GISSEL parameters, we have added the internal reddening for each galaxy by applying the observed attenuation law of local starburst galaxies derived by Calzetti, Kinney & Storchi-Bregmann (1994) and Calzetti (1997). The different values of the reddening excess are listed in Table 1. We have also included the Lyman absorption produced by the intergalactic medium as a function of redshift in the range $0 \leqslant z \leqslant 5$, following Madau (1995).

As a result we obtained a library of 2.5×10^5 spectra, which can be used to derive the colours as a function of redshift for all the model galaxies with an age smaller than the Hubble time at the given redshift (which is cosmology-dependent; the adopted cosmological parameters are also given in Table 1).

2.2.2 Estimating redshifts

To measure the photometric redshifts we used a standard χ^2 fitting procedure comparing the observed fluxes $F_{\rm obs}$ (and corresponding uncertainties) with the GISSEL templates $F_{\rm tem}$:

$$\chi^2 = \sum_i \left[\frac{F_{{\rm obs},i} - sF_{{\rm tem},i}}{\sigma_i}\right]^2, \tag{1}$$

where $F_{{\rm obs},i}$ and σ_i are the fluxes observed in a given filter i and their uncertainties, respectively; $F_{{\rm tem},i}$ are the fluxes of the template in the same filter; the sum runs over the seven filters. The template fluxes have been normalized to the observed ones by choosing the factor s that minimizes the χ^2 value ($\partial\chi^2/\partial s = 0$):

$$s = \sum_i \left[\frac{F_{{\rm obs},i}F_{{\rm tem},i}}{\sigma_i^2}\right] \Big/ \sum_i \left[\frac{F_{{\rm tem},i}^2}{\sigma_i^2}\right]. \tag{2}$$

In the GISSEL library the models provide fluxes emitted per unit mass (in $M_\odot$) and the normalization parameter s, which rescales the template fluxes to the observed ones, provides a rough estimation of the observed galaxy mass. We have limited the range of models accepted in the χ^2 comparison to the interval 10^7–$10^{14}\,M_\odot$. We derived the χ^2 probability function (CPF) as a function of z using the lowest χ^2 values at any redshift. To have an idea of the redshift uncertainties we have derived the interval corresponding to the standard increment $\Delta\chi^2 = 1$. At the same

Table 1. Parameters used for the library of templates.

IMF	Salpeter
Exponential SFR	
Time-scales τ (Gyr)	$1, 2, 3, 5, 9, \infty$, 2 bursts
Ages (Gyr)	.01,.05,.1,.25,.5,.75,1.,1.5,2., 3.,4.,5.,6.,7.,8.,9.,10.,11.,12.,14.
Metallicities	$Z_\odot$, $0.2\,Z_\odot$, $0.02\,Z_\odot$
E_{B-V}	0,0.05,0.1,0.2,0.3,0.4
Extinction law	Calzetti
Cosmology (H_0, q_0)	50, 0.5

time the CPF is analysed to detect the presence, if any, of secondary peaks with a multi-thresholding algorithm (typically we decompose the normalized CPF into ten levels). We notice that our estimates of the photometric redshifts are changed by less than 2 per cent if we adopt a different cosmology:

$$(\Omega_{0m} = 0.3, \Omega_{0\Lambda} = 0) \qquad \text{or} \qquad (\Omega_{0m} = 0.3, \Omega_{0\Lambda} = 0.7)$$

and our mass estimates are nearly unchanged.

3 COMPARISON WITH PREVIOUS WORKS AND SIMULATIONS

3.1 Spectroscopic versus photometric redshifts

In Fig. 1 we show the comparison of our estimates of the photometric redshifts $z_{\rm phot}$ with the 106 spectroscopic redshifts $z_{\rm spec}$ up to $z \simeq 5$ listed in the FLY99 catalogue (see references therein). Our values are generally consistent with the observed spectroscopic redshifts within the estimated uncertainties over the full redshift range. The rms dispersion σ_z for different redshift intervals is reported in Table 2. At redshifts lower than 1.5, two galaxies have photometric redshifts that appear clearly discrepant: galaxy #191 (the number refers to the FLY99 number) with $z_{\rm phot} \simeq 1.05$ versus $z_{\rm spec} \simeq 0.37$ and galaxy #619 with $z_{\rm phot} \simeq 0.95$ versus $z_{\rm spec} \simeq 0.37$. Also FLY99 and SLY97 found for these two objects $z_{\rm phot} \geqslant 0.88$. As discussed in the next section, the techniques used in SLY97, in FLY99 and in the present work are significantly different; therefore, if the spectroscopic redshifts are correct, both objects are expected to have a really peculiar SED. For example, various SEDs used in these works do not include spectra with strong emission lines (starbursts, AGN, …). Yet, based on the observed spectra, the two spectroscopic redshifts are very uncertain (see `http://astro.berkeley.edu/davisgrp/HDF/`). Disregarding these two objects, the photometric accuracy at $z < 1.5$ decreases from $\sigma_z \simeq 0.13$ to $\sigma_z \simeq 0.09$. These values are consistent with the photometric redshift estimates obtained in previous works and compiled by Hogg et al. (1998).

At redshifts $z \geqslant 1.5$ the dispersion is $\sigma_z = 0.24$ if the galaxy #687, which shows catastrophic disagreement (it is found at low redshift also by FLY99, while there is no clear association in the SLY97 catalogue), is discarded. Direct inspection of the original frames shows that in this case the photometry can be incorrect due to the complex morphology of this object, which was assumed to be a single unit.

3.2 Comparison with other photometric redshifts

The relatively good agreement of the photometric redshifts with the spectroscopic ones shows the reliability of our method at bright magnitudes. Obviously, the same accuracy cannot also be expected at fainter magnitudes below the spectroscopic limit $I_{\rm AB} \geqslant 26$. The uncertainty in the identification of the characteristic features (4000 Å and Lyman break) in the observed SEDs necessarily increases when the errors in the photometry become larger. In order to obtain a rough idea of the uncertainty also in the domain inaccessible to spectroscopy we have compared the results of our code with those obtained with other photometric methods.

FLY99 and SLY97 have used the four spectra provided by Coleman, Wu & Weedman (1980), which reproduce different star formation histories or different galaxy types (E/S0, Sbc, Scd and Irr). The wavelength coverage of these template spectra is however too small (1400–10000 Å) to allow direct comparison with the full range of photometric data (3000–25000 Å). To bypass this problem, both authors have extrapolated the infrared SEDs by using the theoretical SEDs of the GISSEL library, corresponding to the four spectral types. In the UV, SLY97 have again used an extrapolation based on GISSEL, while FLY99 have used the observations of Kinney et al. (1993). SLY97 have enlarged the SED library with two spectra of young galaxies with constant star formation (from the GISSEL library) and interpolated

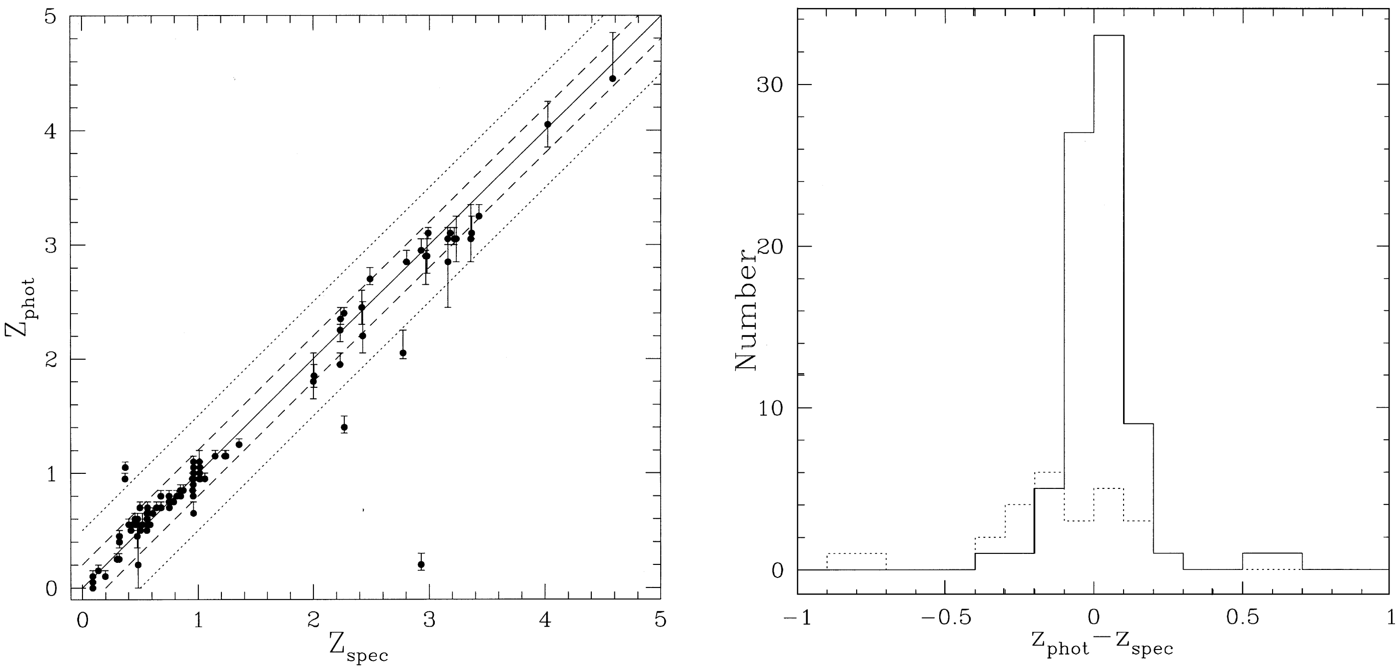


Figure 1. The left panel shows the comparison of our photometric redshifts $z_{\rm phot}$ with the spectroscopic ones $z_{\rm spec}$. Error bars represent the region where $\Delta\chi^2 \leqslant 1$. The dotted and long-dashed lines represent $\Delta z = 0.5$ and $\Delta z = 0.2$, respectively. The right panel shows the histograms with $z_{\rm phot} - z_{\rm spec}$ for galaxies with $z_{\rm spec} \leqslant 1.5$ (solid line) and $z_{\rm spec} > 1.5$ (dotted line).

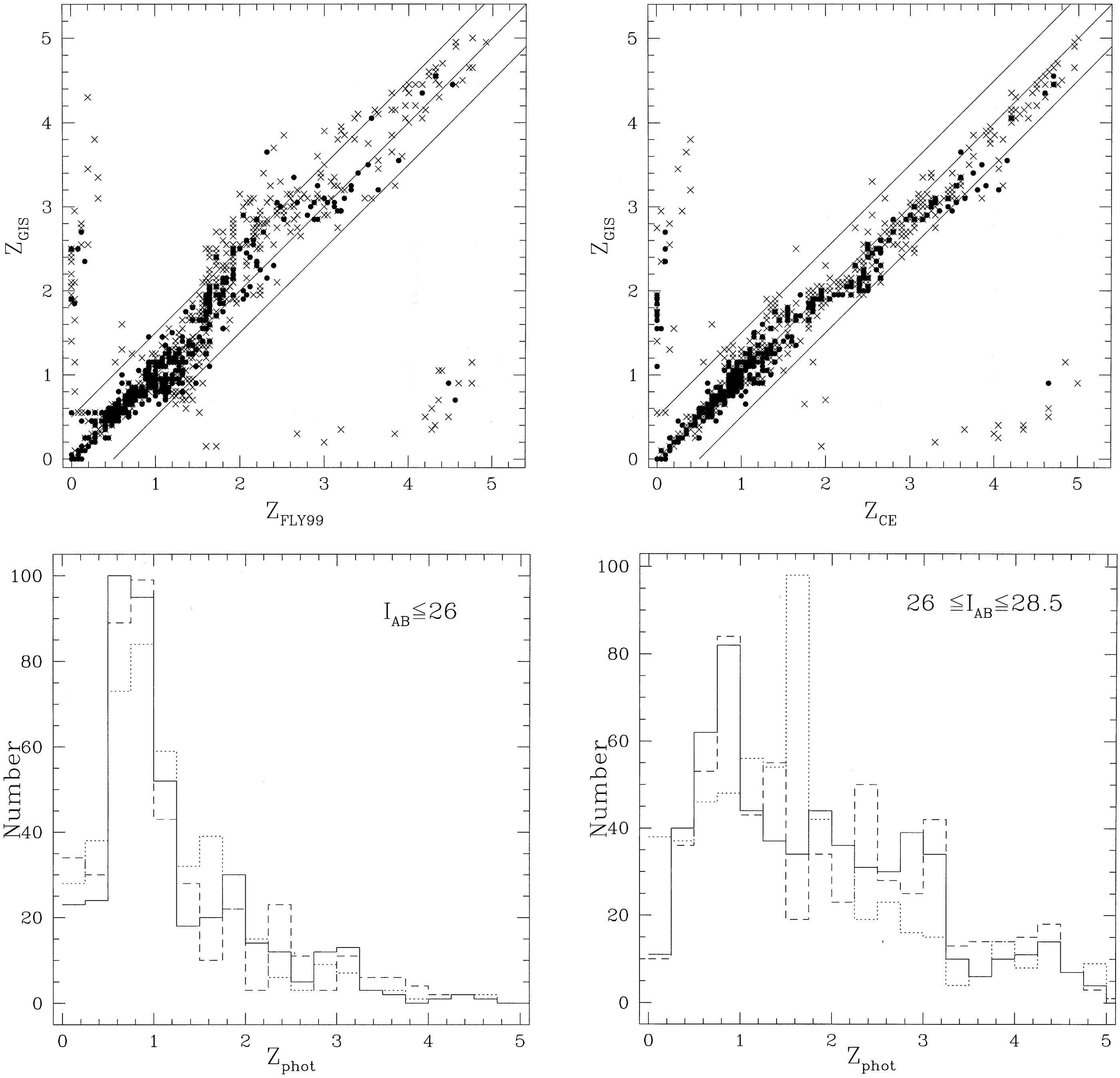


Figure 2. Comparison of our photometric redshifts z_{GIS} (computed using the GISSEL library) with the photometric redshifts obtained by FLY99 (z_{FLY99}; upper left panel) and with our Coleman Extended (CE) libraries similar to the SLY97 method (z_{CE}; upper right panel). Filled circles represent objects with $I_{AB} \leqslant 26$ and crosses refer to objects with $26 \leqslant I_{AB} \leqslant 28.5$. Solid lines correspond to $\Delta z = 0.5$. In the two lower panels, we show the comparison of the three redshift histograms for $I_{AB} \leqslant 26$ (left panel) and $26 < I_{AB} \leqslant 28.5$ (right panel). Our GISSEL model, the CE model and the FLY99 model are shown by solid, dashed and dotted lines, respectively.

between the six spectra to reduce the aliasing effect due to SED sparse sampling.

Comparison of the two approaches with spectroscopic redshifts has been carried out by the authors: the uncertainties are typically $\sigma_z \simeq 0.10$–0.15 at $z \leqslant 1.5$ and reach $\sigma_z \simeq 0.20$–0.25 at higher redshift.

In the SLY97 analysis, only the four optical bands have been used to estimate the photometric redshifts. To carry out a fair comparison, we have set up a code based on a library similar to that used by SLY97 and we recomputed the photometric redshifts with the FLY99 catalogue – hereafter called the Coleman Extended model (CE). Comparison between the three methods is shown in Fig. 2 (upper panels). The three redshift distributions are shown in the lower panels of the the same figure. From these plots, we observe the following.

(i) For $I_{AB} \leqslant 26$ the three methods are compatible within $\Delta z \simeq 0.5$. A small number (~ 2 per cent) of catastrophic discrepancies ($\Delta z \geqslant 1$) is observed. Excluding these objects, we find rms dispersions $\sigma_z \simeq 0.12$ and $\sigma_z \simeq 0.23$ between the GISSEL and CE models at $z \leqslant 1.5$ and $1.5 < z \leqslant 5$, respectively. In the high-redshift range, a systematic shift is observed with

$\langle z_{GIS} - z_{CE}\rangle \simeq -0.15$. Between the GISSEL and FLY99 models, the dispersions are $\sigma_z \simeq 0.16$ and $\sigma_z \simeq 0.26$ at $z \leqslant 1.5$ and $1.5 < z \leqslant 5$, respectively, with a systematic shift in the high-redshift range $\langle z_{GIS} - z_{FLY99}\rangle \simeq +0.18$. These results are compatible with the uncertainties based on the spectroscopic sample. Finally, the three resulting redshift distributions are in good agreement.

(ii) For $I_{AB} \leqslant 28.5$ the number of objects with $\Delta z \geqslant 1$ increases and represents the 6 per cent of the full sample in both cases. Excluding these objects, we find dispersions $\sigma_z \simeq 0.18$ and $\sigma_z \simeq 0.26$ between the GISSEL and CE models at $z \leqslant 1.5$ and $1.5 < z \leqslant 5$, respectively. For the high-redshift range, a systematic shift is still observed with $\langle z_{GIS} - z_{CE}\rangle \simeq -0.11$. Comparing the GISSEL and FLY99 models, the dispersions are $\sigma_z \simeq 0.22$ and $\sigma_z \simeq 0.32$ at $z \leqslant 1.5$ and $1.5 < z \leqslant 5$, respectively, with a larger systematic shift in the high-redshift range $\langle z_{GIS} - z_{FLY99}\rangle \simeq +0.31$.

(iii) The large shift for $z \geqslant 1.5$ observed with FLY99 is due to a feature appearing in their redshift distribution with a large number of sources between $1.2 \leqslant z \leqslant 2$, not observed in the two other models (Fig. 2, lower right panel). The interval $1.2 \leqslant z \leqslant 2$ is critical for the photometric determination of the redshifts, due to the lack of strong features. In fact the Lyman-alpha break is not yet observed in the *F*300*W* band and the break at 4000 Å is located between the *F*814*W* and *J* bands. Therefore, the estimates rest basically on the continuum shape of the templates. As shown by FLY99 in their fig. 6, their photometric redshifts suffer from a systematic underestimate with respect to the spectroscopic ones around $z \simeq 2$. This may be due to an inadequacy of the UV extrapolation used by FLY99 in reproducing the UV shape of the high-z objects. This effect disappears at higher redshift because of the *U*-dropout effect. As a check, we have added to the four templates of FLY99 a spectrum of an irregular galaxy with constant star formation rate (with higher UV flux). In this case, the excess of galaxies with $1.2 \leqslant z \leqslant 2$ disappears and the objects are re-distributed in better agreement with the two other methods.

(iv) Our GISSEL model produces a smaller number of objects at $z \geqslant 3.5$ with respect to the two other approaches. The discrepant objects (found at lower redshift by the GISSEL code) are generally fitted by using a significant fraction of reddening excess: $\langle E(B-V)\rangle \simeq 0.3$. Note that in general objects found at $z \geqslant 3.5$ by the GISSEL code are also at high redshift with the other techniques.

Table 2. Comparison of our photometric redshifts with 106 spectroscopic redshifts up to $z = 5$ for different redshift intervals (Column 1). Here we consider only objects either with $|\Delta z| \equiv |z_{spec} - z_{phot}| \leqslant 1$ or with $|\Delta z| \leqslant 0.5$ (Column 2). The corresponding number of objects in each redshift interval is given in Column 3 and the associated dispersion σ_z in Column 4.

z range	$\lvert\Delta z\rvert$	N_{phot}/N_{spec}	σ_z
0.0–5.0	$\leqslant 1.0$	105/106	0.20
0.0–1.5	$\leqslant 1.0$	79/79	0.13
1.5–5.0	$\leqslant 1.0$	28/29	0.24
0.0–5.0	$\leqslant 0.5$	101/106	0.12
0.0–1.5	$\leqslant 0.5$	77/79	0.09
1.5–5.0	$\leqslant 0.5$	26/29	0.15

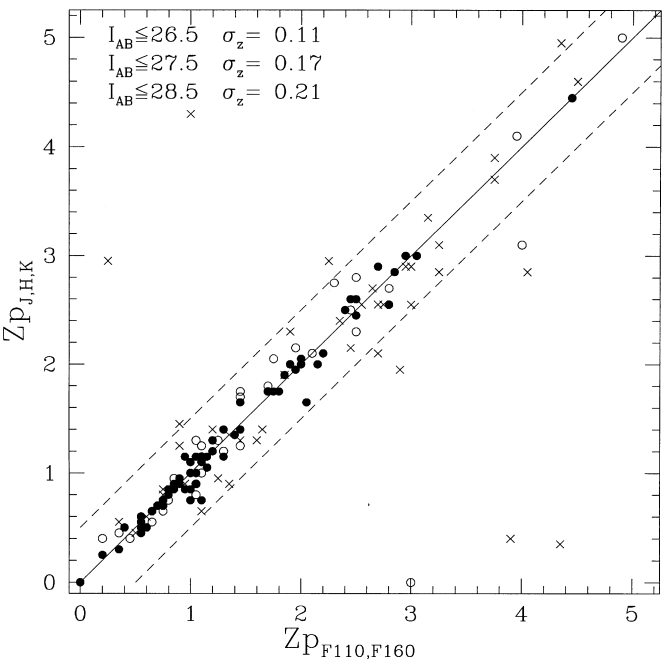


Figure 3. Comparison of the photometric redshifts obtained by using our GISSEL model and those obtained by replacing the *J*, *H* and *Ks* filters with the *F*110*W* and *F*160*W* filters for 164 objects. Filled circles represent galaxies with $I_{AB} \leqslant 26.5$, open circles represent galaxies with $26.5 \leqslant I_{AB} \leqslant 27.5$ and crosses refer to galaxies with $27.5 \leqslant I_{AB} \leqslant 28.5$. The dashed lines represent $\Delta z = 0.5$. The redshift dispersions σ_z for different magnitude limits are given inside the figure.

3.3 Comparison with the NICMOS *F*110*W* and *F*160*W* observations

Recently, deep NICMOS images have been obtained in the area corresponding to chip 4 of the WFPC2 camera in the *HDF North* (Thompson et al. 1999). The observations have been carried out in the two filters *F*110*W* and *F*160*W* and reach $F160W_{AB} \simeq 28.8$ (at 3σ). We have associated each NICMOS detection (from the published catalogue) with the FLY99 catalogue. We consider in our analysis the 164 objects detected in both NICMOS filters. These data provide a crucial check thanks to their depth and high spatial resolution and also to the spectral coverage of the *F*110*W* band. This filter fills the gap between the *F*814*W* filter and the standard *J* filter and makes it possible to detect the 4000 Å break at $z \geqslant 1.2$. We have recomputed the photometric redshifts with our GISSEL models using the four optical bands and replacing the *J*, *H*, and *Ks* filters with the *F*110*W* and *F*160*W* filters. The results are shown in Fig. 3. This subsample shows a good agreement between NICMOS and *J*, *H*, and *Ks* photometry and corroborates the reliability of the infrared measurements performed by FLY99. The redshift agreement in the range $0 \leqslant z \leqslant 5$ is better than $|\Delta z| = 0.5$ up to magnitudes $I_{AB} \leqslant 28.5$ and only 5/164 objects present discrepancies with $|\Delta z| \geqslant 1$.

3.4 Comparison with Monte Carlo simulations

As a final check we performed Monte Carlo simulations to study the effect of photometric errors on our redshift estimates. To do so we have added to the original fluxes of the 1067 galaxies of the FLY99 catalogue a gaussian random noise with rms equal to the flux uncertainties in each band. This operation has been repeated 20 times to produce a catalogue of approximately 21 000

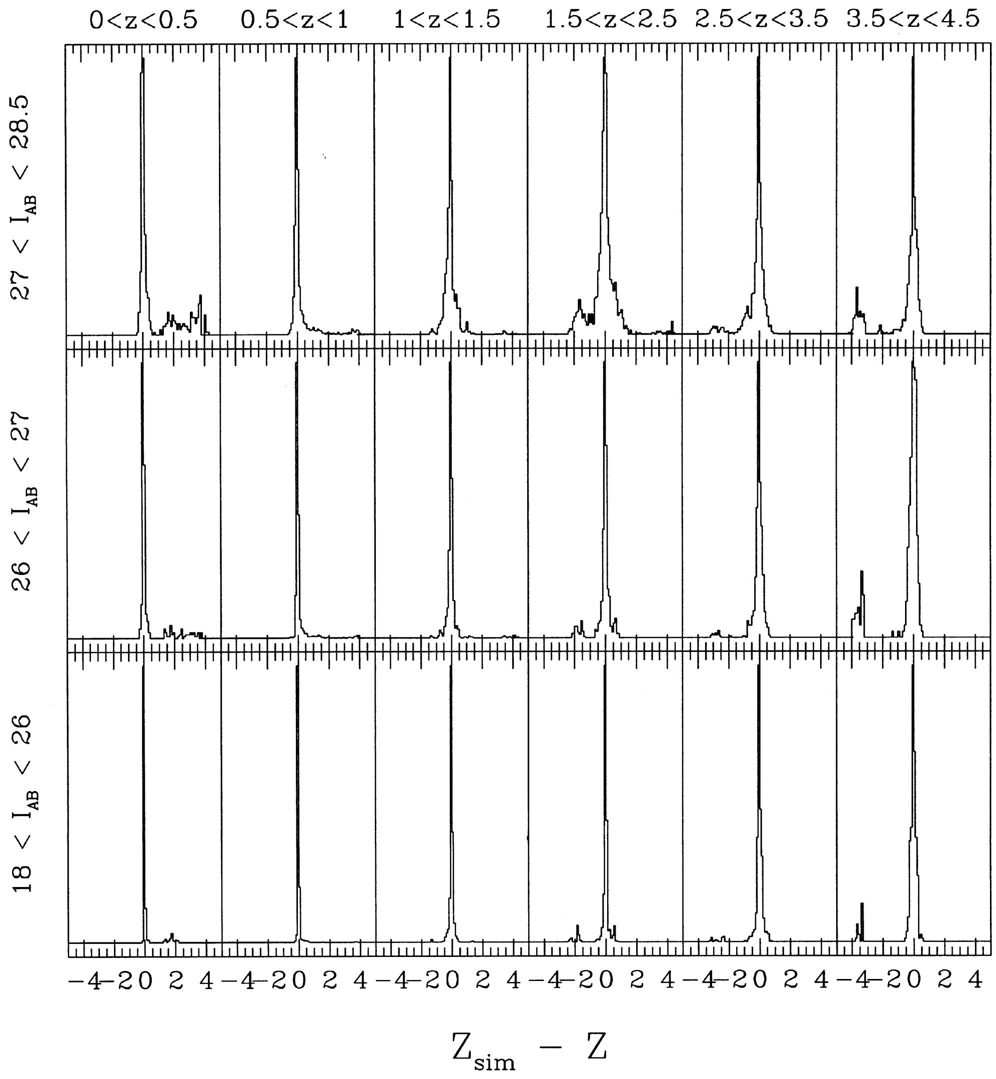


Figure 4. Effect of the photometric errors in the redshift estimates. We have built a catalogue of approximately 21 000 simulated galaxies by adding to the fluxes of each original object a Gaussian random noise with rms equal to the flux uncertainties in each band. The histograms of the differences ($\Delta z = z_{\rm sim} - z$) between the simulated redshift $z_{\rm sim}$ and the original one z are shown for different magnitude and redshift ranges.

simulated galaxies for which we have re-estimated the photometric redshifts with our code. In Fig. 4, we show the distribution of the differences Δz between the simulated redshifts $z_{\rm sim}$ and the original ones z ($\Delta z = z_{\rm sim} - z$) for different magnitude and redshift ranges. Several comments can be made from this figure.

(i) The median value of the redshift difference is very close to zero ($\leqslant 0.05$) for any magnitude and redshift range. The dispersion around the peak, σ_z, is larger for larger magnitudes and redshifts. In Table 3 we report σ_z for galaxies with $I_{\rm AB} \leqslant 28.5$ for different redshift ranges. These dispersions are compatible with the observed ones based on the comparison made above between different codes.

(ii) Table 3 also reports the number of simulated galaxies put in a redshift bin different from their original one because of the photometric errors (Column 3). These results show that the number of lost original galaxies varies between 15 and 25 per cent at any redshift for $I_{\rm AB} \leqslant 28.5$. In the redshift range $0 \leqslant z \leqslant 0.5$, the discrepant objects are distributed in a high redshift tail between $1 \leqslant z \leqslant 4$. For the three bins with $z \geqslant 1.5$, the discordant objects are preferentially located in a secondary peak at low z ($0 \leqslant z_{\rm sim} \leqslant 1$).

Table 3. Contamination effects for different redshift intervals computed from Monte Carlo simulations. Column 1 indicates the redshift range. Column 2 reports the dispersions σ_z around the higher peak in the distribution of the redshift differences Δz for simulated galaxies up to $I_{\rm AB} = 28.5$. Column 3 shows the fractions of objects which are outside the original redshift bin (Lost). Column 4 reports the contamination by objects belonging to another original redshift bin (Cont.). Finally Column 5 reports the contamination by objects belonging to the original adjacent redshift bins (Adj. cont.).

		Simulations	$I_{\rm AB} < 28.5$	
z range	σ_z	Lost (per cent)	Cont. (per cent)	Adj. cont. (per cent)
0.0–0.5	0.20	19.3	30.2	9.4
0.5–1.0	0.20	12.2	11.5	9.3
1.0–1.5	0.25	25.0	15.5	12.5
1.5–2.5	0.35	22.7	27.0	22.7
2.5–3.5	0.32	22.1	19.2	16.9
3.5–4.5	0.26	26.3	21.8	15.1

(iii) The galaxies lost from an original bin are a contaminating factor for the others. We can estimate for each bin this contamination which is also reported in Table 3 (Column 4). In the same table the contaminating fraction due only to the adjacent bins is reported (Column 5). We can see that the contamination plays a different role at different redshifts. For $0 \leqslant z \leqslant 0.5$, the contamination is quite large ($\simeq 30$ per cent) and it is not due to the adjacent bin (representing only one third of the total). In this case the main source of contamination is high-redshift galaxies put at low redshift. For the other bins the contamination is close to 20 per cent and is essentially due to the adjacent bins.

4 THE ANGULAR CORRELATION FUNCTION

4.1 Definition of the redshift bin sizes and subsamples

We have limited our analysis to the region of the *HDF* with the highest signal-to-noise ratio, excluding the area of the PC, the outer part of the three WFPC and the inner regions corresponding to the junction between each chip. In this area we included in our sample all galaxies brighter than $I_{\mathrm{AB}} \simeq 28.5$. This procedure leads to a slight reduction of the overall number of galaxies: our final sample contains 959 out of the 1023 original ones.

To compute the angular correlation function (ACF) correctly, the following details have to be taken into account:

(i) the relatively small field of view of the *HDF* (the angular distance corresponds to $\simeq 1\,h^{-1}$ Mpc at $z \geqslant 1$, with $q_0 = 0.5$);
(ii) the accuracy of the photometric redshifts;
(iii) the number of objects in each redshift bin, in order to reduce the shot noise and achieve sufficient sensitivity to the clustering signal.

As a consequence, relatively large redshift bins are required: according to Fig. 2 and Table 3, a minimum redshift bin size of $\Delta z = 0.5$ (corresponding to $\Delta z \sim 2 \times \sigma_z$) is required for $z \leqslant 1.5$. At higher redshift, due to the uncertainties in the redshift and the relatively low surface densities, a more appropriate bin size is $\Delta z = 1$. Moreover, these large bin sizes can reduce the effects of redshift distortion and, most important, attenuate the sample variance effect caused by the small area covered by the *HDF North* (approximately 4 arcmin^2). A refined approach to treat the sample variance has been recently proposed by Colombi, Szapudi & Szalay (1998).

Finally, we note that the contamination discussed in the previous section can introduce a dilution of the clustering signal. In the worst case, assuming that the contaminating population is uncorrelated, it introduces a dilution of about $(1-f)^2$ (where f corresponds to the contaminating fraction reported in Table 3). This correction factor has been used to define upper limits to the clustering estimates which are shown in the following figures.

4.2 The computation of the angular correlation function

The angular correlation function $\omega(\theta)$ is related to the excess of galaxy pairs in two solid angles separated by the angle θ with respect to a random Poisson distribution. The angular separation used for the computation of $\omega(\theta)$ covers the range from 5 arcsec up to 80 arcsec. We use logarithmic bins with steps of $\Delta \log \theta = 0.3$. The lower limit makes it possible to avoid a spurious signal at small scales due to the multi-deblending of resolved bright spirals and irregulars, the upper cut-off is almost half the size of the *HDF* and corresponds to the maximum separation where the ACF provides a reliable signal.

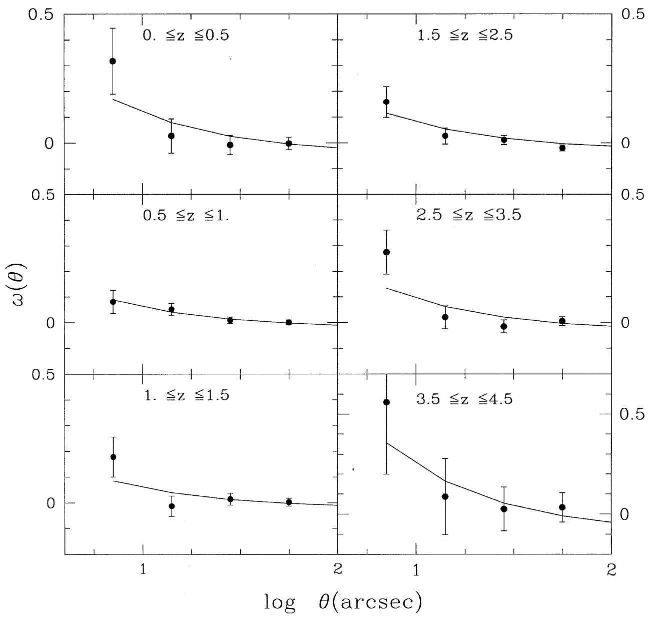


Figure 5. The angular correlation functions $\omega(\theta)$, computed with the estimator of Landy & Szalay (1993), for galaxies with $I_{\mathrm{AB}} \leqslant 28.5$ measured for different redshift ranges (as specified in each panel). The uncertainties are Poisson errors. The solid lines show the best fits obtained by assuming $\omega(\theta) = A_\omega\ \theta^{-\delta}$, with a fixed slope $\delta = 0.8$.

To derive the ACF in each redshift interval, we used the estimator defined by Landy & Szalay (1993):

$$\omega_{\mathrm{est}}(\theta) = \frac{DD(\theta) - 2DR(\theta) + RR(\theta)}{RR(\theta)}, \tag{3}$$

where DD is the number of distinct galaxy–galaxy pairs, DR is the number of galaxy–random pairs and RR refers to random–random pairs with separation between θ and $\theta + \Delta\theta$. The random catalogue contains 20 000 sources covering the same area of our sample. In Fig. 5 we show the measured ACF for each redshift bin. The uncertainties are Poisson errors as shown by Landy & Szalay (1993) for this estimator.

Adopting a power-law form for the ACF as $\omega(\theta) = A_\omega \theta^{-\delta}$, we derive the amplitude A_ω assuming $\delta = \gamma - 1 = 0.8$. Here γ is the slope of the spatial correlation function, which is also assumed to follow a power-law relation. Formally, we can use both A_ω and δ as free parameters to be obtained from the least-square fitting, but, due to the limited sample, we prefer to fix δ and leave as free parameter only A_ω. The value of the slope we assume is larger than the estimates obtained by Le Fèvre et al. (1996) in the analysis of the CFRS catalogue (which covers the interval $0 \leqslant z \leqslant 1$), and is smaller than the estimates obtained for LBGs by G98 at $z \simeq 3$. Nevertheless, the adopted value is still consistent with the respective uncertainties. The value of the slope could also depend on the magnitude, as discussed by Postman et al. (1998).

To estimate the amplitude of the ACF, owing to the small size of the field, we introduce the integral constraint IC in our fitting procedure as $\omega_{\mathrm{est}} \simeq \omega_{\mathrm{true}} - IC = A_\omega \times (\theta^{-0.8} - B)$. The quantity $IC = A_\omega \times B$ has been computed by a Monte Carlo method using the same geometry of the *HDF* and masking the excluded regions. In this computation, we adopt the same value for the slope ($\delta = 0.8$) and we derive $B = 0.044$ (for θ measured in arcsec). The

Table 4. The amplitude of $\omega(\theta)$ at 10 arcsec (A_ω) for different redshift bins with $I_{AB} \leqslant 28.5$. Column 1: the redshift range. Column 2: the number of galaxies in the redshift bin. Column 3: the amplitude of the ACF at 10 arcsec. Columns 4, 5, 6: the comoving correlation length r_0 (in h^{-1} Mpc) as derived from the Limber equation for three different cosmological models (EdS model, open model with $\Omega_{0m} = 0.3$ and a flat model with $\Omega_{0m} = 0.3$ and cosmological constant). All the listed values are not corrected for the contamination.

z range	Number of galaxies	A_ω (at 10 arcsec)	r_0 $\Omega_{0m}=1, \Omega_{0\Lambda}=0$	r_0 $\Omega_{0m}=0.3, \Omega_{0\Lambda}=0$	r_0 $\Omega_{0m}=0.3, \Omega_{0\Lambda}=0.7$
0.0–0.5	96	0.17 ± 0.09	1.63 ± 0.47	1.77 ± 0.51	1.93 ± 0.56
0.5–1.0	294	0.09 ± 0.03	1.37 ± 0.25	1.69 ± 0.31	1.93 ± 0.36
1.0–1.5	157	0.09 ± 0.05	1.21 ± 0.41	1.64 ± 0.56	1.82 ± 0.63
1.5–2.5	202	0.12 ± 0.04	1.92 ± 0.38	3.06 ± 0.61	3.07 ± 0.61
2.5–3.5	142	0.13 ± 0.06	1.69 ± 0.41	3.06 ± 0.75	2.78 ± 0.68
3.5–4.5	35	0.35 ± 0.25	2.56 ± 1.01	5.29 ± 2.08	4.28 ± 1.69
0.0–6.0	959	0.03 ± 0.01			

best fits for the ACF in each redshift bin are shown as solid lines in Fig. 5. The amplitudes A_ω obtained by the best fits are listed in Table 4 with the adopted magnitude limits and the number of galaxies used. We also give the measured amplitude for the galaxies with $I_{AB} \leqslant 28.5$ and with $0 \leqslant z \leqslant 6$. None of these values is corrected for the contamination factor.

In Fig. 6 we compare our values of A_ω (at 10 arcsec) to other published data (Connolly et al. 1998; G98; Magliocchetti & Maddox 1999). The values of A_ω take into account the adopted redshift bin sizes. At a given redshift, a larger Δz implies smaller A_ω owing to the increasing number of foreground and background galaxies with respect to the unchanged number of physically correlated pairs ($A_\omega \propto \Delta z^{-1}$; see e.g. Connolly et al. 1998). Then, if we assume that A_ω does not strongly evolve inside the redshift bin, we can correct the original amplitudes by using $A_\omega \times \Delta z$, which allows a more direct comparison. From this figure we note that our results are in good agreement with those of Connolly et al. (1998) and slightly smaller than those for LBGs obtained by G98. The agreement with Magliocchetti & Maddox (1999) is worse but still consistent with both error bars. In this figure we show the possible effect of the contamination factor discussed in the previous section. This correction increases all the values that should be regarded as upper limits owing to the basic assumption that the contaminating population is uncorrelated. Moreover, we notice that our estimate in the redshift bin $0 \leqslant z \leqslant 0.5$ can be affected by the lack of nearby bright galaxies in the *HDF*. For this reason, this point will be not considered in the following comparison between observational results and model predictions.

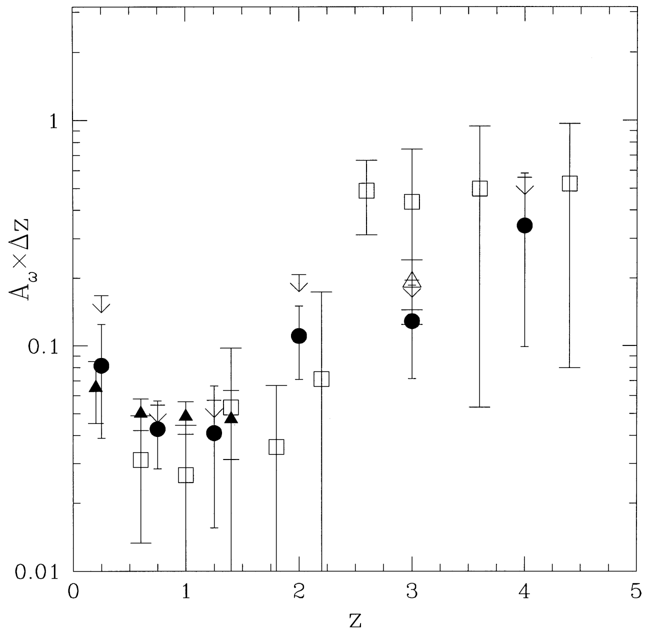


Figure 6. The amplitude of $\omega(\theta)$ at 10 arcsec (A_ω) as a function of redshift. The values are rescaled to the same Δz by applying $A_\omega \times \Delta z$ for a direct comparison (see text). Filled circles represent our results for $I_{AB} \leqslant 28.5$. The arrows show our measurements corrected for the contamination factor and should be considered as upper limits. The filled triangles show the values obtained by Connolly et al. (1998). The open triangle is the value for the LBG sample (Giavalisco et al. 1998) and open squares refer to the values obtained by Magliocchetti & Maddox (1999).

5 COMPARISON WITH THEORETICAL MODELS

5.1 The formalism

We can now predict the behaviour of the angular correlation function $\omega(\theta)$ for our galaxy sample in various cosmological structure formation models. The angular two-point function for a sample extended in the redshift direction over an interval $\mathcal{Z}$ can be written in terms of the spatial correlation function using the relativistic Limber equation (Peebles 1980). We adopt here the Limber formula as given in Matarrese et al. (1997), namely

$$\omega_{\rm obs}(\theta) = N^{-2} \int_{\mathcal{Z}} {\rm d}z \left(\frac{{\rm d}r}{{\rm d}z}\right)^{-1} \mathcal{N}^2(z) \int_{-\infty}^{\infty} {\rm d}u\, \xi_{\rm gal}[r(u,\theta,z),z], \quad (4)$$

where $r(u,\theta,z) = \sqrt{u^2 + r^2(z)\theta^2}$, in the small-angle approximation (e.g. Peebles 1980).

The relation between the comoving radial coordinate r and the redshift z is given with whole generality by

$$r(z) = \frac{c}{H_0\sqrt{|\Omega_{0\mathcal{R}}|}} \mathcal{S}\left(\sqrt{|\Omega_{0\mathcal{R}}|}\int_0^z [(1+z')^2(1+\Omega_{0m}z') - z'(2+z')\Omega_{0\Lambda}]^{-1/2}\,{\rm d}z'\right), \quad (5)$$

where $\Omega_{0\mathcal{R}} \equiv 1 - \Omega_{0m} - \Omega_{0\Lambda}$, with Ω_{0m} and $\Omega_{0\Lambda}$ the density parameters for the non-relativistic matter and cosmological constant components, respectively. In this formula, for an open universe model, $\Omega_{0\mathcal{R}} > 0$, $S(x) \equiv \sinh(x)$, for a closed universe, $\Omega_{0\mathcal{R}} < 0$, $S(x) \equiv \sin(x)$, while in the EdS case, $\Omega_{0\mathcal{R}} = 0$, $S(x) \equiv x$.

In the Limber equation above, $\mathcal{N}(z)$ is the redshift distribution of the catalogue (whose integral over the entire redshift interval is N), which is given by $\mathcal{N}(z) = \int_{\mathcal{M}} {\rm d}\ln M \mathcal{N}(z, M)$, with

$\mathcal{N}(z, M) = 4\pi g_c(z)\phi(z, M)\bar{n}_c(z, M)$ and $\bar{n}_c(z,M)$ is the expected number of galaxies per comoving volume at redshift z; $\phi(z, M)$ is the isotropic catalogue selection function. The quantity $\mathcal{N}(z, M)$ represents the number of objects actually present in the catalogue, with redshift in the range z, $z + \mathrm{d}z$ and intrinsic properties (like mass, luminosity, ...) in the range M, $M + \mathrm{d}M$ ($\mathcal{M}$ representing the overall interval of variation of M). In the latter integral we also defined the comoving Jacobian

$$g_c(z) \equiv r^2(z)\left[1 + \frac{H_0^2}{c^2}\Omega_{0\mathcal{R}}\, r^2(z)\right]^{-1/2} \frac{\mathrm{d}r}{\mathrm{d}z}. \tag{6}$$

In what follows, we assume a simple model for our galaxy distribution, where galaxies are associated in a one-to-one correspondence to their hosting dark-matter haloes. The advantage of this model is that haloes can be simply characterized by their mass M and formation redshift z_f. Since haloes merge continuously into larger mass ones one can safely assume that their formation redshifts coincide with those observed, namely $z_\mathrm{f} = z$. This simple model of galaxy clustering was named the 'transient' model in Matarrese et al. (1997) and Moscardini et al. (1998); Coles et al. (1998) adopted it to describe the clustering of LBGs. The application of this model is more appropriate at high redshift where merging dominates while at low redshift it can only be a rough approximation. Recently Baugh et al. (1999) showed that this simple model under-predicts the clustering properties at low redshift because it does not take into account the possibility that a single halo can host more than one galaxy. Indeed, as discussed in Moscardini et al. (1998), a 'galaxy conserving' bias model is likely to provide a better description of galaxy clustering evolution at low redshift.

In practice, in our modelling we select a minimum mass M_min for the haloes hosting our galaxies, i.e. we take $\phi(z, M) = \theta(M - M_\mathrm{min})$, with θ the Heaviside step function, and we compute the corresponding value of the effective bias b_eff (see the equation below) at each redshift. In what follows we consider two possibilities: (*i*) M_min fixed to a sensible value (we show results obtained by using 10^{10}, 10^{11} and $10^{12}\,h^{-1}\,\mathrm{M}_\odot$); (*ii*) $M_\mathrm{min} = M_\mathrm{min}(z)$ chosen to reproduce a relevant set of observational data. For the latter case, we adopt two different strategies: in the first case we assume $M_\mathrm{min}(z)$ such that the theoretical $\mathcal{N}(z)$ fits that observed in each redshift bin (e.g. Mo & Fukugita 1996; Moscardini et al. 1998; A98; Mo, Mao & White 1999); in the second case we adopt at any redshift the median of the mass distribution estimated by our GISSEL model. Actually this model gives a rough estimate of the baryonic mass. To convert it to the mass of the hosting dark-matter halo we multiply by a factor 10. This value corresponds to a baryonic fraction close to that predicted by the standard theory of primordial nucleosynthesis. Variations in the range from 5 to 20 produce only small changes in the following results.

As a first, though accurate, approximation the galaxy spatial two-point function can be taken as being linearly proportional to that of the mass, namely $\xi_\mathrm{gal}(r, z) \simeq b_\mathrm{eff}^2(z)\xi_\mathrm{m}(r, z)$, where

$$b_\mathrm{eff}(z) \equiv \mathcal{N}(z)^{-1} \int_{\mathcal{M}} \mathrm{d}\ln M' \mathcal{N}(z, M') b(M', z) \tag{7}$$

is the effective bias of our galaxy sample and ξ_m the matter covariance function.

The bias parameter $b(M, z)$ for haloes of mass M at redshift z in a given cosmological model can be modelled as (Mo & White 1996)

$$b(M, z) = 1 + \frac{1}{\delta_c}\left(\frac{\delta_c^2}{\sigma_M^2 D_+^2(z)} - 1\right), \tag{8}$$

where σ_M^2 is the linear mass-variance averaged over the scale M, extrapolated to the present time ($z = 0$), δ_c the critical linear overdensity for spherical collapse ($\delta_c = \mathrm{const} = 1.686$ in the EdS case, while it depends slightly on z for more general cosmologies) and $D_+(z)$ is the linear growth factor of density fluctuations (e.g. $D_+(z) = (1 + z)^{-1}$ in the EdS case). In comparing our theoretical predictions on clustering with the data, we always adopt for the galaxy redshift distribution $\mathcal{N}(z)$ the observed one. Nevertheless, consistency requires that the predicted halo redshift distribution for a given minimum halo mass always exceeds (because of the effects of the selection function) the observed galaxy one. For the calculation of the effective bias, where we need $\mathcal{N}(z, M)$, one might adopt the Press & Schechter (1974) recipe to compute the comoving halo number density (per unit logarithmic interval of mass); it reads

$$\bar{n}_c(z, M) = \sqrt{\frac{2}{\pi}} \frac{\bar{\varrho}_0 \delta_c}{M D_+(z)\sigma_M} \left|\frac{\mathrm{d}\ln\sigma_M}{\mathrm{d}\ln M}\right| \exp\left[-\frac{\delta_c^2}{2D_+^2(z)\sigma_M^2}\right] \tag{9}$$

(with $\bar{\varrho}_0$ the mean mass density of the Universe at $z = 0$).

However, a number of authors have recently shown that the Press–Schechter formula does not provide an accurate description of the halo abundance either in the large- or the small-mass tails (see e.g. the discussion in Sheth & Tormen 1999). Also, the simple Mo & White (1996) bias formula of equation (7) has been shown not to reproduce the correlation of low-mass haloes correctly in numerical simulations. Several alternative fits have been recently proposed (Jing 1998; Porciani, Catelan & Lacey 1999; Sheth & Tormen 1999; Jing 1999). An accurate description of the abundance and clustering properties of the dark-matter haloes corresponding to our galaxy population will be obtained here by adopting the relations introduced by Sheth & Tormen (1999), which have been obtained by fitting to the distribution of the halo population of the GIF simulations (Kauffmann et al. 1999): this technique allows the simultaneous improvement of the performance of both the mass function and the bias factor. The relevant formulae, replacing equations (8) and (9) above, read

$$b(M, z) = 1 + \frac{1}{\delta_c}\left(\frac{a\delta_c^2}{\sigma_M^2 D_+^2(z)} - 1\right) + \frac{2p}{\delta_c} \times \left(\frac{1}{1 + [\sqrt{a}\,\delta_c/(\sigma_M D_+(z))]^{2p}}\right) \tag{10}$$

and

$$\bar{n}_c(z, M) = \sqrt{\frac{2aA^2}{\pi}} \frac{\bar{\varrho}_0 \delta_c}{M D_+(z)\sigma_M} \left[1 + \left(\frac{D_+(z)\sigma_M}{\sqrt{a}\,\delta_c}\right)^{2p}\right] \times \left|\frac{\mathrm{d}\ln\sigma_M}{\mathrm{d}\ln M}\right| \exp\left[-\frac{a\delta_c^2}{2D_+^2(z)\sigma_M^2}\right], \tag{11}$$

respectively. In these formulae $a = 0.707$, $p = 0.3$ and $A \approx 0.3222$, while one would recover the standard (Mo & White and Press & Schechter) relations for $a = 1$, $p = 0$ and $A = 1/2$.

The computation of the clustering properties of any class of objects is completed by the specification of the matter covariance function $\xi_\mathrm{m}(r, z)$ and its redshift evolution. To this purpose we follow Matarrese et al. (1997) and Moscardini et al. (1998), who used an accurate method, based on the Hamilton et al. (1991) original ansatz to evolve $\xi_\mathrm{m}(r, z)$ into the fully nonlinear regime. Specifically, we use here the fitting formulae proposed by Peacock & Dodds (1996).

As recently pointed out by various authors (e.g. Villumsen

1996; Moessner, Jain & Villumsen 1998), when the redshift distribution of faint galaxies is estimated by applying an apparent magnitude limit criterion, magnification bias due to weak gravitational lensing would modify the relation between the intrinsic galaxy spatial correlation function and the observed angular one. Modelling this effect within the present scheme would be highly desirable, but is certainly beyond the scope of our work. Nevertheless, we note that this magnification bias would generally lead to an increase of the apparent clustering of high-z objects above that produced by the intrinsic galaxy correlations, by an amount depending on the amplitude of the fluctuations of the underlying matter distribution.

5.2 Structure formation models

We consider here a set of cosmological models belonging to the general class of Cold Dark Matter (CDM) scenarios. The linear power-spectrum for these models can be represented by $P_{\rm lin}(k,0) \propto k^n T^2(k)$, where we use the fit for the CDM transfer function $T(k)$ given by Bardeen et al. (1986), with 'shape parameter' Γ defined as in Sugiyama (1995). To fix the amplitude of the power spectrum (generally parameterized in terms of σ_8, the rms fluctuation amplitude inside a sphere of $8\,h^{-1}$ Mpc) we either attempt to fit the local cluster abundance, following the Eke, Cole & Frenk (1996) analysis of the temperature distribution of X-ray clusters (Henry & Arnaud 1991), or the level of fluctuations observed by COBE (Bunn & White 1997). In particular, we consider the following models: (1) a version of the standard CDM (SCDM) model with $\sigma_8 = 0.52$, which reproduces the local cluster abundance but is inconsistent with COBE data; (2) the so-called τCDM model (White, Gelmini & Silk 1995), with shape parameter $\Gamma = 0.21$; (3) a COBE normalized tilted model, hereafter called TCDM (Lucchin & Matarrese 1985), with $n = 0.8$, $\sigma_8 = 0.52$ and high (10 per cent) baryonic content (e.g. White et al. 1996; Gheller, Pantano & Moscardini 1998) – the normalization of scalar perturbations, which takes into account the production of gravitational waves predicted by inflationary theories (e.g. Lucchin, Matarrese & Mollerach 1992; Lidsey & Coles 1992), allows the simultaneous fitting of both the CMB fluctuations observed by COBE and the local cluster abundance. The three above models are all flat and without cosmological constant. We also consider here: (4) a cluster-normalized open CDM model (OCDM) with matter density parameter $\Omega_{0\rm m} = 0.3$, and $\sigma_8 = 0.87$, which is also consistent with COBE data; and finally (5) a cluster-normalized low-density CDM model (ΛCDM), with $\Omega_{0\rm m} = 0.3$, but with a flat geometry provided by the cosmological constant, with $\sigma_8 = 0.93$, which is also consistent with COBE data. A summary of the parameters of the cosmological models used here is given in Table 5.

Table 5. The parameters of the cosmological models. Column 2: the present matter density parameter $\Omega_{0\rm m}$; Column 3: the present cosmological constant contribution to the density $\Omega_{0\Lambda}$; Column 4: the primordial spectral index n; Column 5: the Hubble parameter h; Column 6: the shape parameter Γ; Column 7: the spectrum normalization σ_8.

Model	$\Omega_{0\rm m}$	$\Omega_{0\Lambda}$	n	h	Γ	σ_8
SCDM	1.0	0.0	1.0	0.50	0.45	0.52
τCDM	1.0	0.0	1.0	0.50	0.21	0.52
TCDM	1.0	0.0	0.8	0.50	0.41	0.52
OCDM	0.3	0.0	1.0	0.65	0.21	0.87
ΛCDM	0.3	0.7	1.0	0.65	0.21	0.93

5.3 Results

In Fig. 7 we compare the observed amplitude of the ACF with the predictions of the various cosmological models. For consistency with the analysis performed on the observational data shown in the previous section, here the theoretical results have been obtained by fitting the data in the same range of angular separation and using the same stepping $\Delta \log \theta = 0.3$. A fixed slope of $\delta = 0.8$ is also used in the following analysis. Notice that this value is only a rough estimate of the best fit slopes: generally the resulting values are smaller ($\delta \simeq 0.6$) in all redshift intervals and for all the models. The discrepancy is higher for TCDM and τCDM ($\delta \simeq$ 0.3–0.4) and can lead to some ambiguity in the interpretation of the results (see the discussion on the effective bias below).

In each panel the solid lines show the results obtained when we use different (but constant in redshift) values of $M_{\rm min}$ (10^{10}, 10^{11} and $10^{12}\,h^{-1}$ M$_\odot$ from bottom to top). These results can be regarded as a reference on what is the minimum mass of the galaxies necessary to reproduce the observed clustering strength. However, the assumption that the catalogue samples the same class of objects at any redshift, i.e. objects with the same typical minimum mass, cannot be realistic. In fact, we expect that at high redshifts the sample tends to select more luminous, and on average more massive, objects than at low redshift. This is supported by the distribution of the galaxy masses inferred by the GISSEL model, shown in Fig. 8. The solid line, which represents the median mass, is an increasing function of redshift: from $z \simeq 0$ to $z \simeq 4$ its value changes by at least a factor of 30. In Fig. 8 we also show the masses necessary to reproduce at any redshift the observed galaxy density. In general, they are compatible with the GISSEL distribution but the redshift dependence is different for the various cosmological models considered here. For EdS universe models (left panel) the different curves are quite similar and almost constant with typical values of $10^{10.5}\,h^{-1}\,{\rm M}_\odot$. On the contrary, for OCDM and ΛCDM models (shown in the right panel) $M_{\rm min}(z)$ is an increasing function of redshift: at $z \simeq 0$, $M_{\rm min} \simeq 10^{10}\,h^{-1}\,{\rm M}_\odot$, while at $z \simeq 4$, $M_{\rm min} \simeq 10^{11.5}\,h^{-1}\,{\rm M}_\odot$. The amplitudes of the ACF obtained by adopting these $M_{\rm min}(z)$ values are also shown in Fig. 7.

In general, all the models are able to reproduce the qualitative behaviour of the observed clustering amplitudes, i.e. a decrease from $z = 0$ to $z \simeq$ 1–1.5 and an increase at higher redshift. The EdS models are in rough agreement with the observational results when a minimum mass of $10^{11}\,h^{-1}\,{\rm M}_\odot$ is used at any redshift. As discussed above, this mass is slightly larger than the one required to fit the observed $\mathcal{N}(z)$. The situation for OCDM and ΛCDM models is different. The amount of clustering measured would require that the involved objects have, at redshift $z \leqslant$ 1–1.5, minimum masses smaller than $10^{10}\,h^{-1}\,{\rm M}_\odot$, at redshift $1.5 \lesssim z \lesssim 3$, minimum masses of the order of $10^{11.5}\,h^{-1}\,{\rm M}_\odot$, while, at $z \simeq 4$, $M_{\rm min} \geqslant 10^{12}\,h^{-1}\,{\rm M}_\odot$ is needed to reproduce the clustering strength. These small values at low redshift are probably due to the kind of biasing model adopted in an epoch when merging starts to be less important. This is particularly true for open models and flat models with a large cosmological constant, where the growth of perturbations is frozen by the rapid expansion of the universe. On the contrary, the need to explain the high amplitude of clustering at $z \simeq 4$ with very massive objects can be in conflict with the observed abundance of galaxies at this redshift, which requires smaller minimum masses.

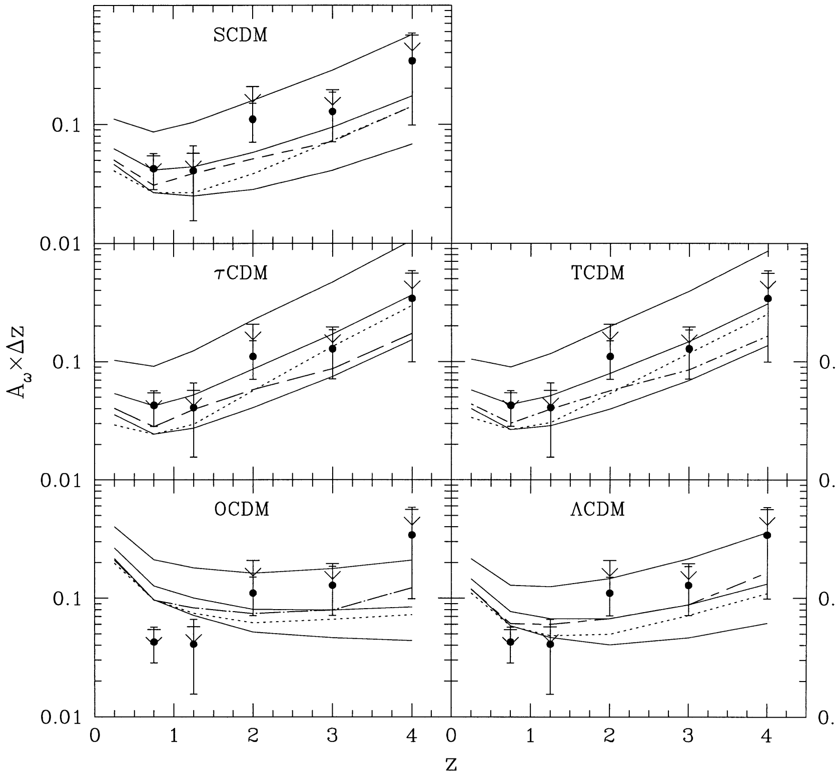


Figure 7. Comparison of the observed A_ω (filled circles for $I_{AB} \leqslant 28.5$, arrows for the upper-limits estimates) with the prediction of the various theoretical models described in the text. The solid lines show the measurements expected when a minimum mass $M_{min} = 10^{10}$, 10^{11} and $10^{12}\,h^{-1}\,M_\odot$ is assumed; the lower curves refer to smaller masses. The dotted lines show the predictions obtained by using the median masses at any redshift estimated by our GISSEL model shown in Fig. 8 (baryonic masses are translated into the masses of the hosting dark-matter haloes by multiplying by a factor of 10; see text). The dashed curves correspond to models where the masses necessary to reproduce the observed density of objects in each redshift bin are used. The dashed curve types used for the different cosmological models are the same as those used in Fig. 8.

If the spatial correlation function can be written in the simple form $\xi_{gal}(r,z) = [r/r_0(z)]^{-\gamma}$, it is possible to obtain the comoving correlation length $r_0(z)$ and the rms galaxy density fluctuation $\sigma_8^{gal}(z)$, with the assumption that the clustering does not strongly evolve inside each redshift bin used for the amplitude measurements (see Magliocchetti & Maddox 1999 for the relevant formulae in the framework of different cosmological models).

The values for the comoving $r_0(z)$ obtained from our data are listed in Table 4 for three different cosmologies. In Fig. 9, we compare our values of r_0 as a function of z to a compilation of values taken from the literature. The results are given under the assumption of an EdS universe. From this figure, one can notice that r_0 shows a small decline from $z \simeq 0$ to $z \simeq 1$–1.5 followed by an increase at higher z. At $z \geqslant 2$ the clustering amplitude is comparable to or higher than that observed at $z \simeq 0.25$.

An implication of the results shown in this figure is that the evolution of galaxy clustering cannot be properly described by the standard parametric form: $\xi(r,z) = \xi(r, z=0)(1+z)^{-(3+\epsilon-\gamma)}$, where ϵ models the gravitational evolution of the structures. Owing to the dependence of the bias on redshift and mass, the evolution of galaxy clustering is related to the clustering of the mass in a complex way. This has already been noticed in G98 from the study of LBGs at $z \simeq 3$ (see also Moscardini et al. 1998 for a theoretical discussion of the problem).

In the plot of the correlation length r_0, we also present the results for $z < 1$ obtained by Le Fèvre et al. (1996) from the estimates of the projected correlation function of the CFRS. We do not show in the figure the correlation lengths obtained by Carlberg et al. (1997), who performed the same analysis using a K selected sample, because they adopted a different cosmological model. Their estimates with $q_0 = 0.1$ of r_0 are approximately a factor of 1.5 larger than the CFRS results in a comparable magnitude and redshift range. Our results are lower than these previous estimates and show that the objects selected by our catalogue at low redshift tend to have different clustering properties. This effect suggests a dependence of the clustering properties on the selection of the sample, which is even more evident at high redshift. In fact our

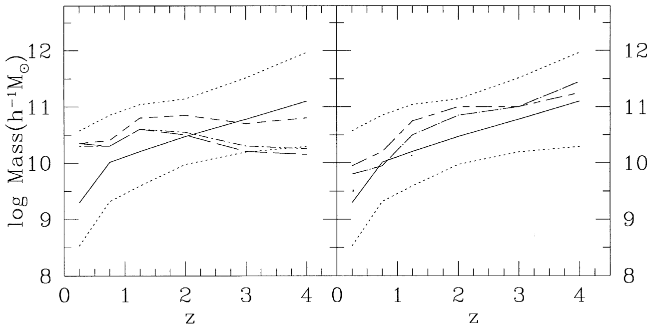


Figure 8. In both panels, the solid line shows, as a function of the redshift, the median of the distribution of the galaxy masses estimated by our GISSEL model; the lower and upper quartiles are shown by dotted lines. The baryonic masses are translated into the masses of the hosting dark-matter haloes multiplying by a factor 10. We show also the mass necessary to reproduce the observed density of objects at any z for SCDM (short-dashed line), τCDM (long-dashed line) and TCDM (short-dashed – dotted line) models in the left panel and for OCDM (long-dashed – dotted line) and ΛCDM (long-dashed – short–dashed line) models in the right panel.

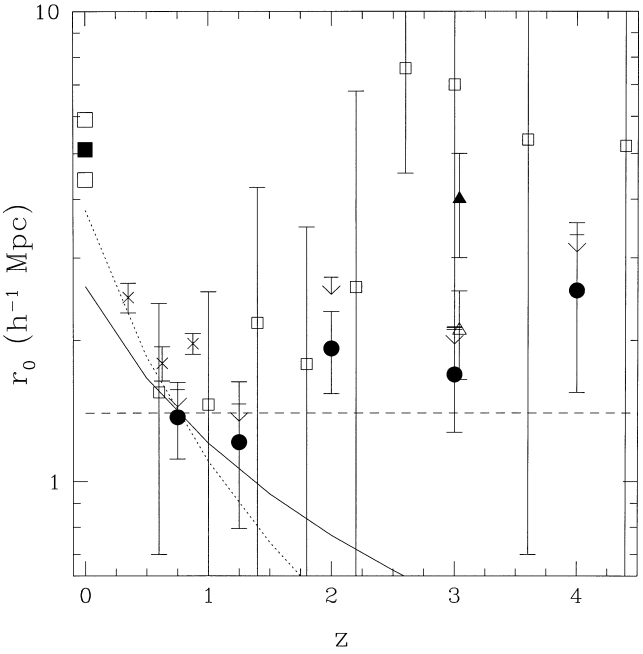


Figure 9. The measured comoving correlation length $r_0(z)$ (in h^{-1} Mpc) as obtained from the values of A_ω by assuming an EdS universe (filled circles; the arrows refer to the upper limits due to the contamination). Previous determinations are shown by the same symbols as in Fig. 6. We add also the values obtained from the analysis of the CFRS (Le Fèvre et al. 1996; cross symbols), from the count-in-cell analysis of the LBG catalogue (Adelberger et al. 1998, filled triangle) and the local values obtained from the APM survey (Loveday et al. 1995; full sample, E/S0 and Sp/Irr subsamples are shown by the filled square, and the high and low open squares, respectively). The curves show the evolution of the clustering using the ϵ model with three values of the parameter ϵ : $\epsilon = 0.8$ (linear growth of clustering, solid line); $\epsilon = 2$ (growth more rapid than the linear prediction, dotted line), $\epsilon = -1.2$ (fixed clustering in comoving coordinates, dashed line). The curves have been arbitrarily scaled to our observed value at $z = 0.75$.

value of r_0 at $z \simeq 3$ is smaller than that obtained in A98 and G98 for their LBG catalogues at the same redshift. To measure the clustering properties, A98 used a bright sample of 268 spectroscopically confirmed galaxies and derived $r_0 \simeq 4\,h^{-1}$ Mpc; G98 used a larger sample of 871 galaxies and derived a value two times smaller ($r_0 \simeq 2\,h^{-1}$ Mpc). Our value, referring to galaxies with $I_{AB} \leqslant 28.5$, is $r_0 \simeq 1.7\,h^{-1}$ Mpc. Notice that this value is a lower limit since it does not take into account the effects of contamination. All these reported values of the correlation length are obtained by assuming an EdS universe. This decrease of r_0 suggests that at fainter magnitudes we observe less massive galaxies that are intrinsically less correlated. This is in qualitative agreement with the prediction of the hierarchical galaxy formation scenario (e.g. Mo et al. 1999). On the contrary, such an interpretation is only marginally consistent with the reported higher value at $z \simeq 3$ of Magliocchetti & Maddox (1999), computed with the same FLY99 catalogue.

In order better to display the relation between the clustering strength and the abundance of a given class of objects (defined as haloes with mass larger than a given mass M_{min}) in Fig. 10, we show, for the different cosmological models, the relation between the predicted correlation length r_0 and the expected surface density, i.e. the number of objects per square arcminute. The quantity r_0 shown in this figure is defined as the comoving separation, where the predicted spatial correlation is unity; the number density is computed by suitably integrating the modified Press–Schechter formula (equation 11) over the given redshift range. In the left panel, showing the results for the interval $2.5 \leqslant z \leqslant 3.5$, we also plot the results obtained in this work (points at high density with their associated upper limits due to contamination effects) and those coming from the LBG analysis of A98 and G98 and corresponding to a lower abundance. All the models are able to reproduce the observed scaling of the clustering length with the abundance and no discrimination can be made between them. Similar conclusions have been reached by Mo et al. (1999). The right panel shows the same plot but at $z \simeq 4$, where the only observational estimates come from this work and from Magliocchetti & Maddox (1999). Here the situation seems to be

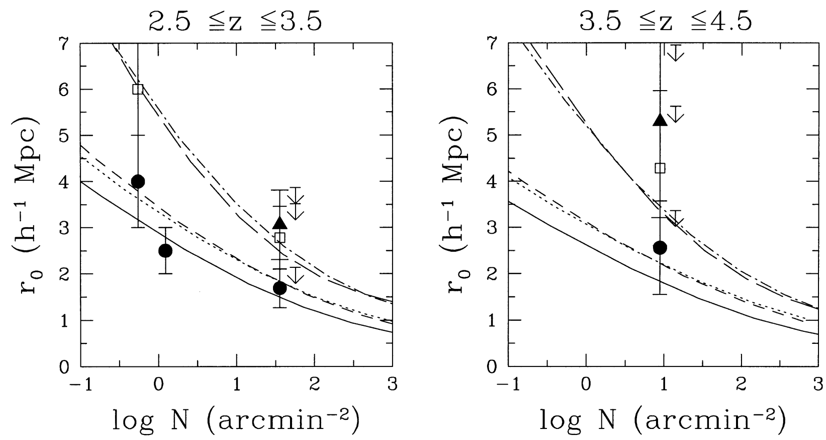


Figure 10. The comoving correlation length r_0 (in h^{-1} Mpc) as a function of surface density (defined as the expected number of objects per square arcminute) in two different redshift intervals: $2.5 \leqslant z \leqslant 3.5$ (left panel) and $3.5 \leqslant z \leqslant 4.5$ (right panel). Different lines refer to the predictions of various cosmological models: SCDM (solid lines), τCDM model (short-dashed lines), TCDM (dotted lines), OCDM(long-dashed lines) and ΛCDM (dotted–dashed lines). For comparison, we show (at $\log N \simeq 1.55$ in the left panel and at $\log N \simeq 0.95$ in the right panel) the results obtained in this work for three cosmologies: filled circles, filled triangles and open squares refer to EdS models, an open universe, and a flat universe with cosmological constant, respectively. The arrows refer to the upper limits due to contamination effects and are shifted by 0.2 in abscissa for clarity. In the left panel, we show also the results obtained from the analysis of the LBG clustering by Adelberger et al. (at $\log N \simeq -0.26$, for an Einstein–de Sitter universe and for a flat universe with cosmological constant), and from Giavalisco et al. (at $\log N \simeq 0.09$, only for an EdS model).

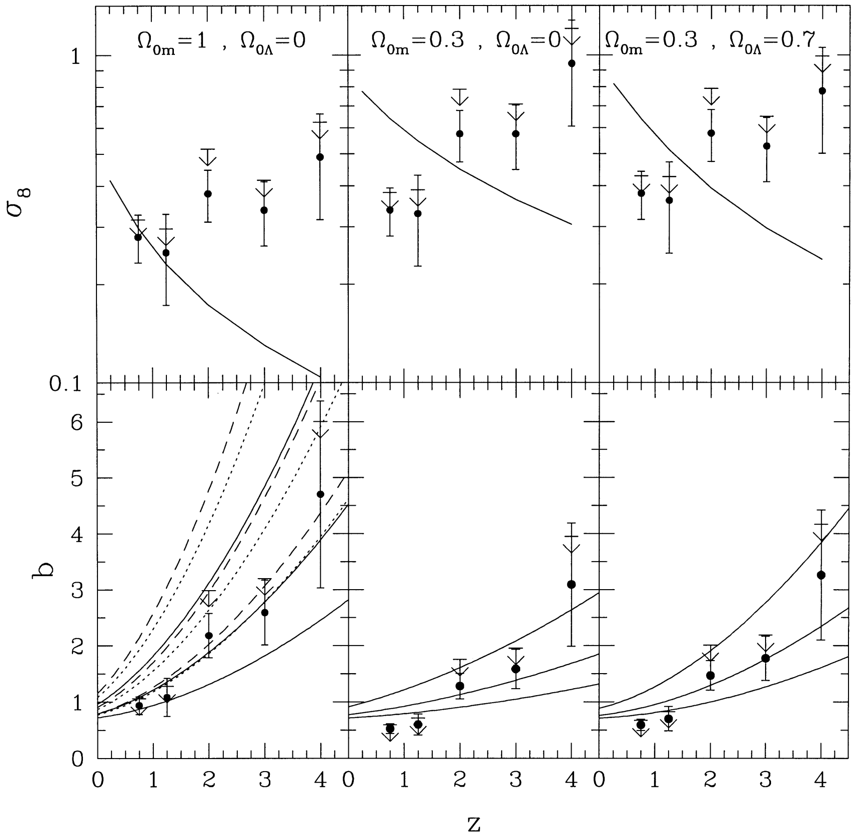


Figure 11. The upper panels show (data as filled circles and upper limits as arrows) our observed rms galaxy density fluctuation $\sigma_8^{\rm gal}(z)$ for three cosmologies (left: EdS universe; middle: open universe with $\Omega_{0m}=0.3$ and vanishing cosmological constant; right: flat universe with $\Omega_{0m}=0.3$ and cosmological constant). The solid lines show the rms mass density fluctuation $\sigma_8^{\rm m}(z)$ for the same cosmologies, as obtained by linear theory. The models are normalized to reproduce the local abundance of the galaxy clusters. The lower panels show the measured bias b as a function of redshift (with $b(z) \equiv \sigma_8^{\rm gal}(z)/\sigma_8^{\rm m}(z)$). The curves show, for different cosmological models, the theoretical effective bias computed with different values of minimum mass. For the EdS universe, three cosmological models are shown in the left panel: SCDM (solid lines), τCDM (dashed lines) and TCDM (dotted lines). The central and right panels refer to OCDM and ΛCDM models, respectively. We show results for $M_{\rm min}=10^{10}$, 10^{11} and $10^{12}\,h^{-1}\,{\rm M}_\odot$ (lower curves are for smaller masses).

more interesting. In fact the observed clustering is quite high and in the framework of the hierarchical models seems to require a low abundance for the relevant objects. This density starts to be in conflict with the observed one (which represents a lower limit due to the unknown effect of the selection function) for some of the models considered here, for example the OCDM model. Thus, if our results are confirmed by future observations, the combination of the clustering strength and galaxy abundance at redshift $z \simeq 4$ could be a discriminant test for the cosmological parameters.

An alternative way to study the clustering properties is given by the observed rms galaxy density fluctuation $\sigma_8^{\rm gal}$. Its redshift evolution is shown in the upper panels of Fig. 11 for three cosmological models: an Einstein–de Sitter universe (left panel); an open universe with $\Omega_{0m}=0.3$ and a vanishing cosmological constant (central panel); and a flat universe with $\Omega_{0m}=0.3$ and cosmological constant (right panel). In the same plot, we also show the theoretical predictions computed by using the linear theory when the cosmological models are normalized to reproduce the local cluster abundance. Since the corresponding values of $\sigma_8^{\rm m}$ at $z=0$ (reported in Table 5) are smaller than unity, we can safely compute the redshift evolution by adopting linear theory. As shown in Moscardini et al. (1998), the differences between these estimates and those obtained by using the fully nonlinear method described above are always smaller than 3 per cent at $z=0$ and consequently negligible at higher redshift. The comparison suggests that, while some anti-bias is present at low redshift, the high-redshift galaxies are strongly biased with respect to the dark matter. This observation strongly supports the theoretical expectation of biased galaxy formation with a bias parameter evolving with z.

Finally, the lower panels of Fig. 11 report directly the values of the bias parameter b as deduced from our catalogue. The results show that b is a strongly increasing function of redshift in all cosmological models: from $z \simeq 0$ to $z \simeq 4$ the bias changes from $b \simeq 1$ to $b \simeq 5$ in the EdS model and from $b \simeq 0.5$ to $b \simeq 3$ in the OCDM and ΛCDM models. This qualitative behaviour is what is expected in the framework of the hierarchical models of galaxy formation, as confirmed by the curves of the effective bias computed by using equation (7), with $M_{\rm min}=10^{10}$, 10^{11} and $10^{12}\,h^{-1}\,{\rm M}_\odot$. The observed bias is well reproduced when a minimum mass of $\simeq 10^{11}\,h^{-1}\,{\rm M}_\odot$ is adopted for SCDM, in agreement with the discussion of the results about the correlation amplitude A_ω. On the contrary, the study of the bias parameter for the other two EdS models (TCDM and τCDM) seems to suggest a smaller value of $M_{\rm min} \simeq 10^{10}\,h^{-1}\,{\rm M}_\odot$. The discrepancy is due to the fact that the computation of the correlation amplitudes has been made by adopting a fixed slope of $\delta=0.8$, which is not a good estimate of the best fit value for these two models. For the OCDM and ΛCDM models, a minimum mass of $M_{\rm min} \simeq 10^{11}\,h^{-1}\,{\rm M}_\odot$ gives an effective bias in agreement with the observations when $1.5 \lesssim z \lesssim 3$, while a smaller (larger) minimum mass is required at lower (higher) redshift.

We can analyse the properties of the present-day descendants of our galaxies at high z, assuming that the large majority of them contain only one of our high-redshift galaxies (see e.g. Baugh et al. 1998). Following Mo et al. (1999), we can obtain the present bias factor of these descendants by evolving $b(z)$ backwards in redshift from the formation redshift z to $z=0$, according to the 'galaxy-conserving' model (Matarrese et al. 1997; Moscardini et al. 1998); this gives

$$b(M,0) = 1 + D_+(z)[b(M,z) - 1], \tag{12}$$

where, for $b(M,z)$, we can use the effective bias obtained for our galaxies by dividing the observed galaxy rms fluctuation on $8\,h^{-1}$ Mpc by that of the mass, which depends on the background cosmology. For galaxies at $z \simeq 3$ we find $b(M,0) \simeq 1.4, 1.3, 1.3$ for the EdS, OCDM and ΛCDM models, respectively. The values of $b(M,0)$ that we obtained can be directly compared with those for normal bright galaxies, which have $b_0 \approx 1/\sigma_8$, i.e. approximately 1.9 in the EdS universe and 1.1 in the OCDM and ΛCDM models. Consequently, the descendants of our galaxies at $z \simeq 3$ appear in the EdS universe to be less clustered than present-day bright galaxies and can be found among field galaxies. On the contrary, the values resulting for the OCDM and ΛCDM models seem to imply that the descendants are clustered at least as much as present-day bright galaxies, so they could be found among the brightest galaxies or inside clusters. This is in agreement with the findings of Mo et al. (1999) for the LBGs (see also Mo & Fukugita 1996; Governato et al. 1998; Baugh et al. 1999). If we repeat the analysis by using our galaxies at redshift $z \simeq 4$, we find that $b(M,0) \simeq 1.8, 1.7, 1.6$ for the EdS, OCDM and ΛCDM models, respectively. The ratio between the correlation amplitudes of the descendants and the normal bright galaxies is $\simeq 0.9, 2.3, 2.2$. This result confirms that for the EdS models they have clustering properties comparable to 'normal' galaxies, while for non-EdS models the descendants seem to be very bright and massive galaxies.

6 DISCUSSION AND CONCLUSIONS

In this paper we have measured over the redshift range $0 \leqslant z \leqslant 4.5$ the clustering properties of a faint galaxy sample in the *HDF North* (Fernández-Soto et al. 1999) by using photometric redshift estimates. This technique makes it possible both to isolate galaxies in relatively narrow redshift intervals, reducing the dilution of the clustering signal (in comparison with magnitude limited samples; Villumsen, Freudling & da Costa 1997) and to measure the clustering evolution over a very large redshift interval for galaxies fainter than the spectroscopic limits. The comparison with spectroscopic measurements shows that, for galaxies brighter than $I_{AB} \leqslant 26$, our accuracy is close to $\sigma_z \sim 0.1$ for $z \leqslant 1.5$ and $\sigma_z \sim 0.2$ for $z \geqslant 1.5$. We have checked the reliability of our photometric redshift in the critical interval $1.2 \leqslant z \leqslant 2$ by replacing the *J*, *H* and *Ks* photometry of Dickinson et al. (in preparation) with the *F*110*W*, *F*160*W* measurements in the *HDF*-N sub-area observed with NICMOS (Thompson et al. 1999). The new photometry is in general consistent with the IR photometry of FLY99 and our photometric redshifts are not significantly changed. In order to infer the confidence level for the galaxies beyond the spectroscopic limits ($26 \leqslant I_{AB} \leqslant 28.5$), we have compared our results first with those obtained by other photometric codes and second with Monte Carlo simulations. The first comparison shows that the resulting dispersion is $\sigma_z \simeq 0.20$ at $z \leqslant 1.5$ and increases at higher redshift ($\sigma_z \simeq 0.30$), with a possible systematic shift ($\langle z_{GIS} - z_{CE}\rangle \simeq -0.15$ and $\langle z_{GIS} - z_{FLY99}\rangle \simeq +0.3$). The second comparison with Monte Carlo simulations (made to determine the effects of photometric errors in the redshift estimates) shows that the rms dispersion obtained in this way is compatible with the previous estimates done by comparing the different codes: for galaxies with $I_{AB} \leqslant 28.5$ we found $\sigma_z \simeq 0.2$–0.3 with a maximum $\sigma_z = 0.35$ for the redshift range $1.5 \leqslant z \leqslant 2.5$. The contamination fraction of simulated galaxies incorrectly put in a bin different from the original one owing to photometric errors is close to ~ 20 per cent. The dominant source of contamination in a given redshift bin is due to the rms dispersion in the redshift estimates with the exception of the bin $0 \leqslant z \leqslant 0.5$, where the contamination is due to the high-redshift galaxies ($z \geqslant 1$) improperly put at low z. Owing to the contamination effect at any redshift, we note that our clustering measurements should be considered as a lower limit. Assuming that the contaminating population is uncorrelated, we have applied a correction $(1-f)^2$ to our original measurements, where f is the contaminating fraction. This correction should be regarded as an upper limit.

As a consequence of the redshift uncertainties, we have chosen to compute the angular correlation function $\omega(\theta)$ in large bins with $\Delta z = 0.5$ at $z \leqslant 1.5$ and $\Delta z = 1.0$ at $z \geqslant 1.5$. The resulting $\omega(\theta)$ has been fitted with a standard power-law relation with fixed slope, $\delta = 0.8$. This latter value can be questioned because of the present lack of knowledge about the redshift evolution of the slope and its dependence on the different classes of object. In order to avoid systematic biases in the analysis of the results, the theoretical predictions have been treated with the same basic assumptions. The behaviour of the amplitude of the angular correlation function at 10 arcsec (A_ω) shows a decrease up to $z \simeq 1$–1.5, followed by a slow increase. The comoving correlation length r_0 computed from the clustering amplitudes shows a similar trend, but its value depends on the cosmological parameters. Finally, we have compared our σ_8^{gal} to that of the mass predicted for three cosmologies to estimate the bias. For all cases, we found that the bias is an increasing function of redshift with $b(z \simeq 0) \simeq 1$ and $b(z \simeq 4) \simeq 5$ (for an EdS universe), and $b(z \simeq 0) \simeq 0.5$ and $b(z \simeq 4) \simeq 3$ (for an open and a Λ universe). This result confirms and extends in redshift the results obtained by A98 and G98 for a Lyman-Break galaxy catalogue at $z \simeq 3$, suggesting that these high-redshift galaxies are located preferentially in the rarer and denser peaks of the underlying matter density field.

We have compared our results with the theoretical predictions of a set of different cosmological models belonging to the class of the CDM scenario. With the exception of the SCDM model, all the other models are consistent with both the local observations and the COBE measurements. We model the bias by assuming that the galaxies are associated in a one-to-one correspondence with their hosting dark-matter haloes defined by a minimum mass (M_{min}). Moreover, we assume that the haloes continuously merge into more massive ones. The values of $M_{min}(z)$ used in these computations refer either to a fixed mass or to the median mass derived by our GISSEL model or to the value required to reproduce the observed density of galaxies at any redshift.

The comparison shows that all galaxy formation models presented in this work can reproduce the redshift evolution of the observed bias and correlation strength. The halo masses required to match the observations depend on the adopted background cosmology. For the EdS universe, the SCDM model reproduces the observed measurements if a typical minimum mass of $10^{11}\,h^{-1}\,M_\odot$ is used, while the τCDM and TCDM models require a lower typical mass of 10^{10}–$10^{10.5}\,h^{-1}\,M_{\odot}$. For the OCDM and ΛCDM models, the mass is a function of redshift, with $M_{\min} \leqslant 10^{10}\,h^{-1}\,M_{\odot}$ at $z \leqslant 1.5$, $10^{11.5}\,h^{-1}\,M_{\odot}$ in the range $1.5 \leqslant z \leqslant 3$ and $10^{12}\,h^{-1}\,M_{\odot}$ at $z \simeq 4$. The higher masses required at high z to reproduce the clustering strength for these models are a consequence of the smaller bias they predict at high redshift compared to the EdS models.

We notice that, at very low z, both the OCDM and ΛCDM models overpredict the clustering and consequently the bias. Two effects may be responsible for this failure. First, the one-to-one correspondence between haloes and galaxies may be an inappropriate description at low z, where a more complex picture might be required. Second, we have assumed that merging continues to be effective at low z when, on the contrary, the fast expansion of the universe acts against this process, particularly for these models.

As a consequence of the bias dependence on the redshift and on the selection criteria of the samples, the behaviour of galaxy clustering cannot provide a straightforward prediction of the behaviour of underlying matter clustering. For this reason, the parametric form $\xi(r,z) = \xi(r, z=0)(1+z)^{-(3+\epsilon-\gamma)}$, where ϵ models the gravitational evolution of the structures, cannot correctly describe the observations for any value of ϵ.

Another prediction of the hierarchical models is the dependence of the clustering strength on the limiting magnitude of the samples. At $z \simeq 3$, we have compared our clustering measurements with the previous results obtained for the LBGs in A98 and G98. The three samples correspond to different galaxy densities (our density in the *HDF* is approximately 65 times higher than the LBGs of A98). Clustering strength shows a decrease with density. This result is in excellent agreement (both qualitatively and quantitatively) with the clustering strength predicted by the hierarchical models as a function of halo density. More abundant haloes are less clustered than less abundant ones (see also Mo et al. 1999). Moreover, this result, which is independent of the adopted cosmology, supports our assumption of a one-to-one correspondence between haloes and galaxies at high redshift (see also Baugh et al. 1999), because otherwise we would expect a higher small-scale clustering at the observed density. As also noticed in A98 for LBGs, such a result seems to be incompatible with a model which assumes a stochastic star formation process, which would predict that observable galaxies have a wider range of masses. In fact, in this case the correlation strength should be lower than the observed one because of the contribution by the most abundant haloes (which are less clustered). Moreover, it seems possible to exclude that a very large fraction (more than 50 per cent) of massive galaxies are lost by observations due to dust obscuration, because the correlation strength would be incompatible with the observed density. Consequently, one of the main results of A98, namely the existence of a strong relation between the halo mass and the absolute UV luminosity due to the fact that more massive haloes host the brighter galaxies, seems to be supported by the present work also for galaxies ten times fainter.

We have estimated the clustering properties at the present epoch of the descendants of our high-redshift galaxies. To do so, we have assumed that only one galaxy is hosted by the descendants. The resulting local bias for the descendants of the galaxies at $z \simeq 3$ is $b(z=0) \simeq 1.4, 1.3, 1.3$ for the EdS, OCDM and ΛCDM models, respectively. Considering the galaxies at $z \simeq 4$, we obtain $b(z=0) = 1.8, 1.7, 1.6$, respectively. These values seem to indicate that in the case of the the EdS universe they are field or normal bright galaxies while for the OCDM and ΛCDM models the descendants can be found among the brightest and most massive galaxies (preferentially inside clusters). As already noted, at $z \simeq 3$, the clustering strength and the observed density of galaxies are in good agreement with the theoretical predictions for any fashionable cosmological model. At $z \simeq 4$, the present analysis seems to be more discriminant. Although our estimation should be regarded as tentative and needs future confirmation, we find a remarkably high correlation strength. For some models the observed density of galaxies starts to be inconsistent with the required theoretical halo density. The relation between clustering properties and number density of very high redshift galaxies therefore provides an interesting way to investigate the cosmological parameters. The difference in the predicted masses ($\simeq$ 15 to 30 at $z \simeq 3$ and 4) between EdS and non-EdS universe models is also in principle testable in terms of measured velocity dispersions. The present results have been obtained in a relatively small field for which the effects of cosmic variance may be important (see Steidel 1998 for a discussion). Nevertheless they show a possibility of challenging cosmological parameters which becomes particularly exciting in view of the rapidly growing wealth of multi-wavelength photometric databases in various deep fields and availability of 10m-class telescopes for spectroscopic follow-up in the optical and near infrared.

ACKNOWLEDGMENTS

We are grateful to H. Aussel, C. Benoist, M. Bolzonella, A. Bressan, S. Charlot, S. Colombi, S. d'Odorico, H. Mo, R. Sheth and G. Tormen for useful general discussions. We acknowledge K. Lanzetta, A. Fernández-Soto and A. Yahil for having made available the photometric optical and IR catalogues of the *HDF*-North. We also thank the anonymous referee for comments which allowed us to improve the presentation of this paper. Many thanks to P. Bristow for carefully reading the manuscript. This work was partially supported by Italian MURST, CNR and ASI and by the TMR programme Formation and Evolution of Galaxies set up by the European Community. S. Arnouts has been supported during this work by a Marie Curie Grant Fellowship.

REFERENCES

Abraham R. G., Tanvir N. R., Santiago B. X., Ellis R. S., Glazebrook K., van den Bergh S., 1996, MNRAS, 279, L47

Adelberger K. L., Steidel C. C., Giavalisco M., Dickinson M. E., Pettini M., Kellogg M., 1998, ApJ, 505, 18(A98)

Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15

Baugh C. M., Cole S., Frenk C. S., Lacey C. G., 1998, ApJ, 498, 504

Baugh C. M., Benson A. J., Cole S., Frenk C. S., Lacey C. G., 1999, MNRAS, 305, L21

Benítez N., 1998, preprint, astro-ph/9811189

Benoist C., Maurogordato S., da Costa L. N., Cappi A., Schaeffer R., 1996, ApJ, 472, 452

Bertin E., Arnouts S., 1996, A&AS, 117, 393

Bruzual G., Charlot S., 1999, in preparation

Bunn E. F., White M., 1997, ApJ, 480, 6

Calzetti D., 1997, in Waller W. H., Fanelli M. N., Hollis J. E., Danks A. C., eds, The Ultraviolet Universe at Low and High Redshift: Probing the Progress of Galaxy Evolution, AIP Conf. Proc. 408. AIP, Woodbury, p. 403

Calzetti D., Kinney A. L., Storchi-Bregmann T., 1994, ApJ, 429, 582

Carlberg R. G., Cowie L. L., Songaila A., Hu E. M., 1997, ApJ, 484, 538

Coleman G. D., Wu C. C., Weedman D. W., 1980, ApJS, 43, 393

Coles P., Lucchin F., Matarrese S., Moscardini L., 1998, MNRAS, 300, 183
Colún P., Carlberg R. G., Couchman H. M. P., 1997, ApJ, 490, 1
Colombi S., Szapudi I., Szalay A. S., 1998, MNRAS, 296, 253
Connolly A. J., Csabai I., Szalay A. S., Koo D. C., Kron R. G., 1995, AJ, 110, 2655
Connolly A. J., Szalay A. S., Brunner R. J., 1998, ApJ, 499, L125
Eke V. R., Cole S., Frenk C. S., 1996, MNRAS, 282, 263
Fasano G., Cristiani S., Arnouts S., Filippi M., 1998, AJ, 115, 1400
Fernández-Soto A., Lanzetta K. M., Yahil A., 1999, ApJ, 513, 34 (FLY99)
Franceschini A., Silva L., Fasano G., Granato G. L., Bressan A., Arnouts S., Danese L., 1998, ApJ, 506, 600
Gheller C., Pantano O., Moscardini L., 1998, MNRAS, 296, 85
Giallongo E., D'Odorico S., Fontana A., Cristiani S., Egami E., Hu E., McMahon R. G., 1998, AJ, 115, 2169
Giavalisco M., Steidel C. C., Adelberger K. L., Dickinson M. E., Pettini M., Kellogg M., 1998, ApJ, 503, 543(G98)
Governato F., Baugh C. M., Frenk C. S., Cole S., Lacey C. G., Quinn T., Stadel J., 1998, Nat, 392, 359
Gwyn S. D. J., Hartwick F. D. A., 1996, ApJ, 468, L77
Hamilton A. J. S., Kumar P., Lu E., Mathews A., 1991, ApJ, 374, L1
Henry J. P., Arnaud K. A., 1991, ApJ, 372, 410
Hogg D. W. et al., 1998, AJ, 115, 1418
Jing Y. P., 1998, ApJ, 503, L9
Jing Y. P., 1999, ApJ, 515, L45
Kaiser N., 1984, ApJ, 284, L9
Kauffmann G., Colberg J. M., Diaferio A., White S. D. M., 1999, MNRAS, 303, 188
Kinney A. L., Bohlin R. C., Calzetti D., Panagia N., Wyse R. F. G., 1993, ApJS, 86, 5
Landy S. D., Szalay A. S., 1993, ApJ, 412, 64
Lanzetta K. M., Yahil A., Fernández-Soto A., 1996, Nat, 381, 759
Le Fèvre O., Hudon D., Lilly S. J., Crampton D., Hammer F., Tresse L., 1996, ApJ, 461, 534
Lidsey J. E., Coles P., 1992, MNRAS, 258, L57
Loveday J., Maddox S. J., Efstathiou G., Peterson B. A., 1995, ApJ, 442, 457
Lucchin F., Matarrese S., 1985, Phys. Rev., D32, 1316
Lucchin F., Matarrese S., Mollerach S., 1992, ApJ, 401, L49
Madau P., 1995, ApJ, 441, 18
Madau P., Ferguson H. C., Dickinson M. E., Giavalisco M., Steidel C. C., Fruchter A., 1996, MNRAS, 283, 1388
Magliocchetti M., Maddox S. J., 1999, MNRAS, astro-ph/9811320, in press
Matarrese S., Coles P., Lucchin F., Moscardini L., 1997, MNRAS, 286, 115
Miralles J. M., Pelló R., 1998, astro-ph/9801062, preprint
Mo H. J., Fukugita M., 1996, ApJ, 467, L9
Mo H. J., White S. D. M., 1996, MNRAS, 282, 347
Mo H. J., Mao S., White S. D. M., 1999, MNRAS, 304, 175
Moessner R., Jain B., Villumsen J. V., 1998, MNRAS, 294, 291
Moscardini L., Coles P., Lucchin F., Matarrese S., 1998, MNRAS, 299, 95
Park C., Vogeley M. S., Geller M. J., Huchra J. P., 1994, ApJ, 431, 569
Peacock J. A., Dodds S. J., 1996, MNRAS, 280, L19
Peebles P. J. E., 1980, The large–scale structure of the Universe, Princeton Univ. Press, Princeton
Porciani C., Catelan P., Lacey C., 1999, ApJ, 513, L99
Postman M., Lauer T. R., Szapudi I., Oegerle W., 1998, ApJ, 506, 33
Press W. H., Schechter P., 1974, ApJ, 187, 425
Roukema B. F., Valls-Gabaud D., Mobasher B., Bajtlik S., 1999, MNRAS, 305, 151
Santiago B. X., da Costa L. N., 1990, ApJ, 362, 386
Sawicki J. J., Lin H., Yee H. K. C., 1997, AJ, 113, 1 (SLY97)
Sheth R. K., Tormen G., 1999, MNRAS, 308, 119
Steidel C., 1998, in Banday A. J., Sheth R. K., da Costa L. N., eds, Proc. MPA/ESO Conf. on Evolution of Large Scale Structure: from Recombination to Garching, astro-ph/9811400, in press
Steidel C., Giavalisco M., Pettini M., Dickinson M. E., Adelberger K., 1996, ApJ, 462, L17
Sugiyama N., 1995, ApJS, 100, 281
Thompson R. I., Storrie-Lombardi L. J., Weymann R. J., Rieke M. J., Schneider G., Stobie E., Lytle D., 1999, AJ, 117, 17
Tucker D. L. et al., 1997, MNRAS, 285, 5
van den Bergh S., Abraham R. G., Ellis R. S., Tanvir N. R., Santiago B. X., Glazebrook K., 1996, AJ, 112, 359
Villumsen J. V., 1996, MNRAS, 281, 369
Villumsen J. V., Freudling W., da Costa L. N., 1997, ApJ, 481, 578
Wang Y., Bahcall N., Turner E. L., 1998, AJ, 116, 2081
White M., Gelmini G., Silk J., 1995, Phys. Rev., D51, 2669
White M., Viana P. T. P., Liddle A. R., Scott D., 1996, MNRAS, 283, 107
Williams R. E. et al., 1996, AJ, 112, 1335

This paper has been typeset from a TEX/LATEX file prepared by the author.

4.3.5 Article 5 : The galaxy-halo connection from a joint lensing, clustering and abundance analysis

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **449,** 1352–1379 (2015) doi:10.1093/mnras/stv276

The galaxy–halo connection from a joint lensing, clustering and abundance analysis in the CFHTLenS/VIPERS field

J. Coupon,[1★] S. Arnouts,[2] L. van Waerbeke,[3] T. Moutard,[2] O. Ilbert,[2] E. van Uitert,[4] T. Erben,[4] B. Garilli,[5] L. Guzzo,[6,7] C. Heymans,[8] H. Hildebrandt,[4] H. Hoekstra,[9] M. Kilbinger,[10] T. Kitching,[11] Y. Mellier,[10,12,13] L. Miller,[14] M. Scodeggio,[5] C. Bonnett,[15] E. Branchini,[16,17,18] I. Davidzon,[19,20] G. De Lucia,[21] A. Fritz,[5] L. Fu,[22] P. Hudelot,[12] M. J. Hudson,[23,24] K. Kuijken,[9] A. Leauthaud,[25] O. Le Fèvre,[2] H. J. McCracken,[12,13] L. Moscardini,[19,20,26] B. T. P. Rowe,[27] T. Schrabback,[4] E. Semboloni[4] and M. Velander[14]

Affiliations are listed at the end of the paper

Accepted 2015 February 8. Received 2015 January 15; in original form 2014 October 23

ABSTRACT

We present new constraints on the relationship between galaxies and their host dark matter haloes, measured from the location of the peak of the stellar-to-halo mass ratio (SHMR), up to the most massive galaxy clusters at redshift $z \sim 0.8$ and over a volume of nearly 0.1 Gpc^3. We use a unique combination of deep observations in the CFHTLenS/VIPERS field from the near-UV to the near-IR, supplemented by ∼60 000 secure spectroscopic redshifts, analysing galaxy clustering, galaxy–galaxy lensing and the stellar mass function. We interpret our measurements within the halo occupation distribution (HOD) framework, separating the contributions from central and satellite galaxies. We find that the SHMR for the central galaxies peaks at $M_{\mathrm{h,peak}} = 1.9^{+0.2}_{-0.1} \times 10^{12}\,\mathrm{M}_{\odot}$ with an amplitude of 0.025, which decreases to ∼0.001 for massive haloes ($M_{\mathrm{h}} > 10^{14}\,\mathrm{M}_{\odot}$). Compared to central galaxies only, the total SHMR (including satellites) is boosted by a factor of 10 in the high-mass regime (cluster-size haloes), a result consistent with cluster analyses from the literature based on fully independent methods. After properly accounting for differences in modelling, we have compared our results with a large number of results from the literature up to $z = 1$: we find good general agreement, independently of the method used, within the typical stellar-mass systematic errors at low to intermediate mass ($M_{\star} < 10^{11}\,\mathrm{M}_{\odot}$) and the statistical errors above. We have also compared our SHMR results to semi-analytic simulations and found that the SHMR is tilted compared to our measurements in such a way that they over- (under-) predict star formation efficiency in central (satellite) galaxies.

Key words: gravitational lensing: weak – galaxies: clusters: general – cosmology: observations – dark matter.

1 INTRODUCTION

The last few years have seen an increasing interest in statistical methods linking observed galaxy properties to their dark matter haloes, owing to the availability of numerous large scale multi-wavelength surveys. Those techniques are based on the assumption that the spatial distribution of dark matter is predictable and one is able to match its statistical properties with those of the galaxies. The halo model (see Cooray & Sheth 2002) is a quantitative representation of the distribution of dark matter, characterized by three main ingredients: the halo mass function describing the number density of haloes per mass, the halo bias tracing the clustering amplitude and the halo density profile.

Galaxies are born and evolve in individual haloes where the baryonic gas condensates, cools and forms stars. Galaxies are gravitationally bound to dark matter and share a common fate with their host, e.g. during mergers. Although we understand qualitatively individual physical processes likely to be involved in galaxy evolution, a number of key answers are missing.

★E-mail: jean.coupon@unige.ch

Observations show that a fraction of galaxies experienced star formation quenching and have become passive, shaping the galaxy population into a bimodal blue/red distribution (Faber et al. 2007; Ilbert et al. 2013). The number of these passive galaxies is higher today than in the past and increases with increasing halo mass. Might feedback processes in massive haloes be responsible for this, or is there a universal critical stellar mass above which star formation ceases, independently of the halo mass? Studying the connection between galaxies and their host haloes is crucial to answer these questions.

Another enigmatic question is the low stellar mass fraction in low-mass haloes, seen in early studies connecting galaxies to their host haloes (Yang, Mo & van den Bosch 2003; Vale & Ostriker 2006; Zheng, Coil & Zehavi 2007). In fact, when measuring the stellar-to-halo mass ratio (SHMR) as a function of time, we observe that stellar mass is building up asymmetrically, first in massive haloes, later on in low-mass haloes (Conroy, Wechsler & Kravtsov 2006; Behroozi et al. 2013b). This asymmetry in the SHMR is one corollary of the so-called galaxy *downsizing* effect (Cowie et al. 1996). In low-mass haloes, stellar winds and supernovae may slow down star formation until the potential well grows deep enough to retain the gas and increase the star formation rate (SFR). Again, it becomes necessary to relate galaxy properties to their host halo mass.

A number of studies have related galaxy properties to dark matter haloes using the *abundance matching* technique (Marinoni & Hudson 2002; Conroy et al. 2006; Behroozi, Conroy & Wechsler 2010; Guo et al. 2010; Moster et al. 2010), which employs the stellar mass (or luminosity) function and the halo mass function to match halo–galaxy properties based on their cumulative abundances. The *conditional luminosity function* technique proposed by Yang et al. (2003) includes a parametrized $M_\star - M_\mathrm{h}$ relationship whose parameters are fitted to the luminosity function. Both this formalism and recent abundance matching studies feature a scatter in $M_\star$ at fixed M_h, which is an important ingredient to account for, given the steep relation between the two quantities at high mass.

More recently, models adopting a similar approach to abundance matching consist of directly populating dark matter haloes from N-body simulations, to reproduce the observed stellar mass functions as a function of redshift, using a parametrized SFR model to account for redshift evolution (Moster, Naab & White 2012; Behroozi et al. 2013b).

Except in some rare cases where central or satellite galaxies can be individually identified (e.g. George et al. 2011; More et al. 2011), in studies based on luminosity or stellar mass distributions, the satellite galaxies' properties cannot be disentangled from those of the central galaxies. To remedy the problem, abundance matching techniques either assume an ad hoc fraction of satellites or use a subhalo mass function estimated from numerical simulations. Unfortunately, as subhaloes may be stripped and disappear after being accreted on to larger haloes, the subhalo mass function at the time considered might not correspond to the distribution of satellites, and one must consider the mass of subhaloes at the time of accretion, further extrapolated to the time considered. Obviously these complications limit the amount of information one can extract about galaxy satellites.

Galaxy clustering, on the other hand, allows separation of the contributions from central and satellite galaxies due to the different typical clustering scales. To model the clustering signal of a given galaxy population, the halo occupation distribution (HOD) formalism assumes that the galaxy number per halo is solely a function of halo mass and that the galaxy satellite distribution is correlated to that of the dark matter (Berlind & Weinberg 2002; Kravtsov et al. 2004).

One achievement of HOD modelling was to demonstrate from simulations (Berlind et al. 2003; Moster et al. 2010) that only a handful of parameters was necessary to fully describe galaxy–halo occupation. This parametric HOD was fitted to a number of observations over a large range of redshifts and galaxy properties. Among the more remarkable results are the local Universe galaxy clustering and abundance matching studies performed on the Sloan Digital Sky Survey (SDSS; see e.g. Zehavi et al. 2011) and at higher redshifts (Foucaud et al. 2010; Wake et al. 2011; Coupon et al. 2012; de la Torre et al. 2013; Martinez-Manso et al. 2015).

However, some underlying assumptions on the distribution of dark matter haloes implied in the HOD formalism are observationally challenging to confirm and one has to rely on N-body simulations. Fortunately, additional techniques may be used to relate galaxy properties to halo masses, among which gravitational lensing is one of the most powerful probes: by evaluating the distortion and magnification of background sources, one is able to perform a direct estimation of the dark matter halo profile (for a review, see Bartelmann & Schneider 2001). The low signal-to-noise ratio associated with individual galaxies, however, forces us to 'stack' them (e.g. binned together within narrow stellar mass ranges), using a technique known as *galaxy–galaxy lensing* (Brainerd, Blandford & Smail 1996; Hudson et al. 1998; Hoekstra, Yee & Gladders 2004; Mandelbaum et al. 2005a; Yoo et al. 2006; van Uitert et al. 2011; Cacciato, van Uitert & Hoekstra 2014; Velander et al. 2014; Hudson et al. 2015).

Clearly, each of the above methods brings a different piece of information and combining all observables together is particularly interesting, although doing so properly is challenging. In a recent study using COSMOS data, Leauthaud et al. (2012) have successfully combined galaxy clustering, galaxy–galaxy lensing and the stellar mass function (see also Cacciato et al. 2009; Mandelbaum et al. 2013; Miyatake et al. 2013; More et al. 2014), fitted jointly and interpreted within the HOD framework: the authors have used a global central galaxy $M_\star - M_\mathrm{h}$ relationship (as opposed to measuring the mean M_h per bin of stellar mass) and extended it in a consistent way to satellite galaxies.

In this paper, we apply this advanced formalism using a new data set covering a uniquely large area of $\sim$25 deg^2 with accurate photometric redshifts in the redshift range $0.5 < z < 1$ and stellar masses $> 10^{10}\,\mathrm{M}_\odot$. Our galaxy properties' measurements are calibrated and tested with 70 000 spectroscopic redshifts from the VIPERS survey and a number of publicly available data sets. Our data span a wide wavelength range of ultraviolet (UV) deep data from *GALEX*, optical data from the Canada–France–Hawaii Telescope (CFHT) Legacy Survey and K_s-band observations with the CFHT WIRCam instrument. This large statistical sample allows us to measure with high precision the stellar mass function, the galaxy clustering, and we use the CFHTLenS shear catalogue to measure galaxy–galaxy lensing signals. The galaxy clustering is measured on the projected sky for the photometric sample and in real space for the spectroscopic sample.

This paper is organized as follows: in Section 2, we describe the observations, the photometric redshift and stellar mass estimates. In Section 3, we present the measurements of the stellar mass function, the galaxy clustering (both from the photometric and spectroscopic samples) and galaxy–galaxy lensing signals. In Section 4, we describe the HOD model, and the Markov chain Monte Carlo (MCMC) model fitting results are given in Section 5. In Section 6, we discuss our results and conclude. Throughout the paper, we adopt the following cosmology: $H_0 = 72$ km s^{-1} Mpc^{-1} and $\Omega_\mathrm{m} = 0.258$, $\Omega_\Lambda = 0.742$ (Hinshaw et al. 2009) unless otherwise stated. To compute stellar masses, we adopt the initial mass function (IMF) of Chabrier (2003) truncated at 0.1 and 100 M$_\odot$, and the stellar population synthesis (SPS) models of Bruzual & Charlot (2003). All magnitudes are given in the AB system. The dark matter halo masses are denoted as M_h and defined within the virial radius enclosing a mean overdensity Δ_vir compared to the mean density background, taking the formula from Weinberg & Kamionkowski (2002). At $z = 0.8$, $\Delta_\mathrm{vir} = 215$. All masses are expressed in unit of M$_\odot$. Measured quantities are denoted as $\widetilde{X}$ and theoretical quantities as X. We call *cosmic variance* the statistical uncertainties caused by the density fluctuations of dark matter and we define the *sample variance* as the sum of the cosmic variance and Poisson noise variance.

2 DATA

In this work, we combine several data sets to build a volume-limited sample of galaxies more massive than $M_\star = 10^{10}$ M$_\odot$ in the redshift range $0.5 < z < 1$. Our galaxy selection is based on NIR ($K_s < 22$) observations, collected in the two fields of the VIMOS Public Extragalactic Redshift Survey ('VIPERS-*W*1' and 'VIPERS-*W*4'), overlapping the (Canada–France–Hawaii Telescope Legacy Survey) CFHTLS-Wide imaging survey, and covering a total unmasked area of 23.1 deg^2. We refer to Arnouts et al. (in preparation) for a complete description of the multiwavelength UV and NIR observations, reduction and photometry.

Our background galaxy selection used for the measurement of the lensing signal is based on the CFHTLS-Wide *i*-band selection in the area that overlaps with the NIR observations.

2.1 The CFHTLS-Wide survey

The CFHTLS[1] is a photometric survey performed with MegaCam (Boulade et al. 2003) on the CFHT telescope in five optical bands $u^\star$, g, r, i, z ($i < 24.5$–25, 5σ detection in 2 arcsec apertures) and covering four independent patches in the sky over a total area of 154 deg^2. In this analysis, we use the photometric and shear catalogues produced by the CFHTLenS[2] team (Heymans et al. 2012). The CFHTLenS photometry is performed with SExtractor (Bertin & Arnouts 1996) on the PSF-homogenized images (Hildebrandt et al. 2012; Erben et al. 2013). Magnitudes are based on the `MAG_ISO` estimator where the isophotal apertures are derived from the *i*-band detection image. This approach optimizes the colour measurements and leads to an improvement in the photometric redshift accuracy (Hildebrandt et al. 2012). To estimate the total magnitude of each source, a global shift is applied to the `MAG_ISO` magnitude in all the bands based on the difference between `MAG_ISO` and `MAG_AUTO` magnitudes, as measured in the *i*-band detection image (Hildebrandt et al. 2012).

As the magnitude errors are measured with SExtractor directly from the local background in the PSF-homogenized image, we need to correct for the noise correlation introduced by the convolution process. To do so, we multiply the CFHTLenS magnitude errors in all bands by the ratio of the *i*-band detection image errors to the *i*-band PSF-homogenized image errors. The correction factor ranges from 3 to 5, where the strongest correction occurs when the seeing difference between the i band and the worse-seeing image is the largest. As the *i*-band image is usually the best-seeing image, this procedure may slightly overestimate the correction in the other bands, however we neglect it here.

In addition, magnitude errors must be rescaled to account for image resampling. Two independent tests have been performed to accurately estimate the correction factor: we measured the dispersion of magnitudes between the *i*-band detection (un-convolved) magnitudes and the CFHTLS-Deep magnitudes, and between duplicated observations of the same object in the overlapping regions of adjacent tiles. We find that the errors must be rescaled by a factor of 2.5.

The footprints of the CFHTLS MegaCam tiles overlapping the VIPERS survey are shown as grey squares in Fig. 1.

2.2 The Near-IR observations

We have conducted a K_s-band follow-up of the VIPERS fields with the WIRCam instrument at CFHT (Puget et al. 2004) for a total allocation time of ∼120 h. The integration time per pixel was 1050 s and the average seeing of all the individual exposures was 0.6 ±0.09 arcsec. The data have been reduced by the Terapix team:[3] the images were stacked and resampled on the pixel grid of the CFHTLS-T0007 release (Hudelot et al. 2012). The images reach a depth of $K_s = 22$ at $\sim 3\sigma$ (Arnouts et al., in preparation). The photometry was performed with SExtractor in dual image mode with a $gri - \chi^2$ image (Szalay, Connolly & Szokoly 1999) as the detection image. To correct for the noise correlation introduced by image resampling, we multiply the errors by a factor 1.5, obtained from the dispersion between the WIRCam K_s-band magnitudes and the magnitudes measured in the deeper ($K < 24.5$) UKIDSS Ultra Deep Survey (UDS; Lawrence et al. 2007). We also used the UDS survey to confirm that our sample completeness based on $gri - \chi^2$ detections reaches 80 per cent at $K_s = 22$. Using the WIRCAM/CFHTLS-Deep data with an *i*-band cut simulating the CFHTLS-Wide data depth, we have checked that this incompleteness is caused by red galaxies above $z = 1$ and does not affect our sample selected in the range $0.5 < z < 1$. The K_s `MAG_AUTO` estimates are then simply matched to their optical counterparts based on position.

In addition to this data set, we also use the CFHTLS-D1 WIRDS data (Bielby et al. 2012), a deep patch of 0.49 deg^2 observed with WIRCam J, H and K_s bands and centred on $02^\mathrm{h}26^\mathrm{m}59^\mathrm{s}$, $-04°30'00''$. All three bands reach 50 per cent completeness at AB magnitude 24.5.

The WIRCam observations are shown in Fig. 1 as the red regions. After rejecting areas with poor WIRCam photometry and those with CFHTLenS mask flag larger than 2, the corresponding effective area used in this work spans over 23.1 deg^2, divided into 15 and 8.1 deg^2 in the VIPERS-*W*1 and VIPERS-*W*4 fields, respectively.

2.3 The UV-*GALEX* observations

When available, we make use of the UV deep imaging photometry from the *GALEX* satellite (Martin et al. 2005; Morrissey et al. 2005). We only consider the observations from the Deep Imaging Survey (DIS), which are shown in Fig. 1 as blue circles (Ø∼1°.1). All the *GALEX* pointings were observed with the NUV channel with exposure times of $T_\mathrm{exp} \geq 30$ ksec. FUV observations are available for 10 pointings in the central part of *W*1.

[1] http://www.cfht.hawaii.edu/Science/CFHTLS/
[2] http://cfhtlens.org/
[3] http://terapix.iap.fr/

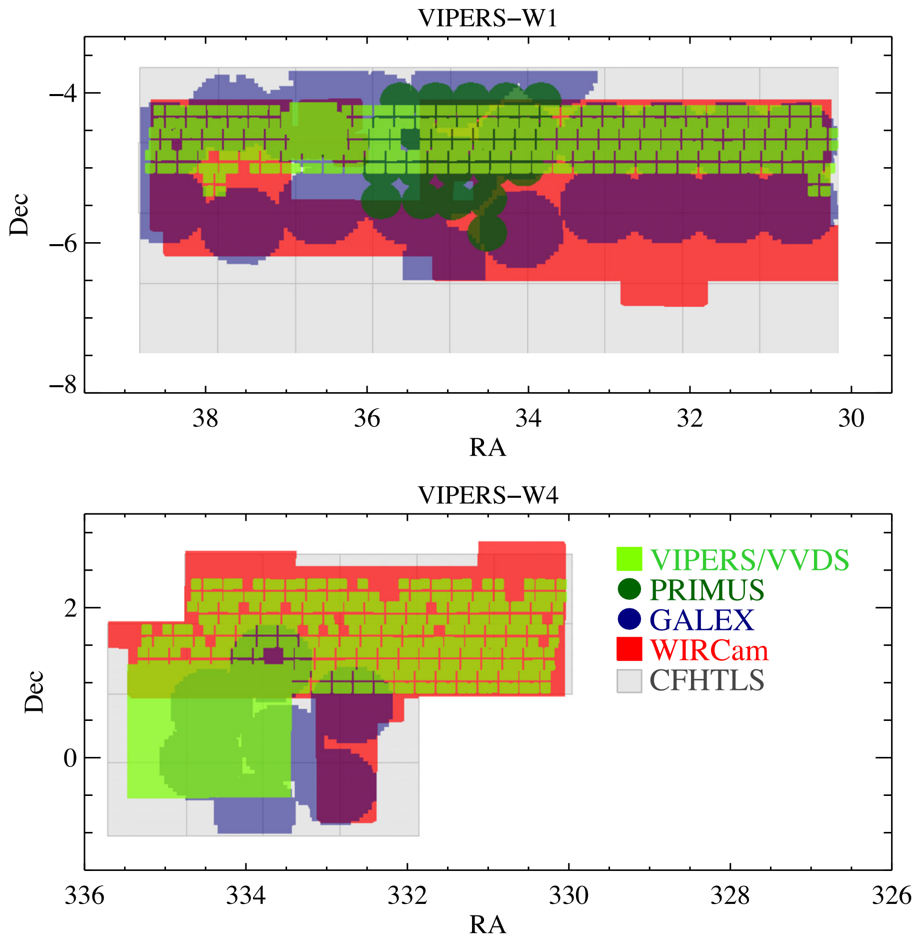


Figure 1. Footprints of the different data sets used in this work. Our selection is based on WIRCam data shown in red and covering approximately 25 deg^2 (23.1 deg^2 after masking). The CFHTLS MegaCam pointings are shown in grey, the *GALEX* DIS observations as large blue circles (in purple if overlapped with WIRCam), the spectroscopic surveys VIPERS/VVDS in light green and PRIMUS in dark green. The SDSS/BOSS coverage is almost complete. The data outside the WIRCam footprint are not used, and shown here only for reference.

Due to the large PSF (FWHM ∼ 5 arcsec), source confusion becomes a major issue in the deep survey. To extract the UV photometry we use a dedicated photometric code, EMPHOT (Conseil et al. 2011) which will be described in a separate paper (Vibert et al., in preparation). In brief, EMPHOT uses U band (here the CFHTLS u band) detected objects as a prior on position and flux. The uncertainties on the flux account for the residual in the [simulated−observed] image. The images reach a depth of $m_{\rm NUV} \sim 24.5$ at $\sim 5\sigma$. As for the WIRCAM data, the *GALEX* sources are matched to the optical counterparts based on position.

The NUV observations cover only part of the WIRCam area with ∼10.8 and 1.9 deg^2 in VIPERS-*W*1 and VIPERS-*W*4, respectively. The UV photometry slightly improves the precision of photometric redshifts and the stellar mass estimates in the *GALEX* area. However, by comparing our measurements inside and outside the *GALEX* area, we have checked that the addition of UV photometry does not make a significant change for the galaxies of interest in this study. Therefore, in the final sample, we mix galaxies inside the *GALEX* area with those outside.

2.4 Spectroscopic data

To optimize the calibration and the validation of our photometric redshifts, we make use of all the spectroscopic redshifts available in the WIRCam area.

The largest sample is based on the VIPERS spectroscopic survey (Garilli et al. 2014; Guzzo et al. 2014) and its first public data release PDR1.[4] VIPERS aims to measure redshift space distortions and explore massive galaxy properties in the range $0.5 < z < 1.2$. The survey is located in the *W*1 and *W*4 fields of the CFHTLS-Wide survey and will cover a total area of 24 deg^2 when completed, with a sampling rate of ∼40 per cent down to $i < 22.5$. In Fig. 1, we show the layout of the VIMOS pointings as the light-green squares. The PDR1 release includes redshifts for ∼54 204 objects. After keeping galaxy spectra within the WIRCam area (44 474) and with the highest confidence flags between 2.0 and 9.5 (95 per cent confidence, see Guzzo et al. 2014), we are left with 35 211 galaxies, which corresponds to a spectroscopic success rate of 80 per cent.

In addition to VIPERS, we also consider the following spectroscopic surveys:

(i) the VIMOS-VLT Deep Survey (VVDS) F02 and Ultra-Deep Survey (Le Fèvre et al. 2005, 2014) which consist of 11 353 galaxies down to $i < 24$ (Deep) and 1125 galaxies down to $i < 24.5$ (Ultra-Deep) over a total area of 0.75 deg^2 in the VIPERS-*W*1 field. We also use part of the VIMOS-VLT F22 Wide Survey with 12 995 galaxies over 4 deg^2 down to $i < 22.5$ (Garilli et al. 2008, shown as the large green square in the southern part of the VIPERS-*W*4 field

[4] http://vipers.inaf.it/rel-pdr1.html

Table 1. Magnitude zero-point offsets measured per CFHTLS MegaCam pointing in VIPERS-*W1* and VIPERS-*W4* (mean and standard deviation). *J*- and *H*-band zero-points were computed for the pointings overlapping WIRDS data.

Filter	VIPERS-*W1*	VIPERS-*W4*
FUV	0.18 ± 0.11	0.02 ± 0.16
NUV	0.11 ± 0.09	0.15 ± 0.10
u	0.10 ± 0.03	0.13 ± 0.03
g	− 0.02 ± 0.01	− 0.01 ± 0.01
r	0.02 ± 0.01	0.01 ± 0.01
i	− 0.01 ± 0.01	− 0.00 ± 0.01
z	− 0.02 ± 0.01	− 0.01 ± 0.01
J	0.08 ± 0.05	–
H	0.02 ± 0.05	–
K	0.02 ± 0.03	0.01 ± 0.05

in Fig. 1). In total, we use 5122 galaxies with secure flags 3 or 4 from the VVDS surveys within the WIRCam area;

(ii) the PRIMUS survey (Coil et al. 2011) which consists of low resolution spectra ($\lambda/\Delta\lambda \sim 40$) for galaxies down to $i \sim 23$ and overlapping our VIPERS-*W1* field. PRIMUS pointings are shown as the dark green circles in Fig. 1. We keep 21 365 galaxies with secure flags 3 or 4;

(iii) the SDSS-BOSS spectroscopic survey based on data release DR10 (Ahn et al. 2014) down to $i < 19.9$, overlapping both VIPERS-*W1* and VIPERS-*W4* fields, totalling 4675 galaxies with `zWarning=0` (99 per cent confidence redshift) within our WIRCam area.

In total, the spectroscopic sample built for this study comprises 62 220 unique galaxy spectroscopic redshifts with the highest confidence flag. We use the spectroscopic redshift value, when available, instead of the photometric redshift value. The galaxies with a spectroscopic redshift represent 6.5 per cent of the full sample, and 12 per cent after selection in the range $0.5 < z < 1$, where most of the galaxies are from the VIPERS sample.

2.5 Photometric redshifts

To compute the photometric redshifts, we use the template fitting code LEPHARE[5] (Arnouts et al. 1999; Ilbert et al. 2006). We adopt similar extinction laws and parameters as Ilbert et al. (2009) used in the COSMOS field (Scoville et al. 2007), and identical priors as in Coupon et al. (2009) based on the VVDS redshift distribution and maximum allowed g-band absolute magnitude. We note that the use of priors is essential for the $z > 1$, low signal-to-noise ratio (or no NIR flux), galaxies used as lensed (background) sources (see Section 3). A probability distribution function (PDF) in steps of 0.04 in redshift is computed for every galaxy.

We use the full spectroscopic sample to adjust the magnitude relative zero-points in all the passbands on a MegaCam pointing-to-pointing basis. For the pointings with no spectroscopic information, we apply a mean correction obtained from all the pointings with spectra. The mean zero-point offsets and standard deviations in all passbands are given in Table 1 for the two fields separately. We further add the zero-point scatter in quadrature to the magnitude errors in each band. We recall that these zero-point corrections may not represent absolute calibration offsets but rather *relative* (i.e. depending on colours) ones and tied to the adopted spectral energy distribution (SED) template set. We come back to the impact of this issue on stellar mass measurements in Section 3.5.

Our SED templates are based on the library used in Ilbert et al. (2009), however the fewer bands used in this study compared to COSMOS necessitate adapting the templates to reduce redshift-dependent biases. The initial templates are based on the SEDs from Polletta et al. (2007), complemented by a number of starburst SEDs from the Bruzual & Charlot (2003) SPS library. Using 35 211 spectroscopic redshifts from VIPERS, we adapt the templates with LEPHARE using the following procedure. First, a best-fitting template from the original set is found for each galaxy and normalized to unity, and the photometry is then corrected into the rest frame given the spectroscopic redshift value. The rest-frame photometry for all galaxies with identical best-fitting templates is combined and the adapted template is constructed from the sliding-window median values as a function of wavelength. The process is repeated iteratively. Given the high number of galaxies with spectroscopic redshifts, we found that only two iterations were necessary to reach convergence. Interestingly, although the improvement in the photometric redshift bias is significant, the new templates appear very similar 'by eye' compared to the original ones, which implies that small features in the SED templates may lead to large photometric errors, as also noted by Ilbert et al. (2006).

In Fig. 2, we show the accuracy of the photometric redshifts by comparing with the spectroscopic redshift sample from VIPERS ($i < 22.5$, left-hand panel) and VVDS Deep/Ultra-Deep ($22.5 < i < 24.5$, right-hand panel). We observe a dispersion[6] of $\sigma/(1+z) \sim 0.03$–0.04 and a fraction of catastrophic redshifts ($|\Delta z| \geq 0.15(1+z)$) of $\eta \sim 1$–4 per cent. The dispersion in both magnitude ranges is significantly better than previous results in the CFHTLS-Wide (Coupon et al. 2009), due to the choice of isophotal magnitudes and PSF homogenization (Hildebrandt et al. 2012) at faint magnitude, and the contribution of NIR data above $z \sim 1$. We note that the faint sample is compared to the VVDS redshifts where deep NIR data from WIRDS are available over a small part ($<1\ \mathrm{deg}^2$) of the field, and with a magnitude distribution biased towards bright galaxies compared to the photometric sample. Therefore, we foresee degraded photometric redshift performance elsewhere, mainly relevant for $z > 1$ galaxies. However, as shown in Appendix C, no systematic bias affecting our lensing measurements is introduced by the use of sources beyond $z = 1$.

2.6 Stellar mass estimates

To compute stellar masses, we adopt the same procedure as Arnouts et al. (2013) and described in detail in their Appendix . In brief, we use the photometric or spectroscopic (when available) redshift and perform a χ^2 minimization on a SED library based on the SPS code from Bruzual & Charlot (2003). The star formation history is either constant or described with an exponentially declining function, with e-folding time $0.01 \leq \tau \leq 15$. We use two metallicities ($Z_\odot$, $0.2\,Z_\odot$) and adopt the Chabrier (2003) IMF. As discussed in Arnouts et al. (2013), the use of various dust extinction laws is critical to derive robust SFR and stellar mass; and in this work, we adopt their choices for differing attenuation curves: a starburst (Calzetti

[5] http://www.cfht.hawaii.edu/arnouts/lephare.html

[6] Defined as the normalized median absolute deviation (Hoaglin, Mosteller & Tukey 1983): $1.48\times\mathrm{Median}(|z_s - z_p|/(1+z_s))$, and robust to outliers.

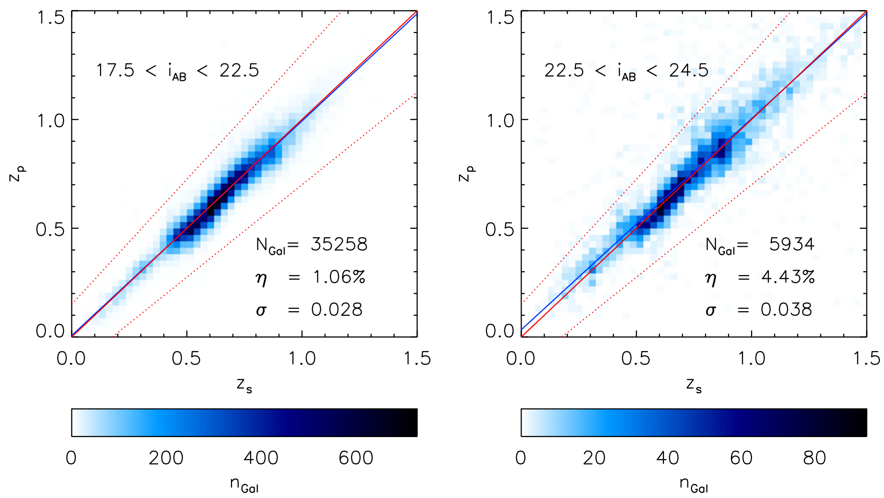


Figure 2. Photometric redshifts measured with *ugrizK* (left) or *ugrizJHK* (right) photometry versus VIPERS and VVDS spectroscopic redshifts. Left: $17.5 < i < 22.5$, where the sample is dominated by galaxies between $0.5 < z < 1.2$ due to the VIPERS selection. Right: $22.5 < i < 24.5$, from the VVDS Deep and Ultra-Deep surveys. The limits for the outliers are shown as red dotted lines.

Table 2. Sample mass definitions in $\log(M_\star / \mathrm{M}_\odot)$ and number of galaxies in each sample. The parent sample comprises a total of 352 585 galaxies.

Sample	Clustering–$w(\theta)$ Mass cut	Number	Clustering–$w_\mathrm{p}(r_\mathrm{p})$ Mass cut	Number	Lensing Mass cut	Number
1	10.00–10.40[a]	23 886	10.60–10.90[a]	2154	10.00–10.40[a]	23 886
2	10.40–10.60	36 560	10.90–11.20[b]	1964	10.40–10.65	45 032
3	10.60–10.80	31 900	11.20–12.00	816	10.65–10.80	23 427
4	10.80–11.00	24 451	–	–	10.80–10.95	19 293
5	11.00–11.20	13 538	–	–	10.95–11.15	16 317
6	11.20–12.00	6326	–	–	11.15–12.00	8654

Notes. [a]$0.5 < z < 0.7$.
[b]$0.5 < z < 0.8$.

et al. 2000), an SMC-like (Prevot et al. 1984) and an intermediate slope ($\lambda^{-0.9}$) law. We consider reddening excess in the range $0 \le E(B - V) \le 0.5$. When fixing the redshift, the typical 68 per cent stellar mass statistical uncertainty, as derived by marginalising the likelihood distribution, ranges from $\sigma(M_\star) \sim 0.05$ to 0.15 for galaxies with $K_s \le 22$ and $z < 1$. This stellar mass uncertainty is an underestimate, since we neglect photometric redshift uncertainties.[7]

In addition to statistical errors, in Section 3.5, we investigate the different sources of systematic effects in the stellar mass estimates, arising from our lack of knowledge of galaxy formation and evolution. The choice of differing dust treatments (and resulting dust attenuation laws) is one of them: Ilbert et al. (2010) have measured a shift of 0.14 dex, with a large scatter, between stellar masses estimated with the Charlot & Fall (2000) dust prescription and the Calzetti et al. (2000) attenuation law. The dust parametrization leads to systematics larger than the statistical errors in the stellar mass function. Even more critical is the choice of the SPS model and the IMF (see more detailed systematic errors analysis in Behroozi et al. 2010; Marchesini et al. 2009; Fritz et al. 2014), leading to systematic differences in stellar mass estimates up to 0.2 dex. One must keep these limitations in mind when comparing results from various authors using different methods, and we come back to these issues when presenting our results.

3 MEASUREMENTS

We aim to compute high signal-to-noise measurements of four distinct observables: the stellar mass function $\phi(M_\star)$, the projected galaxy clustering $w(\theta)$, the real-space galaxy clustering $w_\mathrm{p}(r_\mathrm{p})$ and the galaxy–galaxy lensing $\Delta\Sigma(r)$.

To do so, we select volume-limited samples in the redshift range $0.5 < z < 1$, where the high sampling rate of VIPERS and our NIR data guarantee both robust photometric redshift and stellar mass estimates. As for the stellar mass function, we adopt a lower mass limit of $M_\star = 10^{10}\,\mathrm{M}_\odot$ and employ the V_max estimator to correct for galaxy incompleteness near $z = 1$. The total volume probed in this study is 0.06 Gpc^3.

The stellar mass bins for the clustering and lensing measurements are defined to keep approximately a constant signal-to-noise ratio across the full mass range (which may lead to differing mass cuts depending on the observable), and guarantee complete galaxy samples (see Appendix A). We summarize our samples' properties in Table 2.

To measure each of the observables described below, we use the parallelized code SWOT, a fast tree-code for computing two-point correlations, histograms and galaxy–galaxy lensing signals from large data sets (Coupon et al. 2012). The stellar mass function is expressed in comoving units, whereas the clustering and

[7] We will see in Section 4.2 that our model accounts for such an extra source of uncertainty in stellar mass through a stellar-mass-dependent parametrization of the stellar-to-halo mass scatter.

galaxy–galaxy lensing signal are measured in physical units. We estimate statistical covariance matrices from a jackknife resampling of 64 subregions with equal area (0.35 deg^2 each), by omitting a subsample at a time and computing the properly normalized standard deviation (see more details in Coupon et al. 2012). This number was chosen to meet both requirements of using large enough subregions to capture the statistical variations at large scale, while keeping a sufficient number of subsamples to compute a robust covariance matrix. Nevertheless, we expect the projected galaxy clustering errors to be slightly underestimated on scales larger than the size of our subregions, ∼0.5 deg, and the noise in the covariance matrix to potentially bias the best-fitting χ^2 value.

A random sample with 1 million objects is constructed using our WIRCam observations layout and the union of the WIRCam and CFHTLenS photometric masks. For real-space clustering, measured from VIPERS spectroscopic redshifts, the random sample is constructed using the layout of the VIPERS PDR1 geometry (and photometric masks) plus a random redshift drawn in the range $0.5 < z < 1$ from a distribution following $\mathrm{d}V/\mathrm{d}z$, to match our volume-limited samples. The subregions for the measurements of statistical errors are constructed by SWOT based on the random catalogue: the field is divided into 64 areas with an equal number of random objects.

3.1 The stellar mass function

The stellar mass function $\widetilde{\phi}(M_\star) = \mathrm{d}n/\mathrm{d}\log M_\star$ is measured per unit of comoving volume in 10 equally spaced logarithmic mass bins of width 0.2 dex, centred on the mass mean weighted by the number of galaxies. To correct for the incompleteness in the low-mass galaxy sample ($10^{10} < M_\star/\mathrm{M}_\odot < 10^{10.4}$) occurring near $z = 1$ (see Appendix A), we up-weight low-redshift galaxies by a factor $1/V_{\max}$ defined as

$$V_{\max} = \Omega \int_{0.5}^{z_{\max}} \frac{\mathrm{d}V}{\mathrm{d}z}\mathrm{d}z\,, \tag{1}$$

where Ω is the solid angle of the survey, 23.1 deg^2, V the comoving volume per unit area and $z_{\max}$ the maximum redshift for a galaxy to be observed given a $K_s < 22$ magnitude cut, calculated with LEPHARE.

We have performed a number of tests to check our internal error estimates. In the top panel of Fig. 3, we show our stellar mass function error estimates (square root of the covariance matrix diagonal) as a function of stellar mass compared to the GETCV code estimate of Moster et al. (2011) at $z = 0.8$. The latter code computes the theoretical expectations of cosmic variance[8] assuming a prediction for dark matter clustering and galaxy biasing (Bardeen et al. 1986). We add to the GETCV cosmic variance the theoretical Poisson error and show the resulting (total) sample variance as the thick line in the bottom panel. Our jackknife estimate is represented as the black points, for which we find that the cosmic variance part (after subtracting Poisson noise) needs to be multiplied by a factor of 2 to agree with theoretical expectations (we then multiply the covariance matrix by a factor of 4). We have not found a definitive explanation for the underestimation of the errors from the jackknife resampling, however it is likely caused by the strong correlation between bins (a combined effect of stellar mass scatter and large-scale structure correlations).

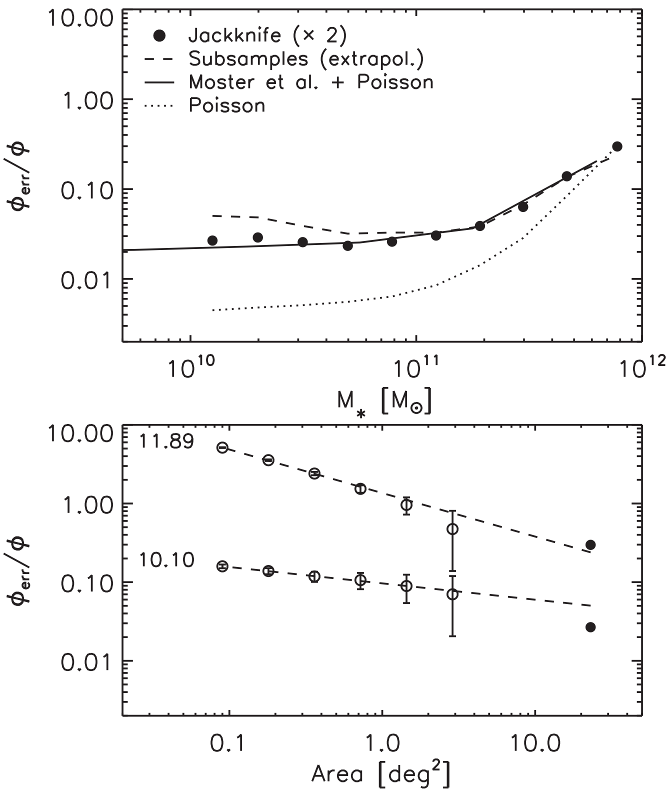


Figure 3. Stellar mass function statistical errors as function of stellar mass (top) and area (bottom). In the top panel, we show the jackknife estimator based on 64 subregions and multiplied by a factor of 2, compared to the theoretical cosmic variance plus Poisson error derived from the Moster et al. (2011) GETCV code (the Poisson error only is shown as the dotted line). The bottom panel shows an alternative internal estimate based on the standard deviation of subregions as a function of their size, in two mass bins ($\log M_\star/\mathrm{M}_\odot = 10.10$ and 11.89), extrapolated to the size of the full survey (dashed lines in both panels). As in the top panel, the black dots are the jackknife estimates, for which the cosmic variance part has been multiplied by 2.

In the bottom panel of Fig. 3, we show an alternative internal estimator as function of area, based on the standard deviation of subsamples with sizes varying from 0.1 to 2.9 deg^2 (the black dots represent our Jackknife estimates in the two mass bins $\langle\log M_\star/\mathrm{M}_\odot\rangle = 10.10$ and 11.89). We use a power-law fit (the amplitudes of the error bars are arbitrarily scaled to the square root of the number of subsamples, ranging from $\sqrt{256}$ to $\sqrt{8}$) to extrapolate to the full size of the survey. The extrapolated values are shown as the dashed line in the top panel of Fig. 3. The bin correlations between small subsamples may tilt the slope of the fit and lead to an overestimate of the extrapolated error estimate, as observed in the low-mass bin. In the high-mass bins, characterized by an uncorrelated sampling variance dominated by Poisson noise, the extrapolated estimate is consistent with both the jackknife estimate and the theoretical Poisson noise.

3.2 Projected galaxy clustering

We measure the two-point correlation function $\widetilde{w}(\theta)$ in 10 logarithmically spaced bins centred on the pair-number weighted averaged separation over the range $0.^\circ002 < \theta < 2^\circ$. The modelled $w(\theta)$ is compared to the measured $\widetilde{w}(\theta)$ by projecting the theoretical spatial

[8] We note that the highest mass bin galaxy bias was estimated a posteriori from our HOD results, since it was not provided by the authors of GETCV, although the contribution of cosmic variance is negligible compared to the Poisson error in this bin, populated by rare massive galaxies.

clustering $\xi(r)$ on to the sample redshift distribution computed as the sum of photometric redshift PDFs (see Section 2.5).

We use the Landy & Szalay (1993) estimator following a similar procedure to that described in Section 3.3 of Coupon et al. (2012). Owing to the limited size of the survey, our measurements are affected by the *integral constraint*, an effect that biases the clustering signal low. Here, we adopt a refined way to correct for it: the correction is calculated directly for every parameter set from the modelled $w(\theta)$ (instead of a pre-determined power law) and integrated over the survey area using random pairs as in Roche et al. (2002), leading to better agreement between the data and the model at large scales. Here, the typical values of the integral constraint range from 10^{-3} to 3×10^{-3}.

We have checked, using the galaxy mocks prepared for the VIPERS sample (de la Torre et al. 2013), that our jackknife error estimates could reproduce within 20 per cent the correct sample variance amplitude of $\widetilde{w}(\theta)$ (this result is in agreement with a number of tests from the literature, e.g. Zehavi et al. 2005; Norberg et al. 2009), and we do not apply any correction.

3.3 Real-space galaxy clustering

We measure the real-space galaxy clustering for the VIPERS spectroscopic sample by integrating the weighted redshift-space correlation function along the line of sight to alleviate redshift-space distortion effects:

$$\widetilde{w_\mathrm{p}}(r_\mathrm{p,phys}) = 2\int_0^{\pi_\mathrm{max}} \widetilde{\xi}(r_\mathrm{p,phys}, \pi_\mathrm{phys})\mathrm{d}\pi_\mathrm{phys}, \tag{2}$$

where $r_\mathrm{p,phys}$ and π_phys are the coordinates perpendicular and parallel to the line of sight, respectively. $r_\mathrm{p,\,phys}$ is expressed in physical coordinates and divided into 10 logarithmically spaced bins centred on the pair-weighted averaged separation over the range $0.2 < r_\mathrm{p,phys}/\mathrm{Mpc} < 10$, and π_phys is divided into linear bins up to $\pi_\mathrm{max} = 40$ Mpc. The value of π_max is consistently used in the derivation of the modelled w_p. As for $\widetilde{w}(\theta)$, $\widetilde{\xi}(r_\mathrm{p,phys}, \pi_\mathrm{phys})$ is computed using the Landy & Szalay estimator and the covariance matrix estimated from the jackknife resampling of 64 subregions.

Each galaxy is weighted to account for the undersampling of the spectroscopic sample: we use the global colour sampling rate (CSR), target sampling rate (TSR) and success sampling rate (SSR), as described in Davidzon et al. (2013), to account for the VIPERS colour selection, the sparse target selection and measurement success as function of signal-to-noise ratio, respectively. In addition, we also use number-count normalized (to prevent global CSR, TSR and SSR double weighting) spatial weights computed for each VIPERS panel by de la Torre et al. (2013) to correct for the position-dependent sampling. Here, the SSR is the most affected quantity, as a function of position in the sky, due to the differing observing conditions at the times of observation.

Small pair incompleteness due to 'slit collision' is corrected by a factor $1 + \widetilde{w}_\mathrm{A}$, such that:

$$1 + \widetilde{w}_\mathrm{p,corr} = \frac{1 + \widetilde{w}_\mathrm{p}}{1 + \widetilde{w}_\mathrm{A}}, \tag{3}$$

where

$$1 + \widetilde{w}_\mathrm{A} = 1 - \frac{0.03}{r_\mathrm{p,phys}} \tag{4}$$

is derived from the projected correlation as function of angular scale by de la Torre et al. (2013) and translated into physical scales at $z = 0.8$. We note that given our conservative small-scale cut of $r_\mathrm{p,phys} > 0.2$, the correction remains below 15 per cent.

3.4 Galaxy–galaxy lensing

The gravitational lensing signal produced by the foreground matter overdensity is quantified by the tangential distortion of background sources behind a sample of stacked 'lens' galaxies, also known as the weighted galaxy–galaxy lensing estimator (e.g. Mandelbaum et al. 2006; Yoo et al. 2006). The excess surface density of the projected dark matter halo relates to the measured tangential shear through:

$$\widetilde{\Delta\Sigma}(r_\mathrm{p,phys}) = \Sigma_\mathrm{crit} \times \widetilde{\gamma}_t(r_\mathrm{p,phys}), \tag{5}$$

(see also Appendix B). We measure the signal in 10 logarithmically spaced bins centred on the number-weighted averaged separation, in the range $0.02 < r_\mathrm{p,\,phys}/\mathrm{Mpc} < 1$. $r_\mathrm{p,\,phys}$ is expressed in physical coordinates.[9]

The critical surface density Σ_crit is given by

$$\Sigma_\mathrm{crit} = \frac{c^2}{4\pi G_\mathrm{N}} \frac{D_\mathrm{OS}}{D_\mathrm{OL}\, D_\mathrm{LS}}, \tag{6}$$

with D_OS the observer-source angular diameter distance, D_OL the observer-lens (foreground galaxy) distance and D_LS the lens-source distance. G_N is the gravitational constant and c the speed of light. All distances are computed in physical coordinates using the photometric (spectroscopic when available) redshift. For photometric redshift values, a cut $z_\mathrm{source} - z_\mathrm{lens} > 0.1 \times (1 + z_\mathrm{lens})$ is adopted. The background source galaxy sample includes all galaxies detected in the i band with a non-zero lensing weight (Miller et al. 2013). Here, we do not restrict our redshift sample to $z_\mathrm{p} < 1.2$, but consider galaxies at all redshifts, taking advantage of the improved photometric redshift estimates in our sample, increasing the background source sample by 30 per cent compared to other CFHTLenS lensing studies, without introducing any systematic bias (see Appendix C).

The galaxy shape measurement was performed on individual exposures using the LENSFIT analysis pipeline (Miller et al. 2007; Kitching et al. 2008; Miller et al. 2013) and systematics checks were conducted by Heymans et al. (2012) for cosmic shear (the projected large-scale structure lensing power spectrum). The lensing (inverse-variance) weights account for shape measurement uncertainties (Miller et al. 2013). Following Velander et al. (2014), who performed extensive systematics checks of the CFHTLenS shear catalogue specifically for galaxy–galaxy lensing (see their Appendix C), we do not reject those CFHTLS-Wide pointings that did not pass the requirements for cosmic shear, and we applied appropriate shape measurement corrections as described in their Section 3.1.

We compute the boost factor (to account for dilution due to sources physically associated with the lens, see Sheldon et al. 2004; Mandelbaum et al. 2006) by randomizing the source positions, and correct the final signal for it. On small scales (<0.1 Mpc), the boost factor reaches up to 20 per cent for the most massive galaxies.

Here, the relatively low source density implies that our errors are dominated by the source galaxy shape noise, originating from ellipticity measurement uncertainties and intrinsic shape dispersion, rather than sample variance. Indeed, when compared to the sum of inverse-variance lensing weights, we have checked that our jackknife estimate was similar at all scales (with small off-diagonal

[9] Note that the galaxy–galaxy lensing signal is measured in physical units, whereas a number of authors assume comoving units, which would require multiplying the excess surface density by a factor of $(1 + z)^{-2}$ compared to our definition.

correlation), confirming the negligible impact of cosmic variance (see Appendix C).

Nevertheless, a correlation exists *between* the mass bins due to the re-use of background source galaxies. We neglect this contribution in the computation of the combined χ^2, but we note that this correlation is likely to lead to underestimation of our parameter confidence limits.

3.5 Systematic errors in stellar mass measurements

In this section, we are concerned with systematic errors affecting the stellar mass measurements caused by the uncertainties in the assumed cosmology (i.e. volume and distance estimates), the dust modelling, and potential biases in the photometry.

To assess the impact of systematics on the measurements of the observables, we propagate the errors affecting the stellar masses by changing one parameter configuration at a time, then re-computing all stellar masses and the observables, and finally measuring the difference with the reference measurements. We repeat the process for the three different kinds of systematics listed below:

(i) assumed cosmology. We explore three Λ cold dark mater (ΛCDM) parameter sets: in addition to the *Wilkinson Microwave Anisotropy Probe* (*WMAP)* cosmology used in this study with $H_0 = 72$ km s^{-1} Mpc^{-1}, $\Omega_{\rm m} = 0.258$, $\Omega_\Lambda = 0.742$ (Hinshaw et al. 2009), a 'concordance' cosmology model with $H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_{\rm m} = 0.3$, $\Omega_\Lambda = 0.7$ and the *Planck* cosmology with $H_0 = 67$ km s^{-1} Mpc^{-1}, $\Omega_{\rm m} = 0.320$, $\Omega_\Lambda = 0.680$ (Planck Collaboration XVI 2014) are tested. In each case, the stellar masses and the observables are consistently re-computed with the same cosmology. We note that the term 'systematics' here refers to the choice for one or another set of parameters that produces a systematic shift in stellar mass and not to systematic errors associated with the measurement of cosmological parameters;

(ii) lens galaxy dust extinction modelling. We compute five different stellar masses for each galaxy by varying one aspect at a time: two different extinction law configurations (among our choice of three laws, see Section 2.6) and three different $E(B-V)$ maximum allowed values (ranging from 0.2 to 0.7);

(iii) photometric calibration. As zero-point offsets do not correct for absolute calibration uncertainties (but do for colours), nor correct for photometric measurement biases (e.g. missing flux of bright objects), a change in the photometric calibration may cause a shift in the best-fitting template and further bias the stellar mass measurements. We re-compute stellar masses applying ad hoc global shifts (in all bands) of -0.05 and $+0.05$ magnitude, which correspond to typical offsets caused by various calibration strategies or photometry measurements (Moutard et al., in preparation).

Results are shown in Fig. 4. For each observable (top left: stellar mass function, top right: projected clustering, bottom left: real-space clustering and bottom right: galaxy–galaxy lensing), we display the re-computed measurements divided by the reference quantities, in each of the 'Cosmology', 'Extinction' and 'Calibration' panels as well as the sum in quadrature of all these effects ('Total'). The grey area corresponds to the maximum value among the differing re-computations, not the standard deviation, as each of the solutions is equally likely to be opted for. Except for the stellar mass function, we only display the results in the most massive bins (where we observe the most significant changes), although the calculations were repeated in all mass bins.

To allow comparison with the statistical errors, we overplot the error bars from our jackknife error estimates. For the stellar mass function (whose jackknife error estimate is multiplied by a factor of 2, see Section 3.1), the systematic errors compared to the statistical errors are striking, with the former being larger by one order of magnitude compared to the latter. The increase of the systematic errors towards the high-mass regime is a direct consequence of the shift in stellar mass and the steep slope of the SMF at the massive end.

It is interesting to note that the different cosmologies lead also to large systematic errors compared to statistical errors. Although many authors in galaxy evolution studies claim to account for cosmological parameter uncertainties by presenting h-free results, we recall that, in a flat Universe, both $\Omega_{\rm m}$ and H_0 enter in the computation of the comoving volume and luminosity distances and, in the precision era of *WMAP* and *Planck*, happen to contribute equally to the distance uncertainties. Comparing our results to the recent literature is therefore not as simple as scaling the different quantities with respect to h, and we must properly account for the more complex dependence of distances on $\Omega_{\rm m}$ and H_0.

In comparison, the projected and real-space galaxy clustering as well as galaxy–galaxy lensing are relatively less prone to systematic errors. For the effect of cosmology, the measurement of projected clustering has no dependence on galaxy distances, and the only difference originates from the modified galaxy selection caused by the stellar mass shift. Interestingly, although the real-space clustering and the galaxy–galaxy lensing do depend on galaxy distance measurements, the change in cosmology also has little impact at the level of our statistical errors. We can draw similar conclusions on the effects of dust extinction modelling and photometric calibration.

Obviously, the stellar mass function is the measured quantity suffering from the largest systematic error contribution, compared to the statistical errors. In particular, we will see in Section 5 that most of the constraints on the central galaxy $M_\star - M_{\rm h}$ relationship emanate from the stellar mass function and taking into account these systematic uncertainties when comparing our results with the literature is necessary.

Ideally, one would like to estimate a best-fitting model for each of the re-computed quantities. Unfortunately, this would be computationally very expensive. Instead, we create two sets of measurements: a 'statistical error' set based on our jackknife error estimate and a 'total error' set for which we add in quadrature the systematic errors (assuming they are Gaussian distributed) and the statistical errors. We present in Section 5 separate results for both.

3.6 Impact of photometric redshift uncertainties

The dispersion of photometric redshifts may also cause systematic effects of several kinds, first on the stellar mass function, as a contribution to the stellar mass scatter, which shifts towards higher masses the high-mass end where the slope is steep, an effect known as *Eddington bias*. Secondly, the projected clustering amplitude is biased low due to the scattering of galaxies falling outside the mass bins.

We will see in Section 4 that our model properly accounts for these systematic effects caused by photometric redshift dispersion, through the parametrization of the stellar mass scatter. However, catastrophic failures and photometric redshift biases may be more problematic. We have demonstrated in Section 2.5 that our catastrophic error rate was not higher than 4 per cent, and based on results from Section 3.2 of Coupon et al. (2012), such a low contamination rate should have no impact on clustering results at our statistical

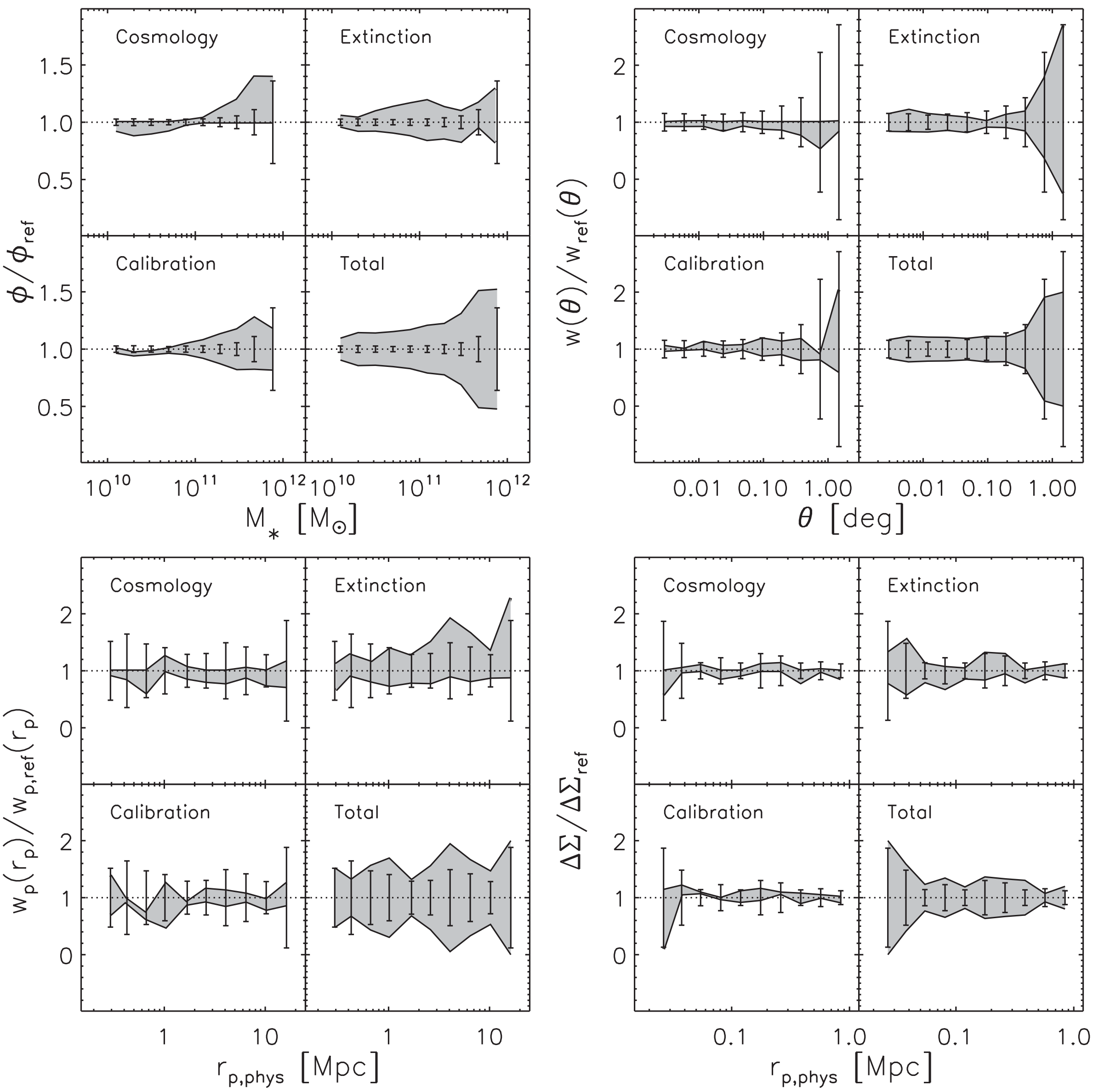


Figure 4. Systematic errors affecting the galaxy stellar mass function (top left), the projected correlation function (top right), the real-space correlation function (bottom left) and the galaxy–galaxy lensing signal (bottom right). In each panel, the grey area symbolizes the envelope (maximum value) of the re-computed measurement compared to the reference. The error bars are statistical errors from the internal jackknife estimator. The 'Total' panel represents the symmetric sum in quadrature of all three contributions. Here, we only show the most massive bins for the clustering and lensing measurements, however we repeated the tests in all mass bins.

error level. To check this statement on the calibration sample (which means the conclusions are limited to the photometric sample with similar properties to the spectroscopic sample), we use the VIPERS galaxies with spectroscopic redshift and re-compute all stellar masses, as well as each observable, using the corresponding photometric redshift. We show the measurements in Fig. 5 (solid lines) divided by the reference measurement made with spectroscopic redshifts and where the error bars are from the statistical jackknife estimator. From left to right, we display the results for the stellar mass function, the projected clustering and the galaxy–galaxy lensing signal, all in the mass range $10^{10} < M_\star/\mathrm{M}_\odot < 10^{12}$ and redshift range $0.5 < z < 1$.

We conclude that for galaxies with similar properties to VIPERS galaxies, none of the observables measured with photometric redshifts display a large bias with respect to the spectroscopic redshift ones. This represents a reassuring sanity check for the calibration procedure. Only the projected clustering presents a slightly low systematic value, expected from the dispersion of redshifts and accounted for in the model, through the projection of the modelled 3D clustering on the redshift distribution constructed from the sum of photometric redshift PDFs (assuming that estimated PDFs are representative of the true PDFs).

4 MODEL AND FITTING PROCEDURE

We use the HOD formalism to connect galaxy properties to dark matter halo masses. Here, we assume that the number of galaxies per halo is solely a function of halo mass, split into central and satellite contributions. The fitting procedure then consists of finding a set of parameters to describe the HOD that best reproduces the observables.

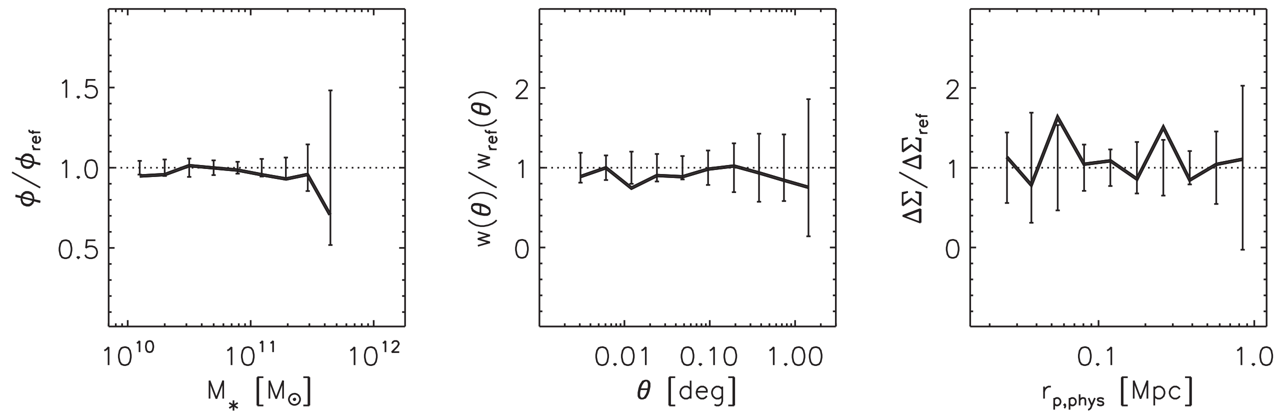


Figure 5. Measurements made with photometric redshifts divided by those made with spectroscopic redshifts. From left to right: the stellar mass function, the projected clustering and the galaxy–galaxy lensing signal, all made with VIPERS galaxies in the mass range $10^{10} < M_\star/\mathrm{M}_\odot < 10^{12}$ and redshift range $0.5 < z < 1$. Error bars represent the statistical error estimates from jackknife resampling.

A key ingredient of the HOD model is the statistical description of the spatial distribution of dark matter. We assume that the matter power spectrum, the halo mass function and the dark matter halo profile are all known quantities over the scales and redshift range ($0.5 < z < 1$) explored in this study. All the technical details about the halo model are given in appendix A.1 of Coupon et al. (2012), with the exception of the large-scale halo bias, for which we use in this study the fitting formula proposed by Tinker et al. (2010).

The exact way to parametrize the HOD is often at the origin of the differences between HOD studies in the literature. In this paper, we follow Leauthaud et al. (2011) who adopted two advanced features:

(i) the HOD is a conditional function of the stellar mass given the halo mass (this formalism is an extension of the conditional luminosity function technique developed by Yang et al. 2003). In this formalism, the central galaxy $M_\star - M_\mathrm{h}$ relationship is a parametrized function representing the mean stellar mass given its host halo mass, $\langle M_\star | M_\mathrm{h}\rangle$;
(ii) all observables, namely the stellar mass function, the projected clustering, the real-space clustering and the galaxy–galaxy lensing signal are fitted jointly.

4.1 The stellar-to-halo mass relationship

To describe the central galaxy $M_\star - M_\mathrm{h}$ relationship, we adopt the parametrized function $f_{\mathrm{SM\text{-}HM}}$ proposed by Behroozi et al. (2010), and defined via its inverse:

$$\log_{10}\left(f^{-1}_{\mathrm{SM-HM}}\right) = \log_{10}(M_\mathrm{h}(M_\star))$$

$$= \log_{10}(M_1) + \beta\,\log_{10}\left(\frac{M_\star}{M_{\star,0}}\right) + \frac{\left(\frac{M_\star}{M_{\star,0}}\right)^{\delta}}{1+\left(\frac{M_\star}{M_{\star,0}}\right)^{-\gamma}} - \frac{1}{2}. \quad (7)$$

M_1 controls the scaling of the relation along the halo mass coordinate, whereas $M_{\star,0}$ controls the stellar mass scaling. β, δ and γ control the low-mass, high-mass and curvature of the relation, respectively.

4.2 The central occupation function

For central galaxies contained in a threshold sample ($M_\star > M_\star^\mathrm{t}$), the HOD is defined as a monotonic function increasing from 0 to 1, with a smooth transition centred on the halo mass value corresponding to $M_\star^\mathrm{t}$:

$$\langle N_\mathrm{cen}(M_\mathrm{h}|M_\star^\mathrm{t})\rangle$$

$$= \frac{1}{2}\left[1 - \mathrm{erf}\left(\frac{\log_{10}(M_\star^\mathrm{t})) - \log_{10}(f_{\mathrm{SM-HM}}(M_\mathrm{h}))}{\sqrt{2}\sigma_{\log M_\star}(M_\star^\mathrm{t})}\right)\right]. \quad (8)$$

The parameter $\sigma_{\log M_\star}$, expresses the scatter in stellar mass at fixed halo mass, which we parametrize as:

$$\sigma_{\log M_\star}(M_\star^\mathrm{t}) = \sigma_{\log M_\star,0}\left(\frac{M_\star^\mathrm{t}}{10^{10}\,\mathrm{M}_\odot}\right)^{-\lambda}, \quad (9)$$

to account for the change in intrinsic stellar mass dispersion as a function of stellar mass.

4.3 The satellite occupation function

We describe the satellite HOD for a threshold sample $M_\star^\mathrm{t}$ with a simple power law as a function of halo mass M_h:

$$\langle N_\mathrm{sat}(M_\mathrm{h}|M_\star^\mathrm{t})\rangle = \left(\frac{M_\mathrm{h} - M_\mathrm{cut}}{M_\mathrm{sat}}\right)^{\alpha}, \quad (10)$$

for which we fix the cut-off mass scale M_cut such that

$$M_\mathrm{cut} = f^{-1}_{\mathrm{SM-HM}}(M_\star^\mathrm{t})^{-0.5}. \quad (11)$$

This assumption is based upon the values reported by Coupon et al. (2012) for their equivalent parameter 'M_0'. We have checked that the exact parametrization of M_cut had very little importance compared to the other parameters and did not change any of our conclusions, in agreement with the loose constraints observed by Coupon et al. (2012).

As in Leauthaud et al. (2011), the normalization M_sat of the satellite HOD follows the halo mass scaling driven by the central $M_\star - M_\mathrm{h}$ relationship, with some degree of freedom controlled by a power law:

$$\frac{M_\mathrm{sat}}{10^{12}\,\mathrm{M}_\odot} = B_\mathrm{sat}\left(\frac{f^{-1}_{\mathrm{SM-HM}}(M_\star^\mathrm{t})}{10^{12}\,\mathrm{M}_\odot}\right)^{\beta_\mathrm{sat}}. \quad (12)$$

4.4 Total occupation functions and observables

Finally, the total HOD is

$$\langle N_\mathrm{tot}(M_\mathrm{h}|M_\star^\mathrm{t})\rangle$$

$$= \langle N_\mathrm{cen}(M_\mathrm{h}|M_\star^\mathrm{t})\rangle + \langle N_\mathrm{sat}(M_\mathrm{h}|M_\star^\mathrm{t})\rangle, \quad (13)$$

and since our measurements are made in bins of stellar mass, we transform the threshold HOD functions into binned functions by writing:

$$\langle N_{\rm tot}(M_{\rm h}|M_\star^{\rm t_1}, M_\star^{\rm t_2})\rangle$$
$$= \langle N_{\rm tot}(M_{\rm h}|M_\star^{\rm t_1})\rangle - \langle N_{\rm tot}(M_{\rm h}|M_\star^{\rm t_2}\rangle. \quad (14)$$

Equivalent relations hold for central and satellite binned HODs.

The stellar mass function, the projected two-point correlation function, the real-space correlation function and the galaxy–galaxy lensing signals are computed from the halo model and the HOD as detailed in Appendix B.

4.5 Systematic errors in the model

As detailed in the previous sections, the HOD formalism relies on an accurate description of the dark matter spatial distribution. Here, we evaluate the impact of our model uncertainties and assumptions on the best-fitting HOD parameters. Ideally, one would like to repeat the fitting procedure to test each of the different assumptions of the model, but to avoid such a time-consuming exercise, we take the simple approach of modifying one feature at a time, and tuning the HOD parameters by hand to reproduce the modelled quantities derived from the best-fitting parameters reported in Section 5. We explore two stellar mass bins ($M_\star = 10^{10}, 10^{11.5}\,{\rm M_\odot}$) and we focus on the two parameters M_1 and $B_{\rm sat}$, controlling the halo-mass scaling of the $M_\star - M_{\rm h}$ relationship, and the normalization of the satellite HOD, respectively. The results are shown in Table 3, and we detail below our calculations for each assumption listed.

The power spectrum normalization parameter, σ_8, is currently known to a precision of a few per cent. This parameter has a strong impact on the large-scale galaxy clustering, and a larger value would lead to an increased number of massive structures, hence shifting the massive end of the halo mass function towards more massive haloes. Choosing *Planck* over *WMAP*7 cosmology (as for the tests in Section 3.5), would result in a 5 per cent increase in σ_8, leading to relatively small changes in best-fitting HOD parameters, of the order of a few per cent.

Halo bias uncertainties originate from the measurement of the bias-to-halo mass relation $b(M_{\rm h})$ using N-body simulations, affected by low-mass resolution, small volume, or the limitations of halo identification techniques. In the low-clustering regime, the typical errors on the bias are as small as a few per cent (Tinker et al. 2010), however the rather shallow slope of bias versus halo mass (see e.g. fig. 18 of Coupon et al. 2012) translates into a larger uncertainty in the deduced halo mass. In the high-mass regime, errors are mainly dominated by the sample variance of simulations, up to ~10 per cent, but have fewer impact on the deduced halo mass owing to the steeper slope in this regime.

The assembly bias (Zentner et al. 2014, and references therein) refers to the correlation between clustering amplitude and time of halo formation, whereas in our model the bias is assumed to vary only with halo mass. The effect is stronger when selecting a population of galaxies based on a parameter correlated with halo formation history, such as the SFR, but moderate when considering the full galaxy population selected by stellar mass only. In this case, and in the mass regime explored in this study, Zentner et al. (2014) found that the systematics caused by assembly bias on HOD parameters do not exceed 10–15 per cent.

In our model, the dark matter halo profile is assumed to follow a Navarro, Frenk & White (1997, NFW) profile. While lensing observations tend to favour NFW profiles (Umetsu et al. 2011; Coupon, Broadhurst & Umetsu 2013; Okabe et al. 2013), the mass–concentration relation – driving the slope of the profile – remains uncertain. We have used a simple mass–concentration relation based on theoretical predictions (updated from Takada & Jain 2003) and empirical redshift evolution (Bullock et al. 2001), but more recent relations such as the work from Muñoz-Cuartas et al. (2010) have been measured. Compared to our concentration values, the difference with Muñoz-Cuartas et al. rises from 11 per cent at $M_{\rm h} \sim 10^{12}\,{\rm M_\odot}$ to 30 per cent at $\sim 10^{15}\,{\rm M_\odot}$ (with a minimum of 2 per cent at $\sim 10^{13}\,{\rm M_\odot}$). These systematics affect the slope of the small-scale clustering and galaxy–galaxy lensing. We estimate that if all of our constraints came from lensing, this may result in a 28 per cent systematic error in M_1.

We assume that the satellite distribution in the halo follows the dark matter density profile. However, this assumption may not be always true and Budzynski et al. (2012) tested this hypothesis from a stacked analysis of massive clusters from the SDSS. They found a typical factor of 2 (with ~50 per cent scatter) lower concentration of the satellite distribution compared to dark matter, whereas Muzzin et al. (2007) measured a value closer to dark matter around $z \sim 0.3$, and van der Burg et al. (2014) a relatively high concentration at $z = 1$. These trends may show a redshift evolution of the concentration or can simply be inherent to the difficulty of observationally measuring the satellite distribution. In Table 3, we report the impact on $B_{\rm sat}$ after setting the satellite concentration a factor of 2 higher than that of dark matter. The effect on $B_{\rm sat}$ does not exceed 11 per cent.

Finally, in our model we neglect the lensing contribution of the subhaloes hosting the satellite galaxies. This effect, first introduced by Mandelbaum et al. (2005b) under the term 'stripped satellite central profile', assumes that a fraction of the satellite haloes survive inside the host halo and further contribute to the lensing signal at

Table 3. Estimated systematic errors from the model on the central halo mass, $\log_{10}M_1$, and the satellite normalization, $B_{\rm sat}$. The total is the sum in quadrature of the errors.

Assumption	Error on $\log_{10}M_1(\sim 12.7)$ $M_\star({\rm M_\odot}) = 10^{10}$	$10^{11.5}$	Error on $B_{\rm sat}(\sim 10)$ $M_\star({\rm M_\odot}) = 10^{10}$	$10^{11.5}$	Affected quantities
σ_8	0.05	0.05	1	0.5	SMF, clustering (small and large scales)
$b(M_{\rm h})$ relation	0.08	0.1	–	–	Clustering (large scale)
Assembly bias[a]	<0.04	<0.04	~1.5	~1.5	SMF, clustering (small and large scales)
$c(M_{\rm h})$ relation	0.11	0.03	0.1	0.4	Clustering (small scale), lensing
Satellite concentration	–	–	1.1	0.9	Clustering (small scale), lensing
Stripped subhaloes	0.09	0.07	–	–	Lensing
Total	0.17	0.14	2.1	1.9	All

Note. [a]From Zentner, Hearin & van den Bosch (2014).

small scales. As a result, the lensing contribution of those subhaloes adds up to the central-galaxy halo term in such a way that the best-fitting *host* halo mass gets reduced compared to a model in which the contribution of subhaloes is neglected. Hudson et al. (2015) quantify the systematic change in best-fitting halo mass as a systematic decrease by a factor of $\sim(1 + f_{\rm sat})$, where $f_{\rm sat}$ is the fraction of satellites in the sample. Assuming a satellite fraction between 20 and 30 per cent, this leads to a systematic error of up to 0.09 in $\log_{10}M_1$. This number must be read as if all the constraints would come from lensing only. In our study where the stellar mass function and the clustering signal-to-noise ratio is higher than that of the lensing, this effect plays relatively little role, and our results would not significantly change if we accounted for it.

The sum in quadrature of these model systematics is shown as 'Total' in Table 3. Intermediate stellar mass bins ($\sim10^{10}\,{\rm M}_{\odot}$) seem to be most affected, with an error of 0.17 for $\log_{10}M_1$ ($\sim$50 per cent in M_1) and 2.1 ($\sim$20 per cent) for $B_{\rm sat}$. We will see below that these values dominate over the typical statistical and systematic errors from the measurements in this mass regime. However, as each of these systematic errors affects the observables in a different way and we fit all of them jointly, one must see these numbers as pessimistic estimates. The high-stellar mass bin ($\sim10^{11.5}\,{\rm M}_{\odot}$) is equally affected but in a regime where statistical errors are large, hence leading to a smaller impact.

4.6 MCMC sampling

We write the combined log-likelihood as the sum of each observable χ^2:

$$-2\ln\mathcal{L} = \chi^2_{\phi} + \sum_{\rm spl}\chi^2_{w(\theta)} + \sum_{\rm spl}\chi^2_{w_{\rm p}(r_{\rm p})} + \sum_{\rm spl}\chi^2_{\Delta\Sigma}\,, \tag{15}$$

where individual χ^2's are computed as

$$\chi^2 = \sum_{i,j}[\widetilde{X}_i - X_i](C^{-1})_{ij}[\widetilde{X}_j - X_j]\,, \tag{16}$$

using the covariance matrices evaluated for each measurement as described in Section 3 ($\widetilde{X}$ and X represent the measured and modelled observables, respectively). Each observable χ^2 is summed over the samples ('spl') as described in Table 2. The 'i' and 'j' subscripts refer to the stellar mass (stellar mass function) or transverse separation (clustering and lensing) binning of each measurement.

We find the best-fitting parameters and posterior distribution (assuming flat priors for all parameters) employing the MCMC sampling technique with the Metropolis–Hastings sampler from the software suite COSMOPMC (Wraith et al. 2009). We check for individual chain convergence and chain-to-chain mixing using the Gelman & Rubin (1992) rule from the R-language CODA package.[10] We find a typical chain-to-chain mixing coefficient (potential scale reduction factor) to be equal to 1.01, and the acceptance rate around 30 per cent.

In practice, we first evaluate a diagonal Fisher matrix at the maximum likelihood point found using the Amoeba algorithm (Press et al. 2002) and run 10 chains in parallel with the inverse Fisher matrix as the MCMC sampler covariance matrix. The acceptance rate is usually very low due to the noisy diagonal Fisher matrix affected by some strong correlations between parameters. Once the chains have converged (after typically 5000–10 000 steps) we compute the final likelihood covariance matrix after rejecting the burn-in phase of the chains (a few thousand steps). This covariance matrix is used as the input sampler covariance matrix of a second and final MCMC run, in which 10 chains of 30 000 steps each are computed in parallel and combined together assuming a burn-in phase of 2000 steps and checking for proper mixing.

We run the full MCMC procedure twice. The first run is performed using the statistical covariance matrices from the jackknife estimator and the second MCMC run uses the total error covariance matrices, which are constructed from the statistical covariance matrices plus the systematic error estimates added in quadrature to the diagonal, as described in Section 3.5.

Table 4. HOD best-fitting parameters and 68 per cent confidence limits (CL) for the statistical errors (top) and total errors (bottom).

Parameter	Mean	Upper CL	Lower CL
Jackknife resampling errors			
$\log_{10}M_1$	12.84	0.020	−0.026
$\log_{10}M_{\star,0}$	10.98	0.015	−0.019
β	0.48	0.017	−0.021
δ	0.63	0.094	−0.073
γ	1.60	0.166	−0.202
$\sigma_{\log M_{\star,0}}$	0.337	0.045	−0.035
λ	0.21	0.047	−0.044
$B_{\rm sat}$	10.87	0.443	−0.416
$\beta_{\rm sat}$	0.83	0.038	−0.035
α	1.17	0.020	−0.021
Total errors			
$\log_{10}M_1$	12.67	0.124	−0.083
$\log_{10}M_{\star,0}$	10.90	0.082	−0.067
β	0.36	0.077	−0.051
δ	0.75	0.193	−0.151
γ	0.81	0.477	−0.386
$\sigma_{\log M_{\star,0}}$	0.394	0.100	−0.074
λ	0.25	0.082	−0.083
$B_{\rm sat}$	9.96	0.938	−0.845
$\beta_{\rm sat}$	0.87	0.078	−0.065
α	1.14	0.040	−0.038

5 RESULTS

Best-fitting parameters with 68 per cent confidence intervals are given in the top panel of Table 4 for the statistical- and total-error MCMC runs. The 1D and 2D likelihood distributions are shown in Fig. D1. The reduced χ^2_ν for the statistical-error fit is $\chi^2/(N_{\rm points} - N_{\rm parameters}) = 260/(160 - 10) = 1.7$, which is an overestimate given the correlations neglected in the computation of the log-likelihood. Firstly, we recall that the lensing and clustering measurements are affected by a sample-to-sample correlation due to the scatter in stellar mass. The re-use of background galaxies in the lensing measurements causes an additional sample-to-sample correlation. Secondly, the projected and real-space clustering are correlated, as both observables bring similar information. This mostly affects the satellite distribution parameter errors, which could be slightly underestimated. Finally, the few number of subsamples (64) used in the computation of a noisy covariance matrix may have biased the inverse estimate and contributed to an increase in χ^2_ν.

5.1 Measurements and best-fitting models

The measured stellar mass function and best-fitting model are displayed in Fig. 6. Statistical error bars and corresponding best-fitting model are shown as thick black lines, whereas total (statistical plus

[10] http://cran.r-project.org/web/packages/coda/citation.html

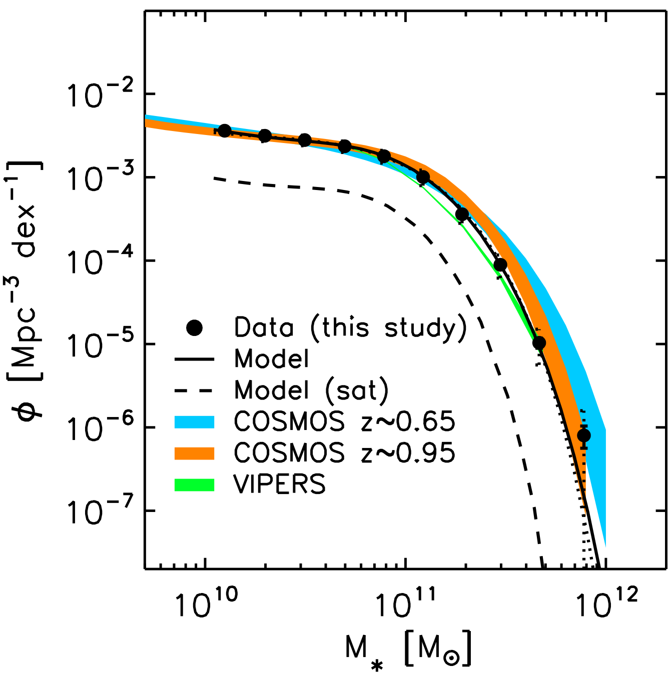


Figure 6. Measured stellar mass function and best-fitting model in the range $0.5 < z < 1$. The statistical errors from the jackknife estimate are shown as black thick lines, whereas the total (statistical plus systematic) error bars as dotted lines. The COSMOS (Ilbert et al. 2013) and VIPERS (Davidzon et al. 2013) mass functions are displayed with their respective statistical errors as shaded areas.

systematic) errors and corresponding best-fitting model are represented in dotted lines. We compare our measurements with the COSMOS mass function evaluated in the ranges $0.5 < z < 0.8$ and $0.8 < z < 1.1$ by Ilbert et al. (2013), and the VIPERS stellar mass function (Davidzon et al. 2013), measured in the range $0.5 < z < 1$ (Davidzon, private communication).

The clustering measurements and best-fitting models are shown in Fig. 7. The projected two-point correlation functions $w(\theta)$ are displayed in the top panels. The mass ranges are given in each top-right corner in units of $\log(M_\star/\mathrm{M}_\odot)$. Similarly, the real-space two-point correlation functions $w_{(r_\mathrm{p})}$ are displayed in the bottom panels.

The galaxy–galaxy lensing measurements and best-fitting models are shown in Fig. 8. The most massive lensing bin features a few data points lower than the model around the transition between the central and the satellites term.

For all observables, we report good agreement between the data and the model. The constraints on the shape of the central $M_\star - M_\mathrm{h}$ relationship (parametrized by $\log_{10}M_1$, $\log_{10}M_{\mathrm{star0}}$, β, γ and δ), are mostly driven by the high signal-to-noise stellar mass function measurements. Satellite HOD parameters (B_sat, β_sat and α) are mainly constrained by the clustering and lensing measurements. The amplitude of clustering at small scale is directly proportional to the relative number of satellites, hence giving strong leverage on the satellite galaxy HOD. Additional information is given on scales $r \sim 0.1$ Mpc from lensing, through the satellite lensing signal. The dispersion in $M_\star$ at fixed M_h, parametrized in amplitude by $\sigma_{\log M_\star,0}$ and in power-law slope by λ, is mainly constrained by the high-mass end of the stellar mass function and the amplitude of the galaxy–galaxy lensing signal in the most massive bins, resulting in a high-mass ($M_\star \sim 10^{11}\,\mathrm{M}_\odot$) scatter of approximately $\sigma_{\log M_\star} \simeq 0.2$ in both the jackknife and total error cases, and a medium mass ($M_\star \sim 10^{10}\,\mathrm{M}_\odot$) scatter of $\sigma_{\log M_\star} \simeq 0.35$.

Because the stellar mass function is most affected by the inclusion of systematics in the error budget, we note a significant increase in uncertainties associated with the parameters driving the central $M_\star - M_\mathrm{h}$ relationship. From Table 4, we report an increase from a factor $\sim$3 in the error in γ, up to a factor $\sim$6 in the error in $\log_{10}M_1$. HOD parameters describing the satellite occupation function such as B_sat, β or α show substantially less sensitivity to the addition of systematic errors in the error budget (a maximum of factor $\sim$2 increase is found). This is explained by the relatively smaller contribution of systematic versus statistical errors affecting the clustering and lensing measurements, compared to the stellar mass function.

The occasional large differences between best-fitting parameters from statistical alone and total errors, seen in Table 4, do not lead to significantly different derived quantities, owing to the strong correlations between parameters. This is confirmed by the almost indistinguishable dotted lines and thick lines in Figs 6–8, and is most probably a consequence of having symmetrically added the systematic errors to the statistical errors.

5.2 Central $M_\star - M_\mathrm{h}$ relationship and the SHMR

In Fig. 9, we show the best-fitting central galaxy $M_\star - M_\mathrm{h}$ relationship (left-hand panel) as parametrized by equation (7), and the SHMR (right-hand panel). The SHMR is shown as function of host halo mass and is derived for the central galaxy in dark grey (from the $M_\star - M_\mathrm{h}$ relationship), the satellites in light grey (integrated over the galaxies above a mass threshold of $M_\star > 10^{10}\,\mathrm{M}_\odot$), and the total in black.

The shaded areas represent the 68 per cent confidence limits, and in the bottom-left panel, we have shown the results obtained with statistical errors in light blue and with total errors in black. As for the stellar mass function, the statistical uncertainties grow by a factor of $\sim$2–4 in the lower mass regime, when incorporating systematics.

The central SHMR peak position is indicated by a black arrow located at $M_{\mathrm{h,peak}} = 1.92^{+0.17}_{-0.14} \times 10^{12}\,\mathrm{M}_\odot$. The SHMR peak value is $\mathrm{SHMR}_\mathrm{peak} = 2.2^{+0.2}_{-0.2} \times 10^{-2}$. When accounting for satellites, the peak position and value do not significantly differ from the estimates for centrals only. However, a remarkable result highlighted in this figure is the increasing contribution of stellar mass enclosed in satellites as function of halo mass. When reaching cluster-size haloes, this contribution reaches over 90 per cent (and presumably higher when accounting for satellite galaxies with masses lower than $10^{10}\,\mathrm{M}_\odot$). However, we stress that we do not take into account the intracluster light, which is challenging to quantify using ground-based photometric data.

5.3 Comparison with the literature

In Figs 10 and 11, we compare our best-fitting $M_\star - M_\mathrm{h}$ relationship for central galaxies with a number of results from the literature. As described in Section 4, our relation describes the mean stellar mass at fixed halo mass which is, due to the scatter in stellar mass, not equivalent to the mean halo mass at fixed stellar mass. This issue becomes particularly important when the slope of the stellar or halo mass distribution is steep (i.e. at high mass). Therefore, we have recomputed our results using the latter definition and we consistently compare our results with the literature in each case.

When required, we convert halo masses to our virial definition using the recipe given by Hu & Kravtsov (2003) in their Appendix C and, following Ilbert et al. (2010), we divide stellar masses by a factor of 1.74 and 1.23 to convert from Salpeter (1955) and 'Diet'

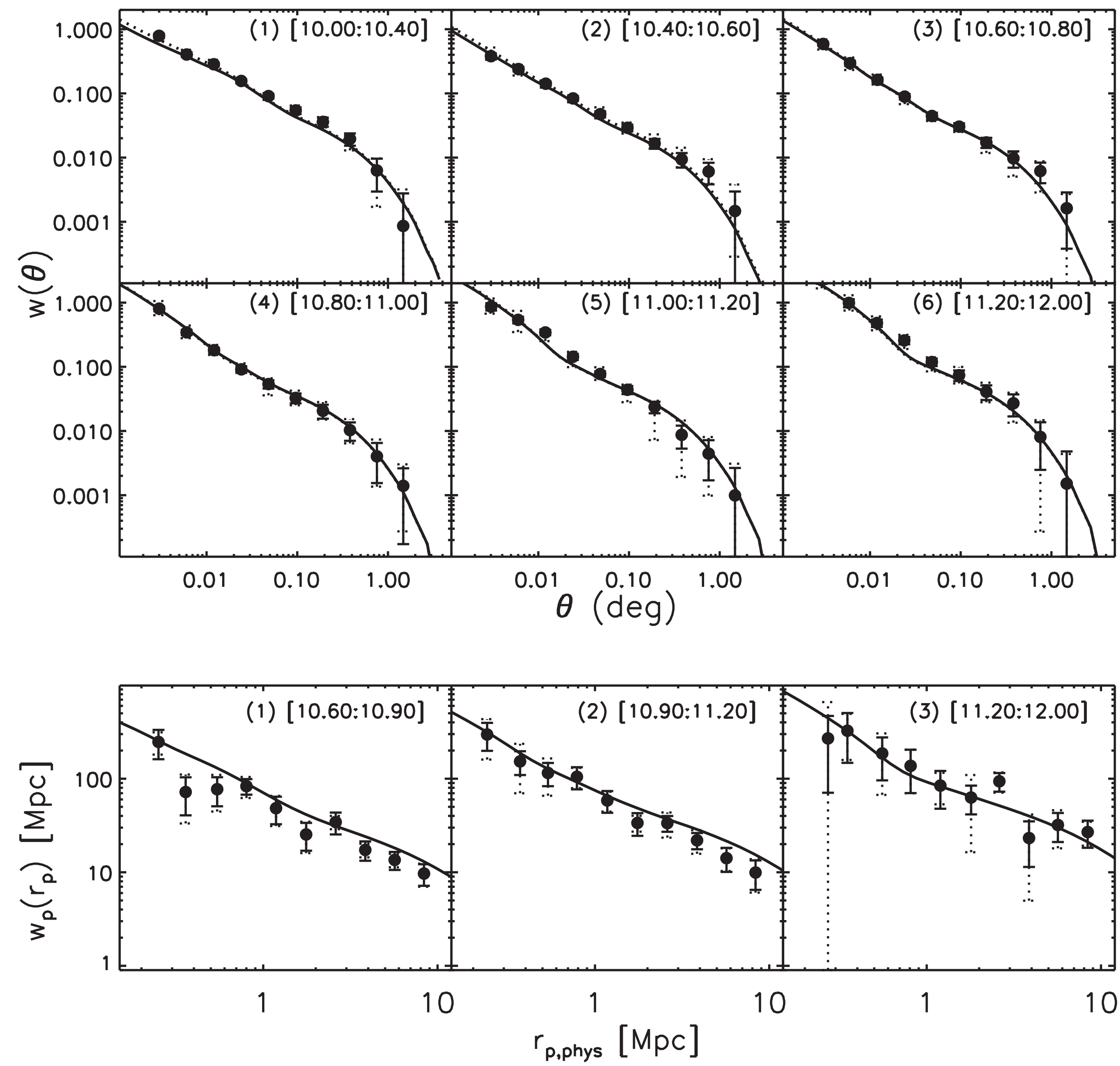


Figure 7. Galaxy clustering measurements (data points with error bars) and best-fitting models (thick lines). The top panels show the projected $w(\theta)$ from the photometric sample (the measurements are corrected for the integral constraint), and the bottom panels show the spectroscopic real-space $w_{\mathrm{p}}(r_{\mathrm{p}})$. The thick error bars associated with thick lines represent the statistical errors and subsequent best-fitting models, whereas dotted lines are for total errors. The mass ranges in the top-right corner of each panel are given in $\log(M_\star/\mathrm{M}_\odot)$.

Salpeter IMFs, respectively, to our Chabrier IMF stellar masses. We apply no correction to Kroupa (2001) IMF stellar masses.

The mean redshift, measured from the sum of the photometric redshift PDFs, is found to be $\langle z\rangle = 0.82$ for our measurements in the range $0.5 < z < 1.0$ ($M_\star > 10^{10.40}\,\mathrm{M}_\odot$) and $\langle z\rangle = 0.65$ in the range $0.5 < z < 0.7$ ($10^{10} < M_\star/\,\mathrm{M}_\odot < 10^{10.40}$). We point out that the lensing signal is more sensitive to lower redshift lens galaxies characterized by a higher signal-to-noise ratio (due to the more numerous background sources), and is likely to be more representative of a lower redshift population, but this effect is assumed to be small compared to the lensing statistical errors.

5.3.1 $\langle M_\star|M_{\mathrm{h}}\rangle$ *results*

We first compare the results for $\langle M_\star|M_{\mathrm{h}}\rangle$ in Fig. 10. The black shaded area shows our results for the central galaxy relationship with 68 per cent confidence limits from the total errors. The total errors consist of the statistical uncertainties from jackknife resampling in addition to three sources of systematic effects from the measurements: the cosmology chosen among widely-used ΛCDM models, the fine-tuning of our dust extinction law modelling, and potential biases in the photometry/calibration. We recall that this list of systematic uncertainties is not exhaustive and, for example, ignores the choice of SPS models, which may be responsible for even larger systematic effects. An estimate of the systematic errors from the model, as detailed in Section 4.5, is also shown in the bottom-right corner.

Behroozi et al. (2013b), shown as the light-blue shaded area, put constraints on the $M_\star$–M_{h} relationship by populating dark matter haloes in simulations and comparing abundances using observed stellar mass functions from a number of surveys. They

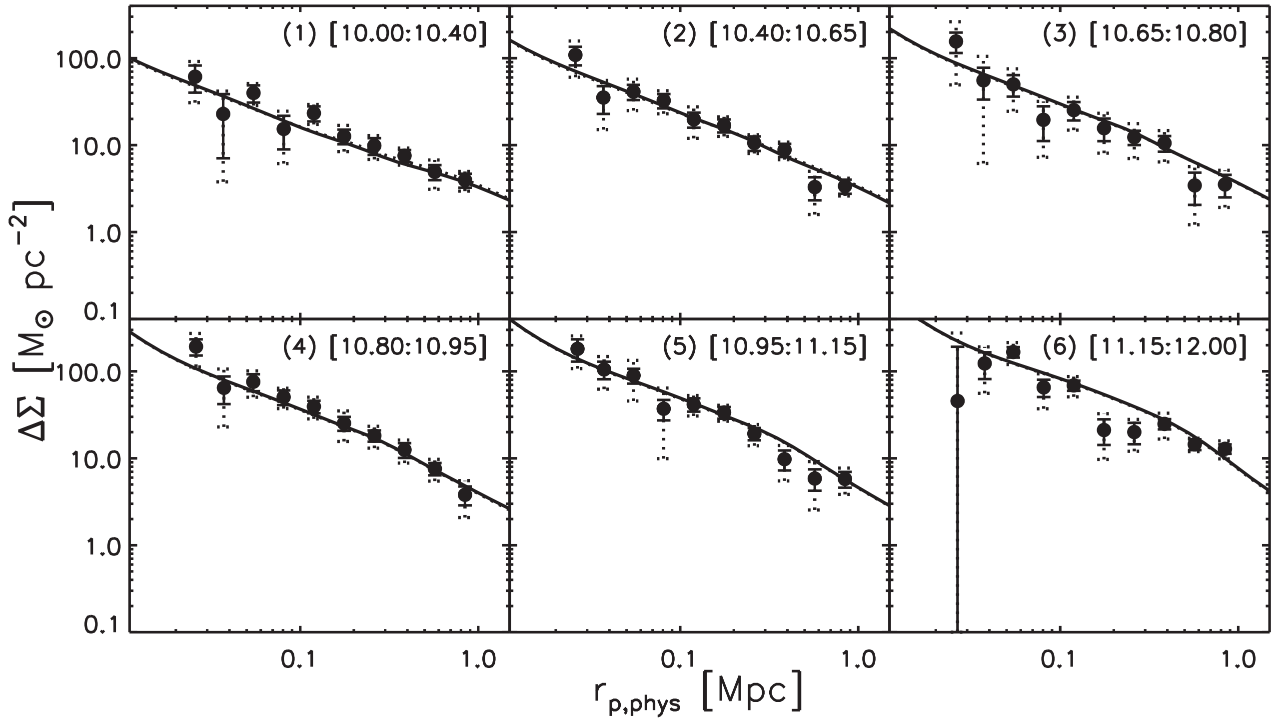


Figure 8. Galaxy–galaxy lensing signal measurements (data points with error bars) and best-fitting models (thick lines). As in Fig. 7, thick and dotted lines are for statistical and total error results, respectively. The mass ranges in the top-right corner of each panel are given in $\log(M_\star/\mathrm{M}_\odot)$.

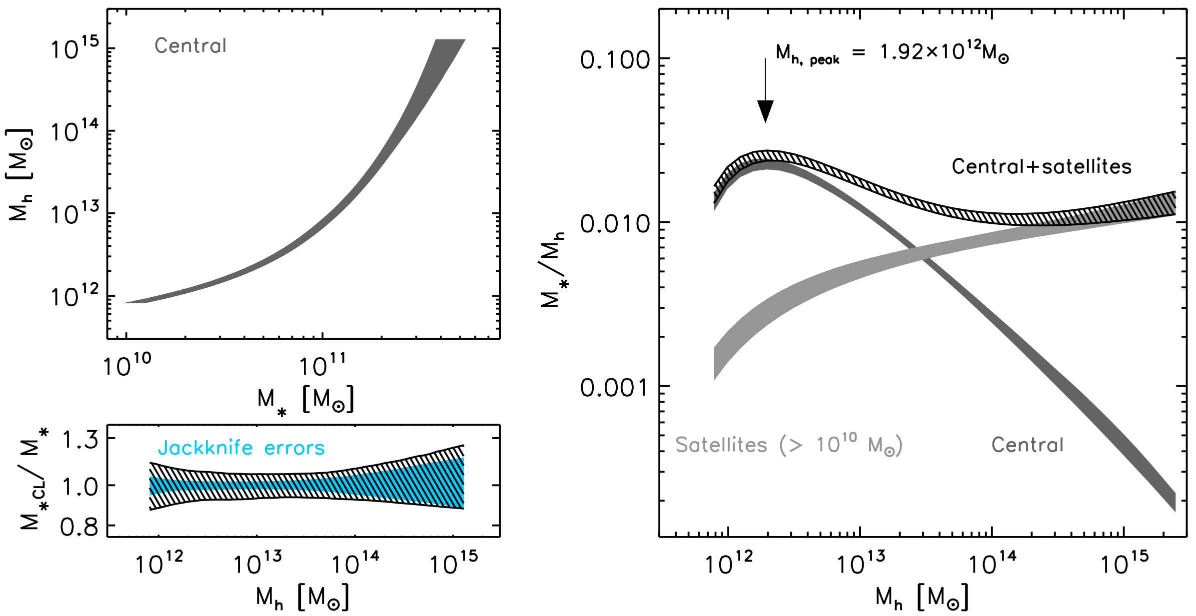


Figure 9. Best-fitting $M_\star - M_\mathrm{h}$ relationship (left) and SHMR (right). The black shaded areas represent the confidence limits from the total errors. The bottom-left panel shows the confidence limits interval as a function of halo mass in the case of statistical errors (from jackknife resampling in light blue) and total errors (in black). The SHMR is derived as function of host halo mass for the central galaxy (dark grey), the satellites (light grey) and the sum of both (black). The peak value of the central SHMR is indicated by the black arrow.

characterized the uncertainties affecting stellar mass estimates by accounting for a number of systematic errors. In particular, unlike in our systematic errors, the authors had to include uncertainties arising from the choice of the IMF and the SPS galaxy templates, necessary when combining the stellar mass functions from several works using different stellar mass measurement methods. Here, we consider their results at $z \sim 1$. A significant difference with our model resides in the assumption that satellite galaxies in larger haloes are seen as central galaxies in subhaloes. To circumvent the difficulty of accurately predicting a subhalo mass function (e.g. complications from tidal stripping), the galaxies in subhaloes at the time of interest are matched to their progenitors at the time of merging on to the central galaxy halo, under the assumption that the $M_\star - M_\mathrm{h}$ evolution at a given stellar mass is identical whether the host halo is isolated or inside a larger halo. In comparison, our model is a 'snapshot' of the galaxy halo occupation at a given time, where the satellite distribution is mainly constrained by galaxy clustering.

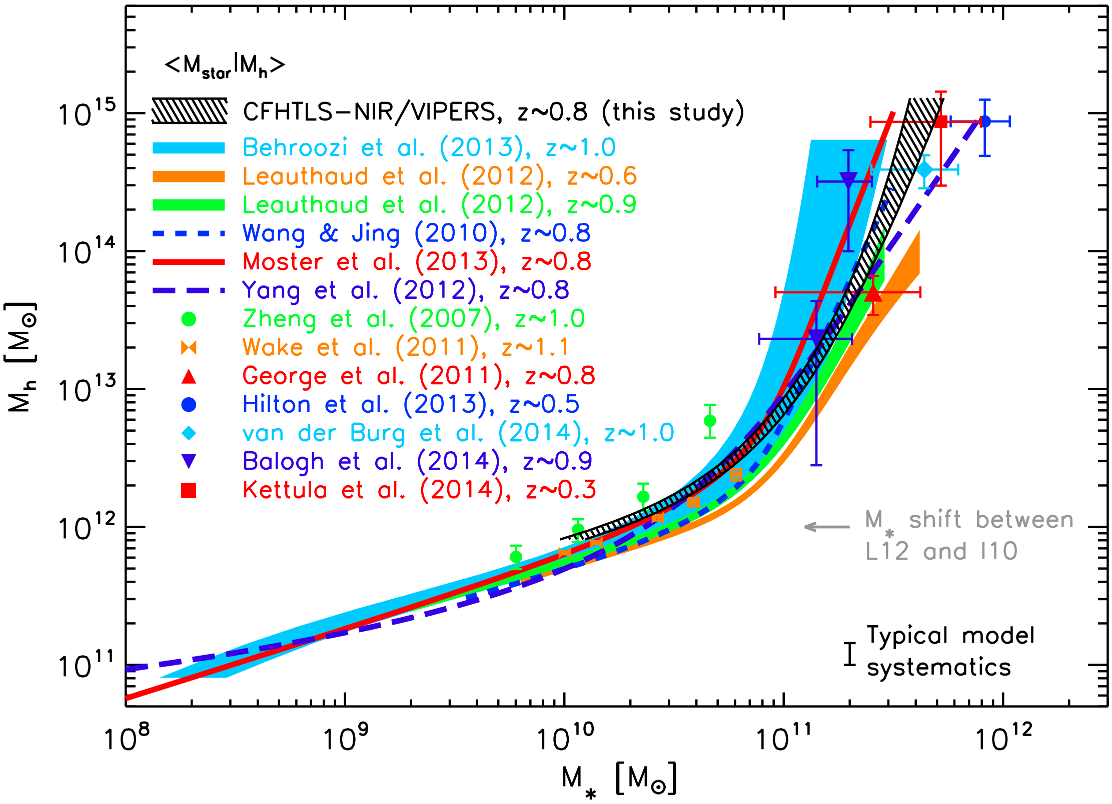


Figure 10. The best-fitting $M_\star - M_\mathrm{h}$ relationship for central galaxies, shown in the black shaded area (total-error based 68 per cent confidence limits), compared with a number of results from the literature at similar redshifts. The results shown here represent the mean stellar mass at fixed halo mass or halo-mass proxy (X-ray temperature or satellite kinematics), $\langle M_\star | M_\mathrm{h}\rangle$, but plotted M_h as function of $M_\star$ to ease the comparison with the literature. We perform appropriate halo mass conversions and IMF stellar mass corrections when required. The length of the grey arrow in the bottom-right corner shows the shift (∼0.2 dex) measured from the direct comparison between stellar masses used in Leauthaud et al. (2012) and George et al. (2011), compared to those in Ilbert et al. (2010) which were estimated in a similar way to this study. The error bar on the bottom-right corner indicates the typical systematic uncertainty arising from the model.

The results from Leauthaud et al. (2012) in COSMOS are shown in brown and green at redshifts $z \sim 0.6$ and ∼0.9, respectively. We observe a small discrepancy which, compared to our results, is unlikely to be explained by differences in the modelling of the HOD (since the model is essentially identical), nor the sample variance as confidence limits do not overlap. A difference in stellar mass estimates on the other hand is more likely to be at the origin of the discrepancy. To check this hypothesis, we have compared the stellar mass estimates from Ilbert et al. (2010), which were measured in a similar way to this study, with those used in Leauthaud et al. (2012) with the method described in Bundy et al. (2006). We measured an offset of ∼0.2 dex, illustrated in Fig. 10 as the grey arrow. Part of the difference seems to be explained by the separate choice for the dust extinction law made in each study (which may typically cause a ∼0.14 dex offset, see Section 2.6). However, we note that in both cases the same IMF and set of SPS models were used, which leaves us without a complete understanding of the difference.

The results by Wang & Jing (2010) are shown as the blue short-dashed line. Their model is based on a HOD modelling of the stellar mass function and real-space galaxy clustering where, as in Behroozi et al. (2013b), the treatment for satellites is not based on the distribution of subhaloes in the host halo but on the $M_\star - M_\mathrm{h}$ relationship at the time of infall.

Moster, Naab & White (2013), shown as the red dot–dashed line, also used abundance matching and provided a redshift-dependent parametrization of the central $M_\star - M_\mathrm{h}$ relationship that we have calculated at $z = 0.8$. As above, the satellites are matched to their haloes at the epoch of merging. Their relation is in good agreement with ours at intermediate mass, however, it shows a steeper dependence on stellar masses at higher mass.

The green dots with error bars are from the HOD modelling results of Zheng et al. (2007), based on real-space clustering and number density measurements. Here, we show their results for DEEP2, a deep spectroscopic survey with high density $z = 1$ galaxies. Without deep NIR data, the authors have computed mean approximate stellar masses for galaxy samples selected in bins of luminosity. This source of uncertainty is not shown on the plot, however, one may expect a large scatter and potential biases due to this conversion.

The orange bow-ties with error bars represent the results[11] by Wake et al. (2011) in the NEWFIRM Medium Band Survey at redshift $z \sim 1.1$, from the combination of NIR-selected galaxy clustering and number density measurements. Their results are in good agreement with ours.

The five next results were produced using galaxy cluster samples associated with their brightest cluster galaxies (BCG). George et al. (2011) built up a catalogue of central versus satellite galaxies in COSMOS, matched to an X-ray detected group/cluster sample with robust halo masses from weak lensing (Leauthaud et al. 2009). From their catalogue, we have computed the mean of stellar mass and halo mass values for clusters in the range $0.5 < z < 1$, shown as the single red triangle (the error bars show the standard deviation

[11] Here, we use updated results compared to the original publication, estimated with Bruzual & Charlot (2003) templates and with rectified h-scaling (Wake, private communication)

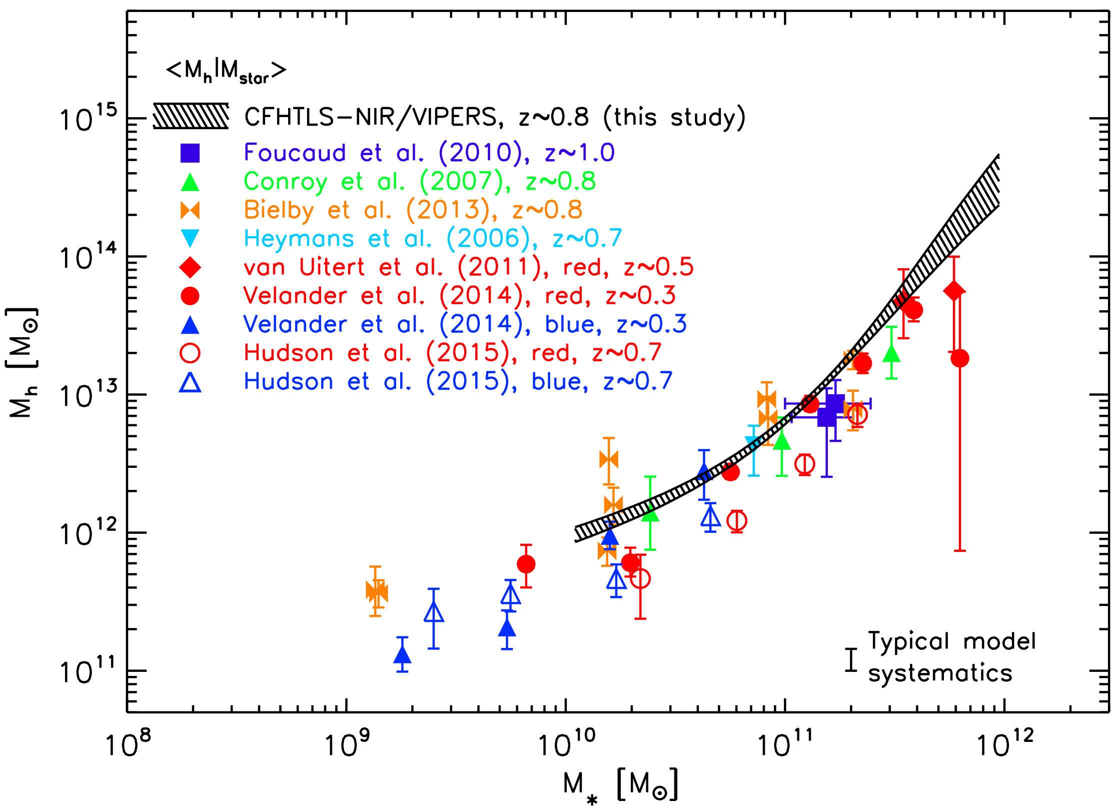


Figure 11. The best-fitting $M_\star - M_\mathrm{h}$ relationship for central galaxies, shown in the black shaded area (total-error based 68 per cent confidence limits), compared with a number of results from the literature at similar redshifts. Unlike in Fig. 10, the results shown here represent the mean *halo mass* at fixed stellar mass $\langle M_\mathrm{h}|M_\star\rangle$. We perform appropriate halo mass conversions and IMF stellar mass corrections when required. The relatively low halo masses found by Hudson et al. (2015) is linked to a different treatment of the satellite subhalo contribution to the lensing signal at small scale (see text for details).

in halo and stellar masses). As they used identical stellar masses to Leauthaud et al. (2012), we also expect a systematic difference in stellar masses compared with our estimates.

From Sunyaev–Zel'dovich detected clusters using the Atacama Cosmology Telescope, Hilton et al. (2013) presented the measurements of the galaxy properties between $0.27 < z < 1.07$. Member galaxies were identified from high-density spectroscopic observations, and stellar masses were measured from *Spitzer* IRAC1-2 mid-infrared (MIR) fluxes. Halo masses were estimated from satellite kinematics. Here, we show the mean halo mass versus mean BCG stellar mass, represented by the single blue dot with errors bars (standard deviations of both masses). Their results appear to be in good agreement with our $M_\star - M_\mathrm{h}$ relationship, although our constraints on such high-mass clusters are extrapolated from the few clusters more massive than $4-5 \times 10^{14}\,\mathrm{M}_\odot$ expected in our sample.

We show as a single light blue diamond the mean halo mass versus mean BCG stellar mass from van der Burg et al. (2014) in the GCLASS/SpARCS cluster sample at $z \sim 1$. Galaxy cluster members were identified from intensive spectroscopic observations, and halo masses were estimated from satellite kinematics. We note that stellar masses were measured from a similar combination of data, redshift range and volume size as ours, however, the methodology used to link halo mass to galaxy stellar masses was rather different. Thus, the agreement with our high-mass $M_\star - M_\mathrm{h}$ relationship within the sample variance is quite remarkable.

Results from Balogh et al. (2014) are shown as the downward purple triangles. Halo mass measurements were made using satellite kinematics for a sample of 11 groups/clusters in the COSMOS field. We show the mean and standard deviation of their measurements split into two halo mass bins (the 11 groups are split into 5 and 6 groups below and above $M_\mathrm{h} = 9 \times 10^{13}\,\mathrm{M}_\odot$, respectively). Although their results suffer from large sample variance, they are in broad agreement with our results and with the rest of the literature.

Finally, the single red square with error bars shows the mean of halo mass measurements from a weak lensing analysis of X-ray selected clusters in the CFHTLenS by Kettula et al. (2014), versus the mean stellar mass of associated BCGs (Mirkazemi et al. submitted). We have re-measured stellar masses of those BCGs in a consistent way to this study (with the exception of missing NIR data for most of the BCGs, which may increase the scatter in stellar mass). Despite the lower redshift range, the identical photometry and lensing catalogue makes the comparison relevant to our results, where the expected difference should arise solely from redshift evolution, although the large statistical uncertainties prevent us from drawing strong conclusions.

5.3.2 $\langle M_\mathrm{h}|M_\star\rangle$ results

We compare the results for $\langle M_\mathrm{h}|M_\star\rangle$ in Fig. 11. To express the mean halo mass at fixed stellar mass $\langle M_\mathrm{h}|M_\star\rangle$ from our results, we derive it from the mean stellar mass at fixed halo mass $\langle M_\star|M_\mathrm{h}\rangle$ using the Bayes theorem relating conditional probability distributions:

$$P(M_\mathrm{h}|M_\star) \propto P(M_\star|M_\mathrm{h}) \times P(M_\mathrm{h})\,. \tag{17}$$

We can then compute $\langle M_{\rm h}|M_\star\rangle$ as the expectation value of $P(M_{\rm h}|M_\star)$:

$$\langle M_{\rm h}|M_\star\rangle = \frac{\int P(M_\star|M_{\rm h})\,P(M_{\rm h})M_{\rm h}{\rm d}M_{\rm h}}{\int P(M_\star|M_{\rm h})\,P(M_{\rm h}){\rm d}M_{\rm h}} \quad (18)$$

with

$$P(M_\star|M_{\rm h}) = \frac{{\rm d}\langle N_{\rm cen}(M_{\rm h}|M_\star)\rangle}{{\rm d}M_\star}\,, \quad (19)$$

the distribution of central galaxies given a halo mass, and

$$P(M_{\rm h}) = \frac{{\rm d}n}{{\rm d}M_{\rm h}}\,, \quad (20)$$

the halo mass function.

We show the results of Foucaud et al. (2010) at $z \sim 1$ from clustering measurements in the UKIDSS-UDS field as the blue squares with error bars. The UKIDSS-UDS field is a small patch of ∼1 deg^2 with deep NIR and optical data. They have converted their clustering amplitude measured in bins of stellar mass into halo masses, using the analytical galaxy-bias halo-mass relationship from Mo & White (1996). As they do not use any constraints from galaxy number density, their error bars are dominated by sample variance and uncertainties on the projected galaxy clustering.

Green upward triangles represent the results by Conroy et al. (2007). Halo masses were derived from satellite kinematics using spectroscopic measurements from the DEEP2 survey. Since the authors have selected their samples based on bins of stellar masses, we can compare their results with our $\langle M_{\rm h}|M_\star\rangle$ $M_\star$–$M_{\rm h}$ relationship. The agreement is found to be good.

Results from clustering measurements in the CFHTLS-DEEP/WIRDS fields by Bielby et al. (2014) are displayed by the brown bow-ties with error bars. We select all mass bin results in the range $0.5 < z < 1$. Although the total field of view is small (∼2.4 deg^2), the combination of four independent fields allowed them to reduce the cosmic variance. As in Foucaud et al., they used an analytical prescription based on the large-scale clustering amplitude to estimate halo masses per bin of stellar mass, so that their results should be compared to our $\langle M_{\rm h}|M_\star\rangle$ results. The two points well above the other results correspond to the measurements at $z \sim 0.7$ and seem to disagree with our constraints and the rest of the literature. The authors claim to have observed an unusually high clustering signal at those redshifts, potentially explained by cosmic variance effects.

Results by Heymans et al. (2006) in the COMBO-17/GEMS field are shown as the downward light-blue triangle with error bars. Here, we have picked their unique measurement at $z > 0.5$. Halo masses were measured using weak lensing with galaxy shapes from the *Hubble Space Telescope* observations.

We show as red diamonds the results for $z \sim 0.5$ red galaxies by van Uitert et al. (2011) in the Red Sequence Cluster Survey 2, a medium-deep CFHT-MegaCam survey in three bands (*gri*) which overlaps 300 deg^2 of the SDSS. The authors have measured the galaxy–galaxy lensing signal for SDSS lens galaxies with a spectroscopic redshift using background source galaxies from the RCS2 survey. Here, the large area permits a high signal-to-noise measurement for very massive galaxies from lensing only. Their results are consistent with ours as this mass bin ($>3 \times 10^{11}$ M$_\odot$) is dominated by red galaxies.

We compare our results with those from Velander et al. (2014) at $z \sim 0.3$, shown as filled symbols (red dots and blue triangles for red and blue galaxies, respectively), and those from Hudson et al. (2015) at $z \sim 0.7$, shown as empty symbols (red dots and blue triangles for red and blue galaxies, respectively). In both studies, halo mass measurements were obtained from galaxy–galaxy lensing measured using the CFHTLenS lensing catalogue and stellar masses computed in a similar way to this study, with the exception that, in both cases, no NIR data were available at the time. This mostly affects the stellar mass estimates of Hudson et al. at $z \sim 0.7$ which, unlike Velander et al. at $z \sim 0.3$, do not benefit from the leverage of the CFHTLS z band. We expect the $M_\star$–$M_{\rm h}$ relationship of the full galaxy population to lie between those of the red and blue populations, however, the results from Hudson et al. lie below our results for both galaxy populations. The bias caused by the scatter in stellar mass partially explains this difference (by shifting their mean stellar mass to higher values), but not entirely: Hudson et al. account for the contribution of subhaloes around satellites occurring at small scale in the lensing signal, whereas we do not (see Section 4.5[12]). As Velander et al. also accounted for subhaloes in their lensing model, we cannot exclude that the apparent good agreement may result from a redshift evolution going in the opposite direction, and requires further investigation.

5.3.3 *The total SHMR*

In Fig. 12, we show the SHMR as function of halo mass compared with observations from the literature. The black shaded area represents the total SHMR as the sum of the central and satellite contributions. The central SHMR (in dashed line on the figure) is simply derived from the central $M_\star$–$M_{\rm h}$ relationship. The satellite SHMR (in dot–dashed line on the figure) is computed from the sum of satellite stellar masses over the halo occupation function at each halo mass, with a lower integration limit of $M_\star = 10^{10}$ M$_\odot$. The total baryon fraction compared to dark matter in the Universe is assumed to be 0.171 and represented on the figure by the grey shaded area on the top (Dunkley et al. 2009, the width of the line represents the uncertainty).

In green, we display the total SHMR from Leauthaud et al. (2012) measured at $z \sim 0.9$. The procedure to compute the total SHMR is identical to ours, i.e. the integrated stellar masses from the satellite HOD were added to the central stellar mass at each halo mass. The authors adopted a mass threshold of $10^{9.8}$ M$_\odot$, which does not change the integrated stellar mass from satellites by a large amount compared with a cut of $>10^{10}$ M$_\odot$. As shown in Fig. 10, part of the vertical shift is explained by the systematic difference in stellar mass estimates.

We show in light blue the central SHMR from Behroozi et al. (2013a). As seen in Fig. 10, the agreement with our central SHMR is good, although their peak is located at a slightly lower halo mass value.

The red triangle shows the results by George et al. (2011) in COSMOS in the redshift range $0.5 < z < 1$. The point represents the mean total stellar mass divided by the halo mass versus the halo mass, and the error bars the standard deviation in each direction. Here, we computed the total stellar mass as the sum of the central galaxy stellar mass plus the stellar masses of associated group members with $M_\star > 10^{10}$. As they used the stellar masses of Leauthaud et al. the agreement is consistently good with their results, however shifted compared to ours.

The single blue dot with error bars marks the mean and standard deviation of estimates by Hilton et al. (2013). Here, the total cluster

[12] This point is also investigated in detail in Appendix D of Hudson et al. (2015).

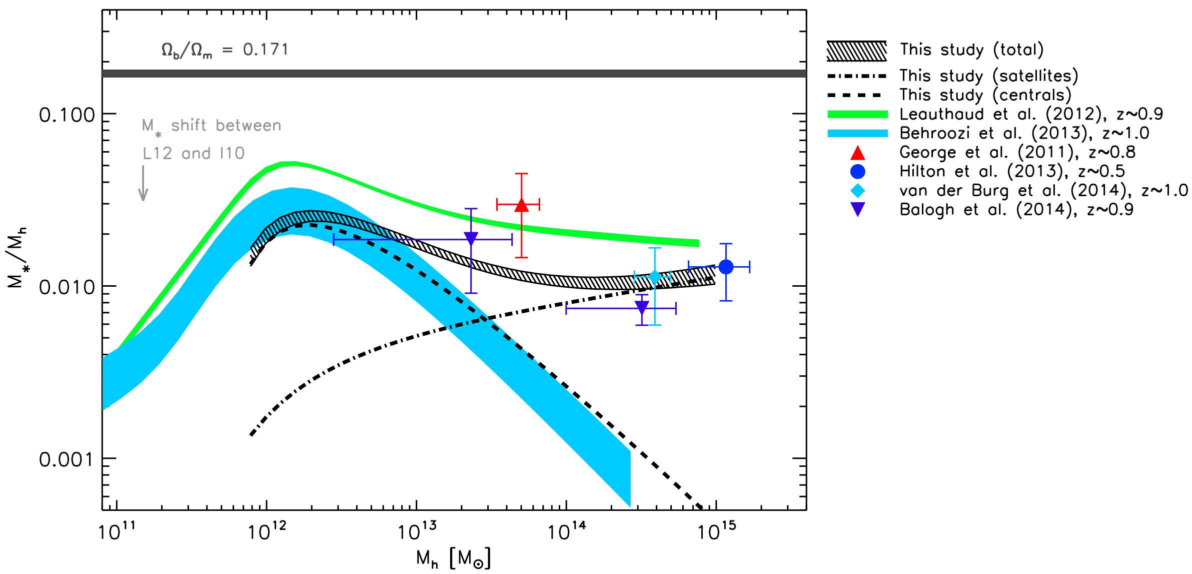


Figure 12. SHMR as function of halo mass compared with observations from the literature. Our best-fitting result for total (central plus satellites) SHMR is shown as the black shaded area. The black dashed line represents the best-fitting central relationship, whereas the dot–dashed line is for the integrated stellar-mass satellite contribution. For Behroozi, Wechsler & Conroy (2013a), only the central SHMR was published and we display it here for comparison with our central SHMR and as an illustration of typical stellar mass systematics. The length of the grey arrow represents the shift to apply to Leauthaud et al. (2012) and George et al. (2011) to reconcile their results with ours, based on the stellar mass comparison with Ilbert et al. (2010).

stellar mass is measured from the background-subtracted sum of galaxy IRAC fluxes within R_{500} from the BCG. Based on the stellar mass completeness computed by Ilbert et al. (2010), an IRAC AB magnitude cut of 24 gives a complete passive galaxy sample down to $M_\star = 10^9\,\mathrm{M}_\odot$ at $z \sim 0.5$. With an IRAC completeness AB magnitude limit of 22.6, it is therefore safe to assume that Hilton et al. are complete above $10^{10}\,\mathrm{M}_\odot$ at $z \sim 0.5$, which matches our sample. We then conclude that their measurements are in good agreement with our results.

Results from van der Burg et al. (2014) are shown as the single light-blue diamond, representing the mean SHMR versus halo mass with its standard deviation. Total stellar masses are computed as the sum of the BCG stellar mass and the stellar mass from galaxy members spectroscopically identified and corrected for TSR. The authors have checked that for $>10^{10}\,\mathrm{M}_\odot$ galaxies, which contribute the most to the total SHMR (see their Fig. 2), the spectroscopic success rate reaches 90 per cent. We note that the median stellar mass completeness $\sim10^{10.16}\,\mathrm{M}_\odot$ is slightly higher than ours (limited by their K_s-band data), however the contribution of satellites compared to a mass limit of $10^{10}\,\mathrm{M}_\odot$ will not significantly change the total SHMR and their measurements can be fairly compared to our results, and we observe an excellent agreement. Interestingly, the authors conclude that when comparing with the literature, no redshift evolution in the total SHMR at high mass is found below $z \sim 1$ and the comparison with our results ($z \sim 0.8$) and those from Hilton et al. ($z \sim 0.5$) confirm their findings.

The two purple downward triangles represent the results from Balogh et al. (2014) in the GEEC2 survey in COSMOS. Here, we show the mean and standard deviation of the SHMR versus halo mass in two halo mass bins. Galaxy members are identified from the spectroscopic redshift when available or using the PDF-weighted photometric redshift computed from the 30-band COSMOS photometric catalogue (Ilbert et al. 2009). The spectroscopic (photometric) sample is complete for group members with $M_\star > 10^{10.3}\,\mathrm{M}_\odot$ ($M_\star > 10^9\,\mathrm{M}_\odot$). Again, since most of the contribution to the total SHMR originates from $M_\star > 10^{10}\,\mathrm{M}_\odot$ galaxies, the comparison with our results is fair. We note a slightly lower value at high mass, and good agreement within the error bars at the group-scale halo mass.

The value of the central SHMR peak may also be compared to that of Coupon et al. (2012) computed from a clustering and galaxy number density analysis of the CFHTLS-Wide. In their study, the authors have measured the evolution of the SHMR peak as function of redshift and have found a lower value compared to ours ($1.1 \times 10^{12}\,\mathrm{M}_\odot$ at redshift $z \sim 0.7$). The difference may not be fully explained by cosmic variance, first because our field significantly overlaps with the full CFHTLS and secondly because the difference is larger than our error bars. In fact, due to their selection in the optical ($i < 22.5$), the SHMR peak above $z = 0.6$ is much less constrained than for our $K_s < 22$ sample, and their peak location suffers from higher uncertainties than in this study, not properly accounted for in their published error bars.

In Fig. 13, we compare our results with a number of semi-analytic predictions from the Millennium simulation (Springel, Frenk & White 2006). In brief, semi-analytic models are anchored to the dark matter halo merger trees provided by N-body simulations, in which empirical recipes of physical processes drive the evolution of galaxies. The fine-tuning of those different processes aim at reproducing the observed galaxy statistical properties. In each case, to derive the total SHMR we compute the sum of the central galaxy stellar mass and the integrated stellar masses of satellites with $M_\star > 10^{10}\,\mathrm{M}_\odot$ to match our sample mass completeness limit. The central SHMR is represented as a dashed line and the shaded area represents the total SHMR with 15 and 85 per cent percentiles. All quantities were computed at redshift $z = 0.8$. The model of Bower et al. (2006) is shown in red (top left), the model of De Lucia & Blaizot (2007) in orange (top right) and the model of Guo et al. (2011) – a modified version of the latter – in green (bottom right). In both De Lucia & Blaizot and Guo et al. models, the contribution from satellites to the total SHMR is significantly below the observations. Despite a different treatment of satellite galaxies and the efficiency of stellar feedback in the latter model, compared to the former, those

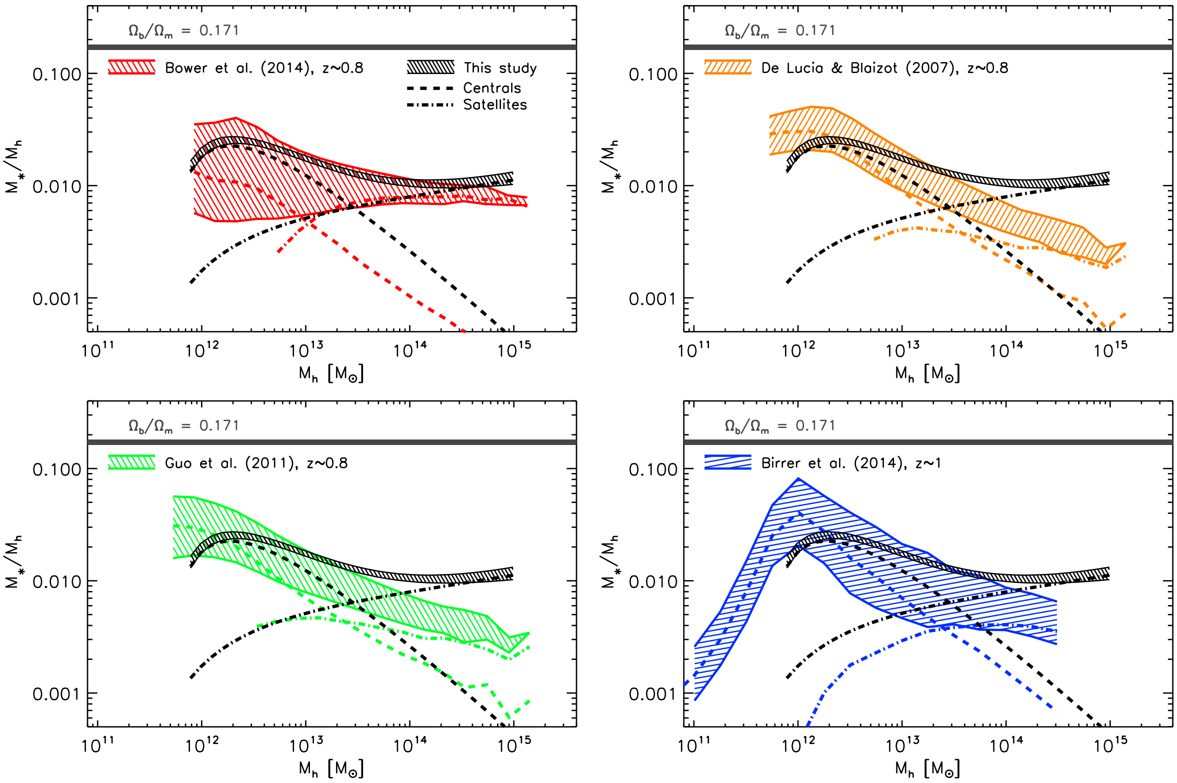


Figure 13. SHMR as function of halo mass compared with simulations from the literature. We compare our total, central and satellite SHMR results with three studies based on semi-analytic models applied to the Millennium simulation (top-row and bottom-left panels) and one study (bottom-right panel) based on the 'gas-regulator' analytical model. For each model, we also display the corresponding central (dashed line) and satellite (dot–dashed line) SHMR.

changes do not show up here. The discrepancy with our results could not arise from a limitation caused by the simulation resolution, as we imposed a cut of $M_\star > 10^{10}$ $\mathrm{M}_\odot$ to match our observations. The model of Bower et al. better reproduces the observed satellite SHMR, however it underestimates the central SHMR and features a significant scatter in the $M_\star - M_\mathrm{h}$ relationship.

We also show the results from the analytical model proposed by Birrer et al. (2014) in blue (bottom right). Their model is an application of the gas-regulator model (Lilly et al. 2013), in which the star formation efficiency is driven by the amount of available gas in the reservoir. In its simplest form, the model describes the inflows and outflows of the gas in the reservoir by two adjustable parameters: a star formation efficiency ϵ, and a mass-loading factor λ that represents the outflows, proportional to the SFR. We show their SHMR at $z = 1$ from the model 'C'.

6 DISCUSSION AND CONCLUSIONS

Using a unique combination of deep optical/NIR data and large area, we have combined galaxy clustering, lensing and galaxy abundance, to put constraints on the galaxy occupation function in the range $0.5 < z < 1$ and to link galaxy properties to dark matter halo masses. Our main result is an accurate measurement of the central galaxy $M_\star - M_\mathrm{h}$ relationship at $z \sim 0.8$ ranging from halo masses at the peak of the SHMR up to the galaxy cluster mass regime. We also provide separate measurements of the SHMR for central and satellite galaxies.

We have shown that the statistical errors (computed using a jackknife estimator) were smaller than systematic errors in the stellar mass measurements caused by uncertainties in the assumed cosmology, dust modelling and photometric calibration. Due to the relatively small amount of statistical uncertainties, the low- to intermediate-mass regime of the stellar mass function is most affected by systematic errors: a factor of $\sim$8 was found between statistical errors and total errors, increasing the error bars of parameters controlling the shape of the $M_\star - M_\mathrm{h}$ relationship by approximately the same amount (see Table 4). Conversely, clustering and lensing measurements feature relatively higher statistical uncertainties and only a factor of $\sim$2 increase in error of the HOD parameters describing the satellite population is observed compared to statistical errors. By probing such a large volume, nearly 0.1 Gpc^3, this study brings unprecedented constraints on the $M_\star - M_\mathrm{h}$ relationship from statistical methods in the cluster mass regime at those redshifts. As shown in Fig. 10, our results make the link between statistical methods based on HOD applied to deep, small-volume surveys, with direct measurements of massive clusters from large-scale surveys.

For central galaxies, we have shown that when properly accounting for halo mass definition, choice of the IMF and the scatter between $M_\star$ and M_h, there is general agreement among results from the literature. We find that stellar mass estimates are the main source of uncertainty, as reflected by the light-blue shaded area from Behroozi et al. (2013b) in Fig. 10, or the stellar mass shift measured between Leauthaud et al. (2012) and Ilbert et al. (2010). We stress, however, that if stellar mass differences may induce a global shift (for instance caused by a separate choice for the IMF), it may also translate into a mass-dependent shift in the more general case (e.g. between two sets of SPS models): hence applying a constant shift may not necessarily reconcile two measurements.

In Fig. 11, stellar mass systematics do not seem to explain all of the observed differences with some results from the literature for which the stellar mass was measured in a similar way to this study. To measure the impact of some of the assumptions made in our model, we have compiled a list of potential systematics propagated through the halo mass and satellite normalization best-fitting parameters. We quote an estimate of 50 per cent error in M_1 and 20 per cent error in B_{sat}, respectively.

For satellite galaxies, the combination of lensing and clustering in this work represents a significant improvement over studies using only the stellar mass function. In Fig. 12, we have shown the measured total SHMR as function of halo mass, compared with a number of results from observations and simulations in the literature. Starting from group-size haloes up to the most massive clusters, we find that the total SHMR is gradually dominated by the contribution from satellites.

Clearly, most SAMs tend to underestimate the total amount of stellar mass produced in medium- to high-mass satellites ($10^{10} < M_\star/\mathrm{M}_\odot < 10^{11}$) at $z \sim 1$ compared to observations. This would suggest that, in SAMs, the bulk of star formation occurs in low-mass galaxies, but is quenched or suppressed at higher mass. Possible explanations for this include either a too strong quenching of haloes in the mass regime $10^{10} < M_\star/\mathrm{M}_\odot < 10^{11}$ (e.g. the work by Henriques et al. 2012, who argue that the gas could be later reincorporated into the haloes), or that low-mass subhaloes are too numerous and would 'catch' the gas in detriment of high-mass subhaloes. It is interesting to link this feature to the overabundance of low-mass galaxies found in numerical simulations compared to observations (see e.g. Guo et al. 2011; Weinmann et al. 2012; De Lucia, Muzzin & Weinmann 2014). In this context, Schive, Chiueh & Broadhurst (2014) recently proposed that cold dark matter could behave as a coherent wave and have shown using *N*-body simulations that this would suppress a large amount of small-mass haloes.

Finally, we can summarize our findings as follows:

(i) the HOD model accurately reproduces the four observables within the statistical error bars in all mass bins over three orders of magnitudes in halo mass and two orders of magnitudes in stellar mass;

(ii) our $M_\star-M_\mathrm{h}$ relationship shows generally good agreement with the literature measurements at $z \sim 0.8$ and we have shown that, when modelling differences are properly accounted for, we are able to make a fair comparison of a number of results derived using independent techniques;

(iii) the systematic errors affecting our measurements were propagated through the whole fitting process. For the parameters describing the $M_\star-M_\mathrm{h}$ relationship, we find that including systematic errors leads to a factor of 8 increase in error bars, and for the parameters describing the satellite HOD a factor of 2 increase in error bars, compared to statistical error bars;

(iv) the sum of systematic errors from the halo model and our model assumptions may be as high (but likely overestimated) as 50 per cent in halo mass and 20 per cent in the satellite number normalization;

(v) the central galaxy SHMR peaks at $M_\mathrm{h} = 1.9 \times 10^{12}\,\mathrm{M}_\odot$, a value slightly larger than the clustering results from the full CFHTLS from Coupon et al. (2012),

(vi) the total (central plus satellites) SHMR is dominated by the satellite contribution in the most massive haloes, in apparent contradiction with SAMs in the Millennium simulation.

We have demonstrated the power of associating a large and deep area with a combination of independent observables to constrain the galaxy–halo relationship with unprecedented accuracy up to $z = 1$. The potential of these data will undoubtedly allow us to extend this analysis to galaxies split by type in future work.

Additionally, studying the evolution in redshift of the SHMR above $z = 1$ is one of the greatest challenge in the near future. If abundance matching already probes the central galaxy–halo relationship up to high redshift, clustering and lensing are necessary to put constraints on the satellite HOD and break some of the degeneracies. Large-scale clustering measurements require wide-field imaging, whereas high-redshift lensing techniques are yet to be improved, but on-going projects such as Hyper Suprime Cam (HSC), Dark Energy Survey (DES) or COSMOS/SPLASH (Spitzer Large Area Survey with Hyper-Suprime-Cam), which will increase by orders of magnitude the currently available data, represent the ideal data sets to address those issues.

ACKNOWLEDGEMENTS

This work is primarily based on observations obtained with WIRCam, a joint project of CFHT, Taiwan, Korea, Canada, France, at the Canada–France–Hawaii Telescope (CFHT) which is operated by the National Research Council (NRC) of Canada, the Institut National des Sciences de l'Univers of the Centre National de la Recherche Scientifique of France and the University of Hawaii. The WIRCAM images have been collected during several semesters and from different programs. We thank G. Morrison, J. Willis and K. Thanjavur for leading some of these programmes as PIs, and a special thanks to the canadian agency who always highly ranked these proposals. We thank the Terapix team for the reduction of all the WIRCAM images and the preparation of the catalogues matching the T0007 CFHTLS data release. The CFHTLenS project is based on observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/IRFU, at the CFHT. This research used the facilities of the Canadian Astronomy Data Centre operated by the National Research Council of Canada with the support of the Canadian Space Agency. We thank the CFHT staff for successfully conducting the CFHTLS observations and in particular J.-C. Cuillandre and E. Magnier for the continuous improvement of the instrument calibration and the Elixir detrended data that we used. We also thank TERAPIX for the quality assessment and validation of individual exposures during the CFHTLS data acquisition period, and E. Bertin for developing some of the software used in this study. CFHTLenS data processing was made possible thanks to significant computing support from the NSERC Research Tools and Instruments grant program, and to HPC specialist O. Toader. Part of the numerical computations were done on the Sciama High Performance Compute (HPC) cluster which is supported by the ICG, SEPNet and the University of Portsmouth.

We thank M. Cousin and S. de la Torre for useful discussions related to this work. We thank S. Birrer and D. Wake for providing us with their measurements. JC acknowledges the support from the laboratoire de Marseille during his stay in 2014 September. The early stages of the CFHTLenS project was made possible thanks to the support of the European Commission's Marie Curie Research Training Network DUEL (MRTN-CT-2006-036133) which directly supported six members of the CFHTLenS team (LF, HH, PS, BR, CB, MV) between 2007 and 2011 in addition to providing travel support and expenses for team meetings. BR +CH and HHo+ acknowledge support from the European Research Council under EC FP7 grant numbers 240672 (BR), 240185 (CH) and 279396 (HHo). HHi is supported by the DFG Emmy Noether grant Hi 1495/2-1. AF acknowledges support by INAF through VIPERS grants PRIN 2008 and PRIN 2010. MJH acknowledges support from the Â Natural Sciences and Engineering Research Council of Canada (NSERC).

Author Contributions: All authors contributed to the development and writing of this paper. The authorship list reflects the lead authors of this paper (JC, SA, LVW, TM, OI, EVU) followed by

two alphabetical groups. The first alphabetical group includes key contributors to the science analysis and interpretation in this paper, the founding core team and those whose long-term significant effort produced the final CFHTLenS data product, and the VIPERS collaboration leaders. The second group covers authors in alphabetical order who made a significant contribution to either the projects, this paper, or both. The CFHTLenS collaboration was co-led by CH and LVW.

REFERENCES

Ahn C. P. et al., 2014, ApJS, 211, 17
Arnouts S., Cristiani S., Moscardini L., Matarrese S., Lucchin F., Fontana A., Giallongo E., 1999, MNRAS, 310, 540
Arnouts S. et al., 2013, A&A, 558, A67
Balogh M. L. et al., 2014, MNRAS, 443, 2679
Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15
Bartelmann M., Schneider P., 2001, Phys. Rep., 340, 291
Behroozi P. S., Conroy C., Wechsler R. H., 2010, ApJ, 717, 379
Behroozi P. S., Wechsler R. H., Conroy C., 2013a, ApJ, 770, 57
Behroozi P. S., Marchesini D., Wechsler R. H., Muzzin A., Papovich C., Stefanon M., 2013b, ApJ, 777, L10
Berlind A. A., Weinberg D. H., 2002, ApJ, 575, 587
Berlind A. A. et al., 2003, ApJ, 593, 1
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Bielby R. et al., 2012, A&A, 545, A23
Bielby R. M. et al., 2014, A&A, 568, A24
Birrer S., Lilly S., Amara A., Paranjape A., Refregier A., 2014, ApJ, 793, 12
Boulade O. et al., 2003, in Iye M., Moorwood A. F. M., eds, Proc. SPIE Conf. Ser. Vol. 4841, Instrument Design and Performance for Optical/Infrared Ground-based Telescopes. SPIE, Bellingham, p. 72
Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645
Brainerd T. G., Blandford R. D., Smail I., 1996, ApJ, 466, 623
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Budzynski J. M., Koposov S. E., McCarthy I. G., McGee S. L., Belokurov V., 2012, MNRAS, 423, 104
Bullock J. S., Kolatt T. S., Sigad Y., Somerville R. S., Kravtsov A. V., Klypin A. A., Primack J. R., Dekel A., 2001, MNRAS, 321, 559
Bundy K. et al., 2006, ApJ, 651, 120
Cacciato M., van den B., Frank C., More S., Li R., Mo H. J., Yang X., 2009, MNRAS, 394, 929
Cacciato M., van Uitert E., Hoekstra H., 2014, MNRAS, 437, 377
Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, ApJ, 533, 682
Chabrier G., 2003, PASP, 115, 763
Charlot S., Fall S. M., 2000, ApJ, 539, 718
Coil A. L. et al., 2011, ApJ, 741, 8
Conroy C., Wechsler R. H., Kravtsov A. V., 2006, ApJ, 647, 201
Conroy C. et al., 2007, ApJ, 654, 153
Conseil S., Vibert D., Amouts S., Milliard B., Zamojski M., Liebaria A., Guillaume M., 2011, in Evans I. N., Accomazzi A., Mink D. J., Rots A. H., eds, ASP Conf. Ser. Vol. 442, Astronomical Data Analysis Software and Systems XX. Astron. Soc. Pac., San Francisco, p. 107
Cooray A., Sheth R., 2002, Phys. Rep., 372, 1
Coupon J. et al., 2009, A&A, 500, 981
Coupon J. et al., 2012, A&A, 542, A5
Coupon J., Broadhurst T., Umetsu K., 2013, ApJ, 772, 65
Cowie L. L., Songaila A., Hu E. M., Cohen J. G., 1996, AJ, 112, 839
Davidzon I. et al., 2013, A&A, 558, A23
de la Torre S. et al., 2013, A&A, 557, A54
De Lucia G., Blaizot J., 2007, MNRAS, 375, 2
De Lucia G., Muzzin A., Weinmann S., 2014, New Astron Rev., 62, 1
Dunkley J. et al., 2009, ApJ, 701, 1804
Erben T. et al., 2013, MNRAS, 433, 2545
Faber S. M. et al., 2007, ApJ, 665, 265
Foucaud S., Conselice C. J., Hartley W. G., Lane K. P., Bamford S. P., Almaini O., Bundy K., 2010, MNRAS, 406, 147
Fritz A. et al., 2014, A&A, 563, A92
Garilli B. et al., 2008, A&A, 486, 683
Garilli B. et al., 2014, A&A, 562, A23
Gelman A., Rubin D., 1992, Stat. Sci., 7, 457
George M. R. et al., 2011, ApJ, 742, 125
Guo Q., White S., Li C., Boylan-Kolchin M., 2010, MNRAS, 404, 1111
Guo Q. et al., 2011, MNRAS, 413, 101
Guzzo L. et al., 2014, A&A, 566, A108
Henriques B. M. B., White S. D. M., Lemson G., Thomas P. A., Guo Q., Marleau G.-D., Overzier R. A., 2012, MNRAS, 421, 2904
Heymans C. et al., 2006, MNRAS, 371, L60
Heymans C. et al., 2012, MNRAS, 427, 146
Hildebrandt H. et al., 2012, MNRAS, 421, 2355
Hilton M. et al., 2013, MNRAS, 435, 3469
Hinshaw G. et al., 2009, ApJS, 180, 225
Hoaglin D. C., Mosteller F., Tukey J. W., 1983, Understanding Robust and Exploratory Data Analysis. Wiley, New York
Hoekstra H., Yee H. K. C., Gladders M. D., 2004, in Ryder S.D., Pisano D.J., Walker M.A., Freeman K.C., eds, Proc. IAU Symp. 220, Dark Matter in Galaxies. Astron. Soc. Pac., San Francisco, p. 439
Hu W., Kravtsov A. V., 2003, ApJ, 584, 702
Hudelot P. et al., 2012, VizieR Online Data Catalog, 2317, 0
Hudson M. J., Gwyn S. D. J., Dahle H., Kaiser N., 1998, ApJ, 503, 531
Hudson M. J. et al., 2015, MNRAS, 447, 298
Ilbert O. et al., 2006, A&A, 457, 841
Ilbert O. et al., 2009, ApJ, 690, 1236
Ilbert O. et al., 2010, ApJ, 709, 644
Ilbert O. et al., 2013, A&A, 556, A55
Kettula K., S. G., van Uitert E., Hoekstra H., A. F., 2014, MNRAS
Kitching T. D., Miller L., Heymans C. E., van Waerbeke L., Heavens A. F., 2008, MNRAS, 390, 149
Kravtsov A. V., Berlind A. A., Wechsler R. H., Klypin A. A., Gottlöber S., Allgood B., Primack J. R., 2004, ApJ, 609, 35
Kroupa P., 2001, MNRAS, 322, 231
Landy S. D., Szalay A. S., 1993, ApJ, 412, 64
Lawrence A. et al., 2007, MNRAS, 379, 1599
Le Fèvre O. et al., 2005, A&A, 439, 845
Le Fèvre O. et al., 2014, preprint (arXiv:e-prints)
Leauthaud A. et al., 2009, ApJ, 709, 97
Leauthaud A., Tinker J., Behroozi P. S., Busha M. T., Wechsler R. H., 2011, ApJ, 738, 45
Leauthaud A. et al., 2012, ApJ, 744, 159
Lilly S. J., Carollo C. M., Pipino A., Renzini A., Peng Y., 2013, ApJ, 772, 119
Mandelbaum R. et al., 2005a, MNRAS, 361, 1287
Mandelbaum R., Tasitsiomi A., Seljak U., Kravtsov A. V., Wechsler R. H., 2005b, MNRAS, 362, 1451
Mandelbaum R., Seljak U., Kauffmann G., Hirata C. M., Brinkmann J., 2006, MNRAS, 368, 715
Mandelbaum R., Slosar A., Baldauf T., Seljak U., Hirata C. M., Nakajima R., Reyes R., Smith R. E., 2013, MNRAS, 432, 1544
Marchesini D., van Dokkum P. G., Förster Schreiber N. M., Franx M., Labbé I., Wuyts S., 2009, ApJ, 701, 1765
Marinoni C., Hudson M. J., 2002, ApJ, 569, 101
Martin D. C. et al., 2005, ApJ, 619, L1
Martinez-Manso J., Gonzalez A. H., Ashby M. L. N., Stanford S. A., Brodwin M., Holder G. P., Stern D., 2015, MNRAS, 446, 169
Miller L., Kitching T. D., Heymans C., Heavens A. F., van Waerbeke L., 2007, MNRAS, 382, 315
Miller L. et al., 2013, MNRAS, 429, 2858
Mirkazemi M. et al., submitted
Miyatake et al., 2013, preprint (arXiv:1311.1480)
Mo H. J., White S. D. M., 1996, MNRAS, 282, 347
More S., van den Bosch F. C., Cacciato M., Skibba R., Mo H. J., Yang X., 2011, MNRAS, 410, 210

More S., Miyatake H., Mandelbaum R., Takada M., Spergel D., Brownstein J., Schneider D. P., 2014, preprint (arXiv:1407.1856)
Morrissey P. et al., 2005, ApJ, 619, L7
Moster B. P., Somerville R. S., Maulbetsch C., van den Bosch F. C., Macciò A. V., Naab T., Oser L., 2010, ApJ, 710, 903
Moster B. P., Somerville R. S., Newman J. A., Rix H.-W., 2011, ApJ, 731, 113
Moster B. P., Naab T., White S. D. M., 2012, MNRAS, 428, 3121
Moster B. P., Naab T., White S. D. M., 2013, MNRAS, 428, 3121
Muñoz-Cuartas J. C., Macciò A. V., Gottlöber S., Dutton A. A., 2010, MNRAS, 411, 584
Muzzin A., Yee H. K. C., Hall P. B., Ellingson E., Lin H., 2007, ApJ, 659, 1106
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Norberg P., Baugh C. M., Gaztanaga E., Croton D. J., 2009, MNRAS, 396, 19
Okabe N., Smith G. P., Umetsu K., Takada M., Futamase T., 2013, ApJ, 769, L35
Planck Collaboration XVI 2014, A&A, 571, 16
Polletta M. et al., 2007, ApJ, 663, 81
Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 2002, Numerical Recipes in C++: The Art of Scientific Computing. Cambridge Univ. Press, Cambridge
Prevot M. L., Lequeux J., Prevot L., Maurice E., Rocca-Volmerange B., 1984, A&A, 132, 389
Puget P. et al., 2004, in Moorwood A. F. M., Iye M., eds, Proc. SPIE Conf. Ser. Vol. 5492, Ground-based Instrumentation for Astronomy. SPIE, Bellingham, p. 978
Roche N. D., Almaini O., Dunlop J., Ivison R. J., Willott C. J., 2002, MNRAS, 337, 1282
Salpeter E. E., 1955, ApJ, 121, 161
Schive H.-Y., Chiueh T., Broadhurst T., 2014, Nat. Phys., 10, 496
Scoville N. et al., 2007, ApJS, 172, 1
Sheldon E. S. et al., 2004, AJ, 127, 2544
Springel V., Frenk C. S., White S. D. M., 2006, Nature, 440, 1137
Szalay A. S., Connolly A. J., Szokoly G. P., 1999, AJ, 117, 68
Takada M., Jain B., 2003, MNRAS, 344, 857
Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlöber S., 2010, ApJ, 724, 878
Umetsu K., Broadhurst T., Zitrin A., Medezinski E., Hsu L.-Y., 2011, ApJ, 729, 127
Vale A., Ostriker J. P., 2006, MNRAS, 371, 1173
van der Burg R. F. J., Muzzin A., Hoekstra H., Wilson G., Lidman C., Yee H. K. C., 2014, A&A, 561, A79
van Uitert E., Hoekstra H., Velander M., Gilbank D. G., Gladders M. D., Yee H. K. C., 2011, A&A, 534, A14
Velander M. et al., 2014, MNRAS, 437, 2111
Wake D. A. et al., 2011, ApJ, 728, 46
Wang L., Jing Y. P., 2010, MNRAS, 402, 1796
Weinberg N. N., Kamionkowski M., 2002, MNRAS, 337, 1269
Weinmann S. M., Pasquali A., Oppenheimer B. D., Finlator K., Mendel J. T., Crain R. A., Macciò A. V., 2012, MNRAS, 426, 2797
Wraith D., Kilbinger M., Benabed K., Cappé O., Cardoso J.-F., Fort G., Prunet S., Robert C., 2009, Phys. Rev. D, 80, 023507
Yang X., Mo H. J., van den Bosch F. C., 2003, MNRAS, 339, 1057
Yoo J., Tinker J. L., Weinberg D. H., Zheng Z., Katz N., Davé R., 2006, ApJ, 652, 26
Zehavi I. et al., 2005, ApJ, 630, 1
Zehavi I. et al., 2011, ApJ, 736, 59
Zentner A. R., Hearin A. P., van den Bosch F. C., 2014, MNRAS, 443, 3044
Zheng Z., Coil A. L., Zehavi I., 2007, ApJ, 667, 760

APPENDIX A: COMPLETENESS OF THE SAMPLES

In this section, we use the CFHTLS-Deep/WIRDS combined data to test our samples' mass completeness. The CFHTLS-Deep/WIRDS data are over 2 mag deeper in all bands compared to our CFHTLS-Wide/WIRCam data and with accurate photometric redshift and stellar mass estimates computed in a similar fashion to this study. Fig. A1 shows the galaxy distribution in WIRDS as function of stellar mass and redshift corresponding to our selection $K_s < 22$ for the photometric sample (top) and $i < 22.5$ for the spectroscopic sample (bottom).

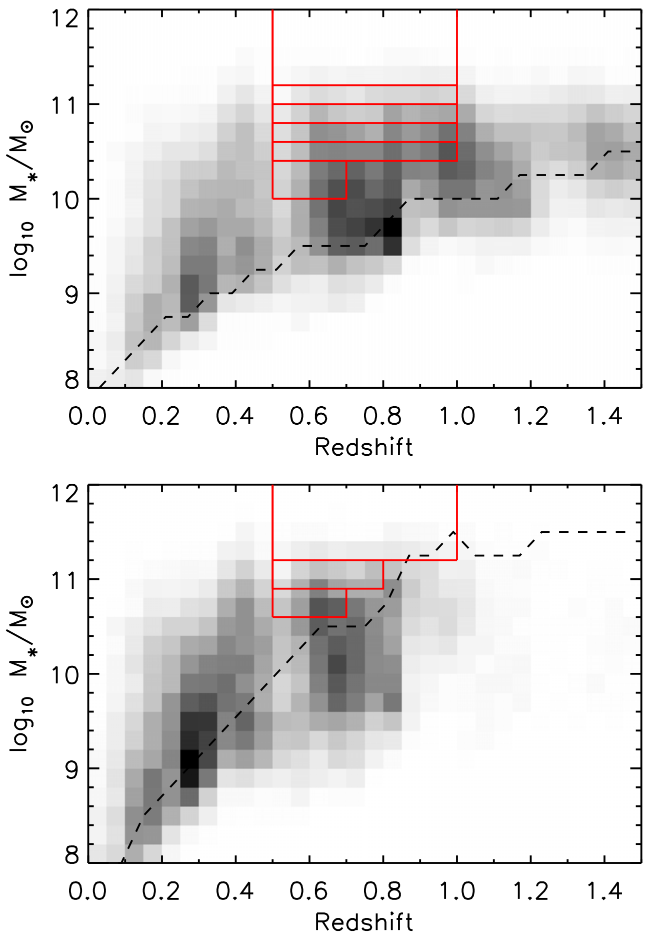


Figure A1. Galaxy distribution as function of stellar mass and redshift in WIRDS. Stellar mass 90 per cent completeness limits of $K_s < 22$ (top) and $i < 22.5$ (bottom) selected samples are represented as the dashed black line and the sample selection as the thick red line.

The density fluctuations seen as function of redshift are due to cosmic variance (the field of view is smaller than 1 deg^2), but we do not expect any significant impact on our completeness assessments. In both panels, we represent the 90 per cent completeness limits as dashed lines, and our samples' selection as red boxes. In the case of the photometric sample, a conservative $z < 0.7$ cut is adopted in the lower mass sample to prevent missing red galaxies caused by the optical incompleteness at the CFHTLS-Wide depth. Overall, these verifications show that all of our samples are complete in mass.

APPENDIX B: DETAILS ON THE DERIVATION OF THE OBSERVABLES

Here, we provide detailed calculations of the four observables used in this study and derived from the HOD model described in Section 4. For the dark matter halo profile and the distribution of satellites, we assume a (Navarro et al. 1997, NFW) profile with the theoretical mass–concentration relation from equation (16) of

Takada & Jain (2003) with $c_0 = 11$ and $\beta = 0.13$, featuring the redshift dependence $(1+z)^{-1}$ (Bullock et al. 2001). All dark matter quantities are derived at the mean redshift of the galaxy sample, computed from the expectation value of the sum of redshift PDFs. All quantities are computed in comoving units ('co'). The clustering and galaxy–galaxy lensing are then converted into physical units ('phys') to match the measurements.

B1 Stellar mass function

The stellar mass function is the integrated HOD over the halo mass function:

$$\phi_{\rm SMF}\left(M_\star^{t_1}, M_\star^{t_2}\right) = \int_0^\infty \left\langle N_{\rm tot}\left(M_{\rm h}|M_\star{}^{t_1}, M_\star{}^{t_2}\right)\right\rangle \frac{{\rm d}n}{{\rm d}M_{\rm h}} {\rm d}M_{\rm h}. \quad \text{(B1)}$$

B2 Galaxy clustering

We describe galaxy clustering using the two-point correlation function, as the sum of the one-halo and two-halo terms:

$$\xi_{\rm gg}(r_{\rm co}) = 1 + \xi_{\rm gg,1}(r_{\rm co}) + \xi_{\rm gg,2}(r_{\rm co}). \quad \text{(B2)}$$

The one-halo term, $\xi_{\rm gg,1}(r_{\rm co})$, expresses the relative contribution of galaxy pairs within the halo $\langle N_{\rm tot}(M_{\rm h})(N_{\rm tot}(M_{\rm h})-1)\rangle/2$ and can be decomposed, assuming Poisson statistics for the satellites, into two terms:

$$\langle N_{\rm cen}N_{\rm sat}\rangle(M_{\rm h}) = \langle N_{\rm cen}(M_{\rm h})\rangle\langle N_{\rm sat}(M_{\rm h})\rangle;$$
$$\langle N_{\rm sat}(N_{\rm sat}-1)\rangle(M_{\rm h})/2 = \langle N_{\rm sat}(M_{\rm h})\rangle^2/2. \quad \text{(B3)}$$

The correlation function for central-satellite pairs is given by

$$1 + \xi_{\rm cs}(r_{\rm co}, z) = \int_{M_{\rm vir}(r)}^\infty {\rm d}M_{\rm h}\, n(M_{\rm h}, z)\frac{\langle N_{\rm cen}\rangle\langle N_{\rm sat}\rangle}{n_{\rm gal}^2/2}\rho_{\rm h}(r_{\rm co}|M_{\rm h}), \quad \text{(B4)}$$

where we assume that the distribution of central-satellite pairs simply follows that of the dark matter halo profile. The lower integration limit $M_{\rm vir}(r_{\rm co})$ accounts for the fact that no halo with a virial mass corresponding to $r_{\rm co}$ would contribute to the correlation function.

For the satellite contribution $\xi_{\rm ss}$, the distribution of satellite pairs is the convolution of the dark matter halo profile with itself, computed here in Fourier space. The satellite power spectrum is

$$P_{\rm ss}(k) = \int_{M_{\rm low}}^{M_{\rm high}} {\rm d}M_{\rm h}\, n(M_{\rm h})\frac{\langle N_{\rm sat}(M_{\rm h})\rangle^2}{n_{\rm gal}^2}|u_h(k|M_{\rm h})|^2, \quad \text{(B5)}$$

where $u_{\rm h}(k|M_{\rm h})$ is the Fourier transform of the dark-matter halo profile $\rho_{\rm h}(r_{\rm co}|M_{\rm h})$. The correlation function $\xi_{\rm ss}$ is then obtained via a Fourier transform.

The one-halo correlation function is the sum of the two contributions,

$$\xi_{\rm gg,1}(r_{\rm co}) = 1 + \xi_{\rm cs}(r_{\rm co}) + \xi_{\rm ss}(r_{\rm co}). \quad \text{(B6)}$$

The two-halo term is computed from the galaxy power spectrum:

$$P_2(k, r_{\rm co}) = P_{\rm m}(k) \times \left[\int_{M_{\rm low}}^{M_{\rm lim}(r_{\rm co})} {\rm d}M_{\rm h} n(M_{\rm h})\frac{\langle N_{\rm tot}\rangle}{n'_{\rm gal}(r_{\rm co})} b_{\rm h}(M_{\rm h}, r_{\rm co})|u_{\rm h}(k|M_{\rm h})|\right]^2, \quad \text{(B7)}$$

where

$$n'_{\rm gal}(r_{\rm co}) = \int_{M_{\rm low}}^{M_{\rm lim}(r_{\rm co})} n(M_{\rm h})\langle N_{\rm tot}\rangle\, {\rm d}M_{\rm h}. \quad \text{(B8)}$$

The upper integration limit $M_{\rm lim}(r_{\rm co})$ accounts for halo exclusion as detailed in Coupon et al. (2012), and references therein.

Finally, the two-halo term $\xi_{\rm gg,2}$ of the galaxy autocorrelation function is the Fourier transform of equation (B7) renormalized to the total number of galaxy pairs:

$$1 + \xi_{\rm gg,2}(r_{\rm co}) = \left[\frac{n'_{\rm gal}(r_{\rm co})}{n_{\rm gal}}\right][1 + \xi_{\rm gg,2}(r_{\rm co})]. \quad \text{(B9)}$$

The projected clustering $w(\theta)$ is derived from the projection of $\xi_{\rm gg}$ on to the estimated redshift distribution from the sum of PDFs, assuming the Limber approximation (see details in Coupon et al. 2012).

The real-space clustering $w_{\rm p}(r_{\rm p,\,co})$ is derived from the projection of the 3D correlation function along the line of sight:

$$w_{\rm p}(r_{\rm p,co}) = 2\int_{r_{\rm p,co}}^\infty r_{\rm co}\,{\rm d}r_{\rm co}\xi_{\rm gg}(r_{\rm co})\,(r_{\rm co}^2 - r_{\rm p,co}^2)^{-1/2}, \quad \text{(B10)}$$

converted into physical units as

$$w_{\rm p,phys} = w_{\rm p,co}/(1+z). \quad \text{(B11)}$$

B3 Galaxy–galaxy lensing

The galaxy–galaxy lensing estimator measures the excess surface density of the projected dark matter halo profile:

$$\Delta\Sigma_{\rm co}(r_{\rm p,co}) = \overline{\Sigma}_{\rm co}(< r_{\rm p,co}) - \overline{\Sigma}_{\rm co}(r_{\rm p,co}), \quad \text{(B12)}$$

where $\overline{\Sigma}_{\rm co}(< r_{\rm p,co})$ is the projected mean surface density within the comoving radius $r_{\rm p,\,co}$ and $\overline{\Sigma}_{\rm co}(r_{\rm p,co})$ the mean surface density at the radius $r_{\rm p,co}$.

To compute the analytical projected dark matter density Σ, we write

$$\Sigma_{\rm co}(r_{\rm p,co}) = \int \rho\left(\sqrt{r_{\rm p,co}^2 + \pi_{\rm co}^2}\right){\rm d}\pi_{\rm co}$$
$$= \overline{\rho}\int \left[1 + \xi_{\rm gm}\left(\sqrt{r_{\rm p,co}^2 + \pi_{\rm co}^2}\right)\right]{\rm d}\pi_{\rm co}, \quad \text{(B13)}$$

where $r_{\rm p,co}$ is the transverse comoving distance, $\pi_{\rm co}$ the line-of-sight comoving distance, $\overline{\rho}$ the mean density of the Universe, so that $\Delta\Sigma_{\rm co}(r_{\rm p,\,co})$ is related to the galaxy-dark matter cross-correlation function $\xi_{\rm gm}$ through

$$\Delta\Sigma_{\rm co}(r_{\rm p,co}) = \overline{\Sigma}_{\rm co}(< r_{\rm p,co}) - \overline{\Sigma}_{\rm co}(r_{\rm p,co})$$
$$= \overline{\rho}\left[\frac{4}{r_{\rm p,co}^2}\int_0^{r_{\rm p,co}}\int_0^{\pi_{\rm max}} r'_{\rm p,co}\xi_{\rm gm}\left(\sqrt{r'^2_{\rm p,co}+\pi_{\rm co}^2}\right){\rm d}\pi_{\rm co}{\rm d}r'_{\rm p,co}\right.$$
$$\left.-2\int_0^{\pi_{\rm max}}\xi_{\rm gm}\left(\sqrt{r_{\rm p,co}^2+\pi_{\rm co}^2}\right){\rm d}\pi_{\rm co}\right]. \quad \text{(B14)}$$

The integration along the line of sight is performed up to the scale $\pi_{\rm max} = 80$ Mpc.

The excess surface density in physical units writes

$$\overline{\Delta\Sigma}_{\rm phys} = \overline{\Delta\Sigma}_{\rm co}\times(1+z_{\rm L})^2, \quad \text{(B15)}$$

where $z_{\rm L}$ is the redshift of the lens galaxy.

As for $\xi_{\rm gg}$, $\xi_{\rm gm}$ can be written as the sum of the one- and two-halo terms:

$$\xi_{\rm gm}(r) = 1 + \xi_{\rm gm,1}(r) + \xi_{\rm gm,2}(r). \quad \text{(B16)}$$

$\xi_{\rm gm,1}(r)$ is itself decomposed into a contribution from the cross-correlation of the central galaxy-dark matter and from that of the

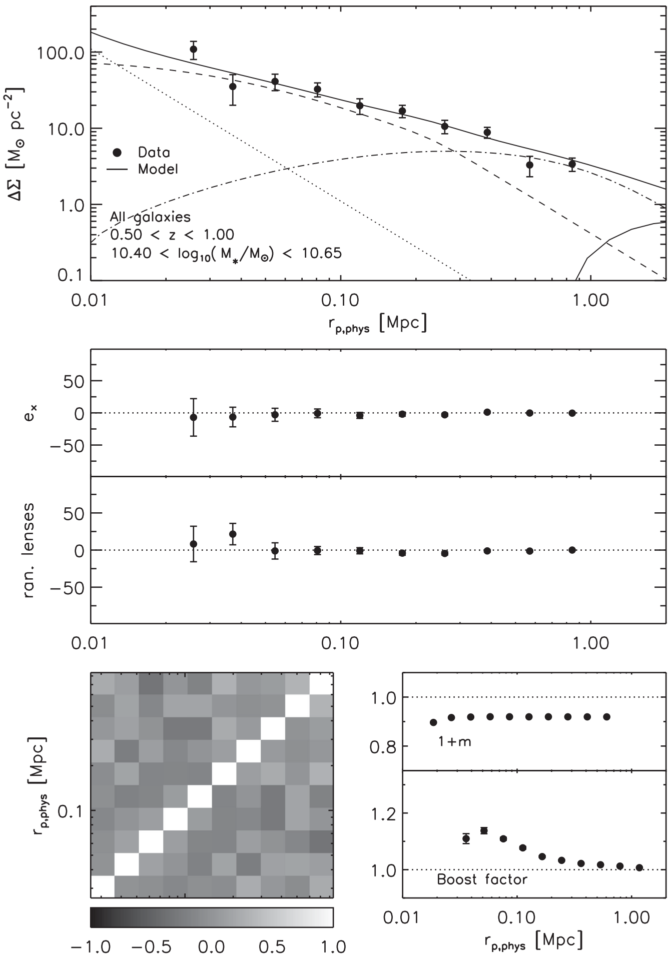


Figure C1. Galaxy–galaxy lensing measurements and systematics checks for the sample $10.40 < \log(M_\star/\mathrm{M}_\odot) < 10.65$. In the top panel, we show the data (dots with error bars) and the model (thick line) split into the stellar term in dotted line, the central term in dashed line, the satellite term in dot–dashed line and the two-halo term in black solid line at bottom-right corner. The lower panels show the systematic tests (rotated-shape signal and random lens positions), calibration factor (multiplicative bias correction and boost factor) and the lower-left corner the correlation coefficients of the correlation matrix from the jackknife estimate.

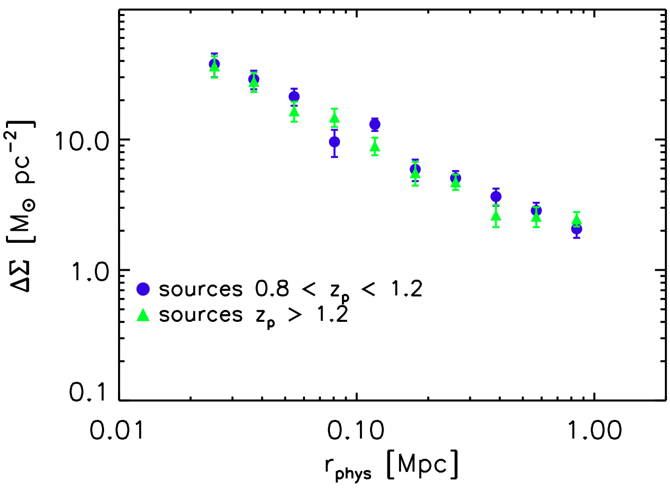


Figure C2. Galaxy–galaxy lensing measurements separating the background sample into $0.8 < z_\mathrm{p} < 1.2$ sources (purple dots) and $z_\mathrm{p} > 1.2$ sources (green triangles), keeping the same lens galaxy foreground sample (low-redshift galaxies with spectroscopic redshifts).

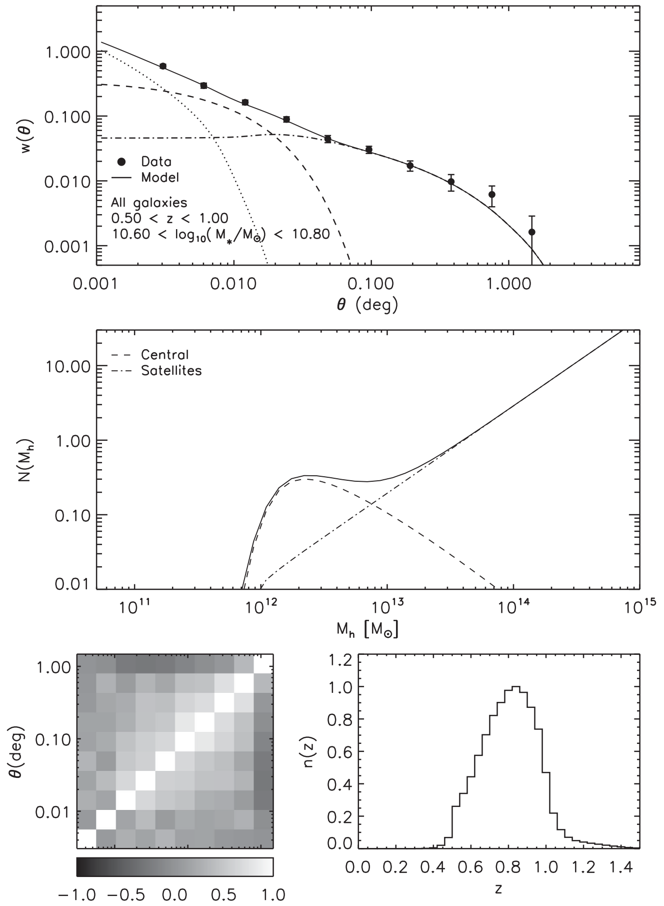


Figure C3. $w(\theta)$ measurements and the corresponding HOD function for the sample $10.60 < \log(M_\star/\mathrm{M}_\odot) < 10.80$. In the top panel, we show the data points with error bars and the best-fitting model: the dotted line represents the central-satellite cross-correlation, the dashed line the satellite–satellite autocorrelation, and the dot–dashed line the central–central autocorrelation (or 2-halo term). The middle panel displays the corresponding HOD, the dashed line shows the central galaxy HOD and the dot–dashed line the satellites' HOD. The lower-right panel shows the corresponding redshift distribution constructed from the sum of individual PDFs. The lower-left panel shows the correlation coefficients of the covariance matrix from the jackknife estimate.

satellite-dark matter, both assuming an NFW profile. We write the former as

$$1 + \xi_{\mathrm{gm,cen}}(r, z) = \int_{M_{\mathrm{vir}}(r)}^{M_{\mathrm{high}}} \mathrm{d}M_\mathrm{h}\, n(M_\mathrm{h}, z) \frac{\langle N_\mathrm{cen}\rangle}{n_\mathrm{gal}} \rho_\mathrm{h}(r|M_\mathrm{h}) \frac{M_\mathrm{h}}{\overline{\rho}} \tag{B17}$$

and the latter $\xi_{\mathrm{gm,sat}}$ from the Fourier transform of its power spectrum

$$P_{\mathrm{gm,ss}}(k) = \int_{M_{\mathrm{low}}}^{M_{\mathrm{high}}} \mathrm{d}M_\mathrm{h}\, n(M_\mathrm{h}) \frac{\langle N_\mathrm{sat}(M_\mathrm{h})\rangle}{n_\mathrm{gal}} \frac{M_\mathrm{h}}{\overline{\rho}} |u_\mathrm{h}(k|M_\mathrm{h})|^2. \tag{B18}$$

Finally, we compute the two-halo term $\xi_{\mathrm{gm,\,2}}(r)$ from the Fourier transform of the galaxy–dark matter cross-correlation power spectrum:

$$P_{\mathrm{gm,2}}(k, r) = P_\mathrm{m}(k) \times \int_{M_{\mathrm{low}}}^{M_{\mathrm{lim}}(r)} \mathrm{d}M_\mathrm{h} n(M_\mathrm{h}) \frac{\langle N_\mathrm{tot}(M_\mathrm{h})\rangle}{n'_\mathrm{gal}(r)} b_\mathrm{h}(M_\mathrm{h}, r) |u_\mathrm{h}(k|M_\mathrm{h})|, \tag{B19}$$

with a similar treatment of halo exclusion to that of the galaxy power spectrum.

APPENDIX C: SYSTEMATICS CHECKS ON LENSING AND CLUSTERING

We have performed systematics checks for the lensing and clustering measurements. In Fig. C1, we detail the galaxy–galaxy lensing measurement for the sample $10.40 < \log(M_\star/\mathrm{M}_\odot) < 10.65$ as an example. The top panel shows the data (dots with error bars) and best-fitting model (thick line) with the different components of the model the central galaxy term, the satellite term and the two-halo term. The lower panels show a number of systematics checks. The '$e_\times$' panel shows the signal measured after rotating the ellipticities by 45° and the 'ran. lenses' panel shows the signal measured by randomizing the lenses positions, both consistent with zero. The '1+m' panel shows the multiplicative bias correction applied to the galaxy–galaxy lensing measurement, estimated after replacing the ellipticities by the multiplicative calibration factor $1 + m$. The 'boost factor' was estimated from randomizing the background source positions and measuring the ratio of the number of real sources over random objects as a function of distance from the lenses, and applied to the galaxy–galaxy lensing measurement. The covariance matrix from the jackknife estimate is shown in the

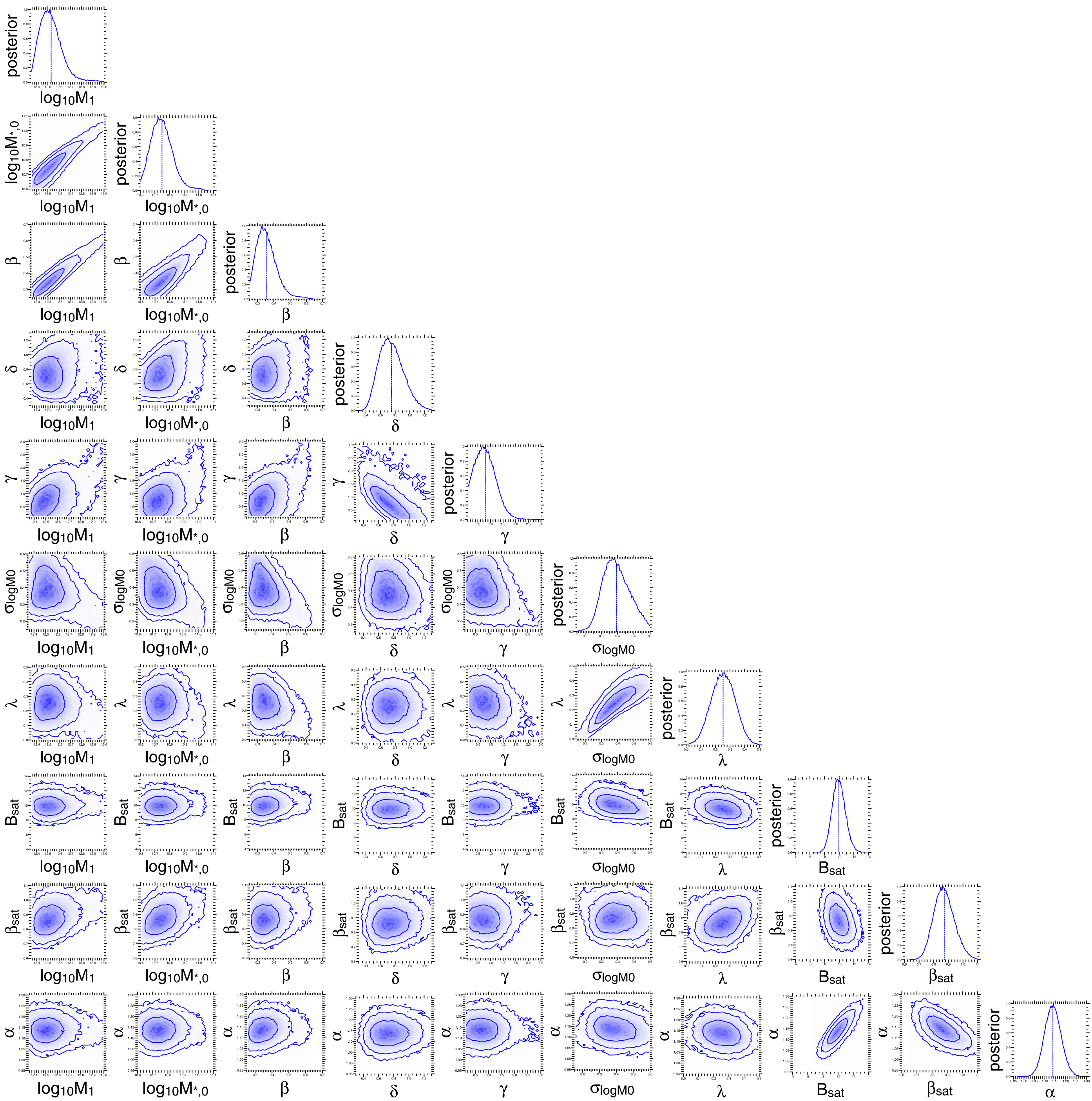


Figure D1. 1D (diagonal) and 2D likelihood distributions of best-fitting HOD parameters in the case of total errors. The 2D contours represent the 68.3, 95.5 and 99.7 per cent confidence limits. We used flat priors within the ranges shown on the figure for all parameters.

left-bottom corner of the figure. The relatively small off-diagonal values show the low correlation between data points. We repeated identical tests for all mass bins. In all cases, systematics are found to be consistent with zero.

In Fig. C2, we test the impact of including high-redshift sources beyond $z > 1.2$. To do so, we select an arbitrary sample of low-redshift lens galaxies with a spectroscopic redshift and we measured the galaxy–galaxy lensing signal using all sources with $0.8 < z_{\rm p} < 1.2$ (purple dots in the figure) and all sources with $z_{\rm p} > 1.2$ (green triangles in the figure). We see no significant difference between the two signals, meaning that the photometric redshifts and shape measurements in our catalogue are robust enough beyond $z_{\rm p} > 1.2$.

In Fig. C3, we show the projected clustering in the mass bin $10.60 < \log(M_\star/\mathrm{M}_\odot) < 10.80$. The top panel shows the data points with error bars and the best-fitting model, with the different components of the model: the one-halo term split into the central-satellite and satellite–satellite terms and the two-halo term. In the middle panel, we show the corresponding HOD, as a dashed line for the central contribution and as a dot–dashed line for the satellites' contribution.

APPENDIX D: 2D CONTOURS

We show in Fig. D1, the likelihood distributions of the best-fitting HOD parameters. Here, the results are shown for the MCMC run done with total (statistical plus systematic) errors.

[1]*Astronomical Observatory of the University of Geneva, ch. d'Ecogia 16, CH-1290 Versoix, Switzerland*
[2]*Aix Marseille Université, CNRS, LAM (Laboratoire d'Astrophysique de Marseille) UMR 7326, 13388, Marseille, France*
[3]*Department of Physics and Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, BC V6T 1Z1,Canada*
[4]*Argelander Institute for Astronomy, University of Bonn, Auf dem Hügel 71, D-53121 Bonn, Germany*
[5]*INAF Istituto di Astrofisica Spaziale e Fisica Cosmica Milano, via Bassini 15, I-20133 Milano, Italy*
[6]*INAF Osservatorio Astronomico di Brera, via Brera 28, 20122 Milano, via E. Bianchi 46, I-23807 Merate, Italy*
[7]*3 Dipartimento di Fisica, Università di Milano-Bicocca, P.zza della Scienza 3, I-20126 Milano, Italy*
[8]*The Scottish Universities Physics Alliance, Institute for Astronomy, University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ, UK*
[9]*Leiden Observatory, Leiden University, Niels Bohrweg 2, NL-2333 CA Leiden, the Netherlands*
[10]*CEA Saclay, Service d'Astrophysique (SAp), Orme des Merisiers, Bât 709, F-91191 Gif-sur-Yvette, France*
[11]*Mullard Space Science Laboratory, University College London, Holmbury St Mary, Dorking, Surrey RH5 6NT, UK*
[12]*Sorbonne Universités, UPMC Univ Paris 06, UMR 7095, Institut d'Astrophysique de Paris, F-75014 Paris, France*
[13]*CNRS, UMR 7095, Institut d'Astrophysique de Paris, F-75014 Paris, France*
[14]*University of Oxford, Department of Physics, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK*
[15]*IFAE, Campus UAB, E-08193 Bellaterra, Spain*
[16]*Dipartimento di Matematica e Fisica, Università degli Studi Roma Tre, via della Vasca Navale 84, I-00146 Roma, Italy*
[17]*INFN, Sezione di Roma Tre, via della Vasca Navale 84, I-00146 Roma, Italy*
[18]*INAF Osservatorio Astronomico di Roma, via Frascati 33, I-00040 Monte Porzio Catone (RM), Italy*
[19]*Dipartimento di Fisica e Astronomia - Università di Bologna, viale Berti Pichat 6/2, I-40127 Bologna, Italy*
[20]*INAF Osservatorio Astronomico di Bologna, via Ranzani 1, I-40127 Bologna, Italy*
[21]*INAF Osservatorio Astronomico di Trieste, via G. B. Tiepolo 11, I-34143 Trieste, Italy*
[22]*Shanghai Key Lab for Astrophysics, Shanghai Normal University, 100 Guilin Road, 200234, Shanghai, China*
[23]*Department of Physics and Astronomy, University of Waterloo, Waterloo, ON N2L 3G1, Canada*
[24]*Perimeter Institute for Theoretical Physics, 31 Caroline Street N, Waterloo, ON N2L 1Y5, Canada*
[25]*Kavli Institute for the Physics and Mathematics of the Universe, Todai Institutes for Advanced Study, the University of Tokyo, Kashiwa 277-8583, Japan*
[26]*INFN, Sezione di Bologna, viale Berti Pichat 6/2, I-40127 Bologna, Italy*
[27]*Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK*

This paper has been typeset from a TEX/LATEX file prepared by the author.

4.3.6 Article 6 : The cosmic Web reconstruction in the GAMA survey

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **474,** 547–571 (2018) doi:10.1093/mnras/stx2638
Advance Access publication 2017 October 10

Galaxy evolution in the metric of the cosmic web

K. Kraljic,[1★] S. Arnouts,[1] C. Pichon,[2,3] C. Laigle,[4] S. de la Torre,[1] D. Vibert,[1] C. Cadiou,[2] Y. Dubois,[2] M. Treyer,[1] C. Schimd,[1] S. Codis,[5] V. de Lapparent,[2] J. Devriendt,[4] H. S. Hwang,[6] D. Le Borgne,[2] N. Malavasi,[7] B. Milliard,[1] M. Musso,[2,8] D. Pogosyan,[9] M. Alpaslan,[10] J. Bland-Hawthorn[11] and A. H. Wright[12]

[1]*Aix Marseille University, CNRS, LAM, Laboratoire d'Astrophysique de Marseille, Marseille, France*
[2]*Institut d'Astrophysique de Paris, UMR 7095 CNRS et Université Pierre et Marie Curie, 98bis Bd Arago, F-75014, Paris, France*
[3]*School of Physics, Korea Institute for Advanced Study (KIAS), 85 Hoegiro, Dongdaemun-gu, Seoul, 02455, Republic of Korea*
[4]*Sub-department of Astrophysics, University of Oxford, Keble Road, Oxford OX1 3RH, United Kingdom*
[5]*Canadian Institute for Theoretical Astrophysics, University of Toronto, 60 St. George Street, Toronto, ON M5S 3H8, Canada*
[6]*Quantum Universe Center, Korea Institute for Advanced Study (KIAS), 85 Hoegiro, Dongdaemun-gu, Seoul 02455, Republic of Korea*
[7]*Department of Physics and Astronomy, Purdue University, 525 Northwestern Avenue, West Lafayette, IN 47907, USA*
[8]*Institut de physique théorique, Université Paris Saclay and CEA, CNRS, 91191 Gif-sur-Yvette, France*
[9]*Department of Physics, University of Alberta, 412 Avadh Bhatia Physics Laboratory, Edmonton, AB T6G 2J1, Canada*
[10]*NASA Ames Research Center, Moffett Field, Mountain View, CA 94035, USA*
[11]*Sydney Institute for Astronomy, School of Physics, University of Sydney, NSW 2006, Australia*
[12]*Argelander-Institut für Astronomie (AIfA), Universität Bonn, Auf dem Hügel 71, D-53121 Bonn, Germany*

Accepted 2017 October 6. Received 2017 September 8; in original form 2017 July 19

ABSTRACT

The role of the cosmic web in shaping galaxy properties is investigated in the Galaxy And Mass Assembly (GAMA) spectroscopic survey in the redshift range $0.03 \leq z \leq 0.25$. The stellar mass, $u - r$ dust corrected colour and specific star formation rate (sSFR) of galaxies are analysed as a function of their distances to the 3D cosmic web features, such as nodes, filaments and walls, as reconstructed by DisPerSE. Significant mass and type/colour gradients are found for the whole population, with more massive and/or passive galaxies being located closer to the filament and wall than their less massive and/or star-forming counterparts. Mass segregation persists among the star-forming population alone. The red fraction of galaxies increases when closing in on nodes, and on filaments regardless of the distance to nodes. Similarly, the star-forming population reddens (or lowers its sSFR) at fixed mass when closing in on filament, implying that some quenching takes place. These trends are also found in the state-of-the-art hydrodynamical simulation HORIZON-AGN. These results suggest that on top of stellar mass and large-scale density, the traceless component of the tides from the anisotropic large-scale environment also shapes galactic properties. An extension of excursion theory accounting for filamentary tides provides a qualitative explanation in terms of anisotropic assembly bias: at a given mass, the accretion rate varies with the orientation and distance to filaments. It also explains the absence of type/colour gradients in the data on smaller, non-linear scales.

Key words: large-scale structure of Universe – cosmology: observations – galaxies: evolution – galaxies: high-redshift – galaxies: statistics.

1 INTRODUCTION

Within the Λ cold dark matter (ΛCDM) cosmological paradigm, structures in the present-day Universe arise from hierarchical clustering, with smaller dark matter haloes forming first and progressively merging into larger ones. Galaxies form by the cooling and condensation of baryons that settle in the centres of these haloes (White & Rees 1978) and their spin is predicted to be correlated with that of the halo generated from the tidal field torques at the moment of proto-halo collapse (tidal torque theory, TTT; e.g. Peebles 1969; Doroshkevich 1970; Efstathiou & Jones 1979; White 1984). However, dark matter haloes, and galaxies residing within them, are not isolated. They are part of a larger-scale pattern, dubbed the cosmic web (Jõeveer, Einasto & Tago 1978; Bond, Kofman & Pogosyan 1996), arising from the anisotropic collapse

★E-mail: katarina.kraljic@lam.fr

of the initial fluctuations of the matter density field under the effect of gravity across cosmic time (Zel'dovich 1970).

This web-like pattern, brought to light by systematic galaxy redshift surveys (e.g. De Lapparent, Geller & Huchra 1986; Geller & Huchra 1989; Colless et al. 2001; Tegmark et al. 2004), consists of large nearly-empty void regions surrounded by sheet-like walls framed by filaments which intersect at the location of clusters of galaxies. These are interpreted as the nodes, or high-density peaks of the large-scale structure pattern, containing a large fraction of the dark matter mass (Bond et al. 1996; Pogosyan et al. 1996). The baryonic gas follows the gravitational potential gradients imposed by the dark matter distribution, then shocks and winds up around multistream, vorticity-rich filaments (Codis et al. 2012; Hahn, Angulo & Abel 2015; Laigle et al. 2015). Filamentary flows, along specific directions dictated by the geometry of the cosmic web, advect angular momentum into the newly formed low mass galaxies with spins typically aligned with their neighbouring filaments (Pichon et al. 2011; Stewart et al. 2013). The next generation of galaxies forms through mergers as they drift along these filaments towards the nodes of the cosmic web with a post merger spin preferentially perpendicular to the filaments, having converted the orbital momentum into spin (e.g. Aubert, Pichon & Colombi 2004; Navarro, Abadi & Steinmetz 2004; Aragón-Calvo et al. 2007b; Codis et al. 2012; Libeskind et al. 2012; Trowland, Lewis & Bland-Hawthorn 2013; Aragon-Calvo & Yang 2014; Dubois et al. 2014; Welker et al. 2015).

Within the standard paradigm of hierarchical structure formation based on ΛCDM cosmology (Blumenthal et al. 1984; Davis et al. 1985), the imprint of the (*past*) large-scale environment on galaxy properties is therefore, to some degree, expected via galaxy mass assembly history. Intrinsic properties, such as the mass of a galaxy (and internal processes that are directly linked to its mass), are indeed shaped by its build-up process, which in turn is correlated with its *present* environment. For instance, more massive galaxies are found to reside preferentially in denser environments (e.g. Dressler 1980; Postman & Geller 1984; Kauffmann et al. 2004; Baldry et al. 2006). This mass-density relation can be explained through the biased mass function in the vicinity of the large-scale structure (LSS; Kaiser 1984; Efstathiou et al. 1988) where the enhanced density of the dark matter field allows the proto-halo to pass the critical threshold of collapse earlier (Bond et al. 1991) resulting in an overabundance of massive haloes in dense environments. However, what is still rightfully debated is whether the large-scale environment is also driving other observed trends such as morphology-density (e.g. Dressler 1980; Postman & Geller 1984; Dressler et al. 1997; Goto et al. 2003), colour-density (e.g. Blanton et al. 2003; Baldry et al. 2006; Bamford et al. 2009) or star formation-density (e.g. Hashimoto et al. 1998; Lewis et al. 2002; Kauffmann et al. 2004) relations, and galactic 'spin' properties, such as their angular momentum vector, their orientation or chirality (trailing versus leading arms).

On the one hand, there are evidences that the cosmic web affects galaxy properties. Void galaxies are found to be less massive, bluer, and more compact than galaxies outside of voids (e.g. Rojas et al. 2004; Beygu et al. 2016); galaxies infalling into clusters along filaments show signs of some physical mechanisms operating even before becoming part of these systems, that galaxies in the isotropic infalling regions do not (Porter et al. 2008; Martínez, Muriel & Coenda 2016); Kleiner et al. (2017) find systematically higher HI fractions for massive galaxies ($M_\star > 10^{11}\,\mathrm{M}_\odot$) near filaments compared to the field population, interpreted as evidence for a more efficient cold gas accretion from the intergalactic medium; Kuutma, Tamm & Tempel (2017) report an environmental transformation with a higher elliptical-to-spiral ratio when moving closer to filaments, interpreted as an increase in the merging rate or the cut-off of gas supplies near and inside filaments (see also Aragon-Calvo, Neyrinck & Silk 2016); Chen et al. (2017) detect a strong correlation of galaxy properties, such as colour, stellar mass, age and size, with the distance to filaments and clusters, highlighting their role beyond the environmental density effect, with red or high-mass galaxies and early-forming or large galaxies at fixed stellar mass having shorter distances to filaments and clusters than blue or low-mass and late-forming or small galaxies, and Tojeiro et al. (2017) interpret a steadily increasing stellar-to-halo mass ratio from voids to nodes for low mass haloes, with the reversal of the trend at the high-mass end, found for central galaxies in the GAMA survey (Driver et al. 2009, 2011), as an evidence for halo assembly bias being a function of geometric environment. At higher redshift, a small but significant trend in the distribution of galaxy properties within filaments was reported in the spectroscopic survey VIPERS ($z \simeq 0.7$; Malavasi et al. 2017) and with photometric redshifts ($0.5 < z < 0.9$) in the COSMOS field (with a 2D analysis; Laigle et al. 2017). Both studies find significant mass and type segregations, where the most massive or quiescent galaxies are closer to filaments than less massive or active galaxies, emphasizing that large-scale cosmic flows play a role in shaping galaxy properties.

On the other hand, Alpaslan et al. (2015) find in the GAMA data that the most important parameter driving galaxy properties is stellar mass as opposed to environment (see also Robotham et al. 2013). Similarly, while focusing on spiral galaxies alone, Alpaslan et al. (2016) do find variations in the star formation rate (SFR) distribution with large-scale environments, but they are identified as a secondary effect. Another quantity tracing different geometric environments that was found to vary is the luminosity function. However, while Guo, Tempel & Libeskind (2015) conclude that the filamentary environment may have a strong effect on the efficiency of galaxy formation (see also Benítez-Llambay et al. 2013), Eardley et al. (2015) argue that there is no evidence of a direct influence of the cosmic web as these variations can be entirely driven by the underlying *local* density dependence. These discrepancies are partially expected: the present state of galaxies must be impacted by the effect of the past environment, which in turn does correlate with the present environment, if mildly so, but these environmental effects must first be distinguished from mass-driven effects which typically dominate.

The TTT, naturally connecting the large-scale distribution of matter and the angular momentum of galactic haloes (e.g. Jones & Efstathiou 1979; Barnes & Efstathiou 1987; Heavens & Peacock 1988; Porciani, Dekel & Hoffman 2002a,b; Lee 2004), in its recently revisited, conditioned formulation (Codis, Pichon & Pogosyan 2015) predicts the angular momentum distribution of the forming galaxies relative to the cosmic web, which tend to first have their angular momentum aligned with the filament's direction while the spin orientation of massive galaxies is preferentially in the perpendicular direction. Despite the difficulty to model properly the halo-galaxy connection, due to the complexity, non-linearity and multiscale character of the involved processes, modern cosmological hydrodynamic simulations confirm such a mass-dependent angular momentum distribution of galaxies with respect to the cosmic web (Dubois et al. 2014; Welker et al. 2014, 2017). On galactic scales, the dynamical influence of the cosmic web is therefore traced by the distribution of angular momentum and orientation of galaxies, when measured relative to their embedding large-scale environment. The impact of such environment on the spins of galaxies has only recently started to be observed (confirming the spin alignment for spirals and preferred perpendicular orientation for ellipticals; Trujillo et al. 2006; Lee & Erdogdu 2007; Paz et al. 2008; Tempel et al. 2013; Tempel & Libeskind 2013; Pahwa et al. 2016, but see also Jones, van de Weygaert & Aragón-Calvo 2010; Cervantes-Sodi, Hernandez & Park 2010; Andrae & Jahnke 2011, for contradictory results). What is less obvious is whether observed integrated scalar properties such as morphology or physical properties (SFR, type, metallicity, which depend not only on the mass but also on the past and present gas accretion) are also impacted.

Theoretical considerations alone suggest that local density as a sole and unique parameter (and consequently any isotropic definition of the environment based on density alone) is not sufficient to account for the effect of gravity on galactic scale (e.g. Mo, van den Bosch & White 2010) and therefore capture the environmental diversity in which galaxies form and evolve: one must also consider the relative past and present orientation of the tidal tensor with respect to directions pointing towards the larger-scale structure principal axes. At the simplest level, on large scales, gravity should be the dominant force. Its net cumulative impact is encoded in the tides operating on the host dark matter halo. Such tides may be decomposed into the trace of the tidal tensor, which equals the local density, and its traceless part, which applies distortion and rotation to the forming galaxy. The effect of the former on increasing scales has long been taken into account in standard galaxy formation scenarios (Kaiser 1984), while the effect of the latter has only recently received full attention (e.g. Codis et al. 2015). Beyond the above-discussed effect on angular momentum, other galaxy's properties could in principle be influenced by the large-scale traceless part of the tidal field, which modifies the accretion history of a halo depending on its location within the cosmic web. For instance, the tidal shear near saddles along the filaments feeding massive haloes is predicted to slow down the mass assembly of smaller haloes in their vicinity (Hahn et al. 2009; Borzyszkowski et al. 2017; Castorina et al. 2016). Bond & Myers (1996) integrated the effect of ellipsoidal collapse (via the shear amplitude), which may partially delay galaxy formation, in the Extended Press-Schechter (EPS) theory. Yet, in that formulation, the geometry of the delay imposed by the specific relative orientation of tides imposed by the large-scale structure is not accounted for, because time delays are ensemble-averaged over all possible geometries of the LSS. The anisotropy of the large-scale cosmic web – voids, walls, filaments, and nodes (which shape and orient the tidal tensor beyond its trace) – should therefore be taken into account explicitly, as it impacts mass assembly. Despite of the above-mentioned difficulty in properly describing the connection between galaxies and their host dark matter haloes, this anisotropy should have direct observational signatures in the differential properties of galaxies with respect to the cosmic web at fixed mass and local density. Quantifying these signatures is the topic of this paper. Extending EPS to account for the geometry of the tides beyond that encoded in the density of the field is the topic of the companion paper (Musso et al. 2017).

This paper explores the impact of the cosmic web on galaxy properties in the GAMA survey, using the Discrete Persistent Structure Extractor code (DisPerSE; Sousbie 2011; Sousbie, Pichon & Kawahara 2011) to characterize its 3D topological features, such as nodes, filaments and walls. GAMA is to date the best data set for this kind of study, given its unique spectroscopic combination of depth, area, target density and high completeness, as well as its broad multiwavelength coverage. Variations in stellar mass and colour, red fraction and star formation activity are investigated as a function of galaxy distances to these three features. The rest of the paper is organized as follows. Section 2 summarizes the data and describes the sample selection. The method used to reconstruct the cosmic web is presented in Section 3. Section 4 investigates the stellar-mass and type/colour segregation and the star formation activity of galaxies within the cosmic web. Section 5 shows how these results compare to those obtained in the HORIZON-AGN simulation (Dubois et al. 2014). Section 6 addresses the impact of the density on the measured gradients towards filaments and walls. Results are discussed in Section 7 jointly with predictions from Musso et al. (2017). Finally, Section 8 concludes. Additional details on the matching technique and the impact of the boundaries to the measured gradients are provided in Appendices A and B, respectively. Appendix C investigates the effect of smoothing scale on the found gradients, Appendix D briefly presents the HORIZON-AGN simulation, Appendix F provides tables of median gradients and a short summary of predicted gradient misalignments is presented in Appendix E.

Throughout the study, a flat ΛCDM cosmology with $H_0 = 67.5\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_M = 0.31$ and $\Omega_\Lambda = 0.69$ is adopted (Planck Collaboration XIII 2016). All statistical errors are computed by bootstrapping, such that the errors on a given statistical quantity correspond to the standard deviation of the distribution of that quantity re-computed in 100 random samples drawn from the parent sample with replacement. All magnitudes are quoted in the AB system, and by log we refer to the 10-based logarithm.

2 DATA AND DATA PRODUCTS

This section describes the observational data and derived products, namely the galaxy and group catalogues, that have been used in this work.

2.1 Galaxy catalogue

The analysis is based on the GAMA survey[1] (Driver et al. 2009, 2011; Hopkins et al. 2013; Liske et al. 2015), a joint European-Australian project combining multiwavelength photometry (UV to far-IR) from ground and space-based facilities and spectroscopy obtained at the Anglo-Australian Telescope (AAT, NSW, Australia) using the AAOmega spectrograph. GAMA provides spectra for galaxies across five regions, but this work only considers the three equatorial fields G9, G12 and G15 covering a total area of 180 deg^2 (12 × 5 deg^2 each), for which the spectroscopic completeness is >98 per cent down to a r-band apparent magnitude $m_r = 19.8$. The reader is referred to Wright et al. (2016) for a complete description of the spectro-photometric catalogue constructed using the LAMBDAR[2] code that was applied to the 21-band photometric data set from the GAMA Panchromatic Data Release (Driver et al. 2016), containing imaging spanning the far-UV to the far-IR.

The physical parameters for the galaxy sample such as the absolute magnitudes, extinction corrected rest-frame colours, stellar masses and specific star formation rate (sSFR) are derived using a grid of model spectral energy distributions (SED; Bruzual & Charlot 2003) and the SED fitting code LEPHARE[3] (Arnouts et al. 1999; Ilbert et al. 2006). The details used to derive these physical parameters are given in the companion paper (Treyer et al. in preparation).

[1] http://www.gama-survey.org/
[2] Lambda Adaptive Multi-Band Deblending Algorithm in *R*.
[3] http://cesam.lam.fr/lephare/lephare.html

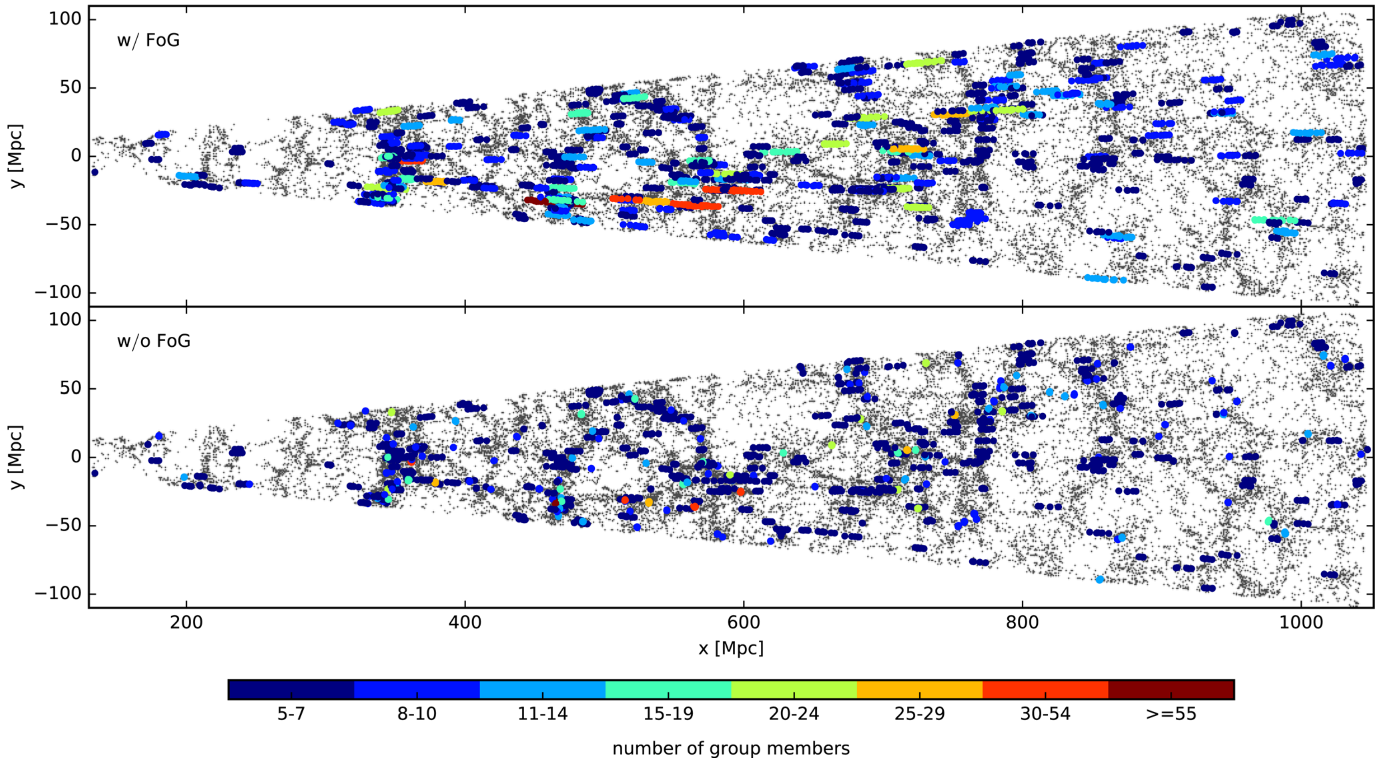


Figure 1. Spatial distribution of whole galaxy population with $m_r < 19.8$ in the GAMA field G12 in the redshift range $0.03 \leq z \leq 0.25$ (grey points). Overplotted are galaxy group members, colour coded by the size of their group. Only groups having five or more members are shown. The top and bottom panels illustrate the galaxy group members before and after correcting for the FoG effect, respectively.

The classification between the active (star-forming) and passive (quiescent) populations is based on a simple colour cut at $u - r = 1.8$ in the rest-frame extinction corrected $u - r$ versus r diagram that is used to separate the two populations. This colour cut is consistent with a cut in sSFR at $10^{-10.8}$ yr^{-1} (see Treyer et al. in preparation). Hence, in what follows, the terms red (blue) and quiescent (star-forming) will be used interchangeably.

The analysis is restricted to the redshift range $0.03 \leq z \leq 0.25$, totalling 97 072 galaxies. This is motivated by the high galaxy sampling required to reliably reconstruct the cosmic web. Beyond $z \sim 0.25$, the galaxy number density drops substantially (to 2×10^{-3} Mpc^{-3} from 8×10^{-3} Mpc^{-3} at $z \leq 0.25$, on average), while below $z \sim 0.03$, the small volume does not allow us to explore the large scales of the cosmic web.

The stellar mass completeness limits are defined for the passive and active galaxies as the mass above which 90 per cent of galaxies of a given type (blue/red) reside at a given redshift $z \pm 0.004$. This translates into mass completeness limits of $\log(M_\star/\mathrm{M}_\odot) = 9.92$ and $\log(M_\star/\mathrm{M}_\odot) = 10.46$ for the blue and red populations at $z \leq 0.25$, respectively.

2.2 Group catalogue

Since the three-dimensional distribution of galaxies relies on the redshift-based measures of distances, it is affected by their peculiar velocities. In order to optimize the cosmic web reconstruction, one needs to take into account these redshift-space distortions. On large scales, these arise from the coherent motion of galaxies accompanying the growth of structure, causing its flattening along the line of sight, the so-called Kaiser effect (Kaiser 1987). On small scales, the so-called Fingers of God (FoG; Jackson 1972; Tully & Fisher 1978) effect, induced by the random motions of galaxies within virialized haloes (groups and clusters) causes the apparent elongation of structures in redshift space, clearly visible in the galaxy distribution in the GAMA survey (Fig. 1, top panel). While the Kaiser effect tends to enhance the cosmic web by increasing the contrast of filaments and walls (e.g. Subba Rao et al. 2008; Shi et al. 2016), the FoG effect may lead to the identification of spurious filaments. Because the impact of the Kaiser effect is expected to be much less significant than that of the FoG (e.g. Subba Rao et al. 2008; Kuutma et al. 2017), for the purposes of this work, we do not attempt to correct for it and we focus on the compression of the FoG only. To do so, the galaxy groups are first constructed with the use of an anisotropic Friends-of-Friends (FoF) algorithm operating on the projected perpendicular and parallel separations of galaxies, that was calibrated and tested using the publicly available GAMA mock catalogues of Robotham et al. (2011) (see also Merson et al. 2013, for details of the mock catalogues construction). Details on the construction of the group catalogue and related analysis of group properties can be found in the companion paper (Treyer et al. in preparation). Next, the centre of each group is identified following Robotham et al. (2011) (see also Eke et al. 2004, for a different implementation). The method is based on an iterative approach: first, the centre of mass of the group (CoM) is computed; next its projected distance from the CoM is found iteratively for each galaxy in the group by rejecting the most distant galaxy. This process stops when only two galaxies remain and the most massive galaxy is then identified as the centre of the group. The advantage of this method, as shown in Robotham et al. (2011), is that the iteratively defined centre is less affected by interlopers than luminosity-weighted centre or the central identified as the most luminous group galaxy. The groups are then compressed radially so that the dispersions in transverse and radial directions are equal, making the galaxies in the groups isotropically distributed about their centres (see e.g. Tegmark et al. 2004). In practice, since the elongated FoG effect affects mostly the largest groups, only groups with more than six members are compressed. Note that the precise correction of the FoG effect is not sought. What is needed for the purpose of this work is the elimination of these elongated structures that could be misidentified as filaments.

Fig. 1 displays the whole galaxy population and the identified FoF groups (coloured by their richness) in the GAMA field G12. The top and bottom panels show the groups before and after correcting for the FoG effect, respectively. For the sake of clarity, only groups having at least five members are shown. The visual inspection reveals that most of the groups are located within dense regions, often at the intersection of the apparently filamentary structures.

3 THE COSMIC WEB EXTRACTION

With the objective of exploring the impact of the LSS on the evolution of galaxy properties, one first needs to properly describe the main components of the cosmic web, namely the high-density peaks (nodes) which are connected by filaments, framing the sheet-like walls, themselves surrounding the void regions. Among the various methods developed over the years, two broad classes can be identified. One uses the geometrical information contained in the *local* gradient and the Hessian of the density or potential field (e.g. Novikov, Colombi & Doré 2006; Aragón-Calvo et al. 2007a,b; Hahn et al. 2007a,b; Sousbie et al. 2008a,b; Forero-Romero et al. 2009; Bond, Strauss & Cen 2010a,b), while the second exploits the topology and connectivity of the density field by using the watershed transform (Aragón-Calvo, van de Weygaert & Jones 2010) or Morse theory (e.g. Colombi, Pogosyan & Souradeep 2000; Sousbie et al. 2008a; Sousbie 2011). The theory for the former can be built in some details, (see e.g. Pogosyan et al. 2009), shedding some light on physical interpretation, while the latter avoids shortcomings of a second-order Taylor expansion of the field and provides a natural metric in which to compute distances to filaments. Within these broad categories, some algorithms deal with discrete data sets, while others require that the density field must be first estimated (possibly on multiple scales). An exhaustive description of several cosmic web extraction techniques and a comparison of their classification patterns as measured in simulations are presented in Libeskind et al. (2017). While this paper found some differences between the various algorithms, which should in principle be accounted for as modelling errors in this work, these differences remain small on the scales considered.

3.1 Cosmic web with disperse

This work uses the Discrete Persistent Structure Extractor (DisPerSE; see Sousbie et al. 2011, for illustrations in a cosmological context), a geometric three-dimensional ridge extractor dealing directly with discrete data sets, making it particularly well adapted for astrophysical applications. It allows for a scale and parameter-free coherent identification of the 3D structures of the cosmic web as dictated by the large-scale topology. For a detailed description of the DisPerSE algorithm and its underlying theory, the reader is referred to Sousbie (2011); its main features are summarized below.

DisPerSE is based on discrete Morse and persistence theories. The Delaunay tessellation is used to generate a simplicial complex, i.e. a triangulated space with a geometric assembly of cells, faces, edges and vertices mapping the whole volume. The Delaunay Tessellation Field Estimator (DTFE; Schaap & van de Weygaert 2000; Cautun & van de Weygaert 2011) allows for estimating the density field at each vertex of the Delaunay complex. The Morse theory enables to extract from the density field the critical points, i.e. points with a vanishing (discrete) gradient of the density field (e.g. maxima, minima and saddle points). These critical points are connected via the field lines tangent to the gradient field in every point. They induce a geometrical segmentation of space, where all the field lines have the same origin and destination, known as the Morse complex. This segmentation defines distinct regions called ascending and descending k-manifolds.[4] The morphological components of the cosmic web are then identified from these manifolds: ascending 0-manifolds trace the voids, ascending 1-manifolds trace the walls and filaments correspond to the ascending 2-manifolds with their extremities plugged on to the maxima (peaks of the density field). In addition to its ability to work with sparsely sampled data sets while assuming nothing about the geometry or homogeneity of the survey, DisPerSE allows for the selection of retained structures on the basis of the significance of the topological connection between critical points. DisPerSE relies on persistent homology theory to pair critical points according to the birth and death of a topological feature in the excursion. The ‘persistence’ of a feature or its significance is assessed by the density contrast of the critical pair chosen to pass a certain signal-to-noise threshold. The noise level is defined relative to the RMS of persistence values obtained from random sets of points. This thresholding eliminates less significant critical pairs, allowing to simplify the Morse complex, retaining its most topologically robust features. Fig. 2 shows that filaments outskirt walls, themselves circumventing voids. The filaments are made of a set of connected segments and their end points are connected to the maxima, the peaks of the density field where most of clusters and large groups reside. Each wall is composed of the facets of tetrahedra from the Delaunay tessellation belonging to the same ascending 2-manifold. In this work, DisPerSE is run on the flux-limited GAMA data with a 3σ persistence threshold. Fig. 3 illustrates the filaments for the G12 field, overplotted on the density contrast of the underlying galaxy distribution, $1 + \delta$, where the local density is estimated using the DTFE density estimator. Even in this 2D projected visualization, one can see that filaments trace the ridges of the 3D density field connecting the density peaks between them.

3.2 Cosmic web metric

Having identified the major cosmic web features, let us now define a new metric to characterize the environment of a galaxy, which will be referred to as the ‘cosmic web metric’ and into which galaxies are projected. Fig. 4 gives a schematic view of this framework. Each galaxy is assigned the distance to its closest filament, $D_{\rm skel}$. The impact point in the filament is then used to define the distances along the filament towards the node, $D_{\rm node}$ and towards the saddle point, $D_{\rm saddle}$. Similarly, $D_{\rm wall}$ denotes the distance of the galaxy to its closest wall. In this work, only distances $D_{\rm node}$, $D_{\rm skel}$ and $D_{\rm wall}$ are used. Other investigations of the environment in the vicinity of the saddle points are postponed to a forthcoming work.

The accuracy of the reconstruction of the cosmic web features is sensitive to the sampling of the data set. The lower the

[4] The index k refers to the critical point the field lines emanate from (ascending) or converge to (descending), and is defined as its number of negative eigenvalues of the Hessian: a minimum of the field has index 0, a maximum has index 3 and the two types of saddles have indices 1 and 2.

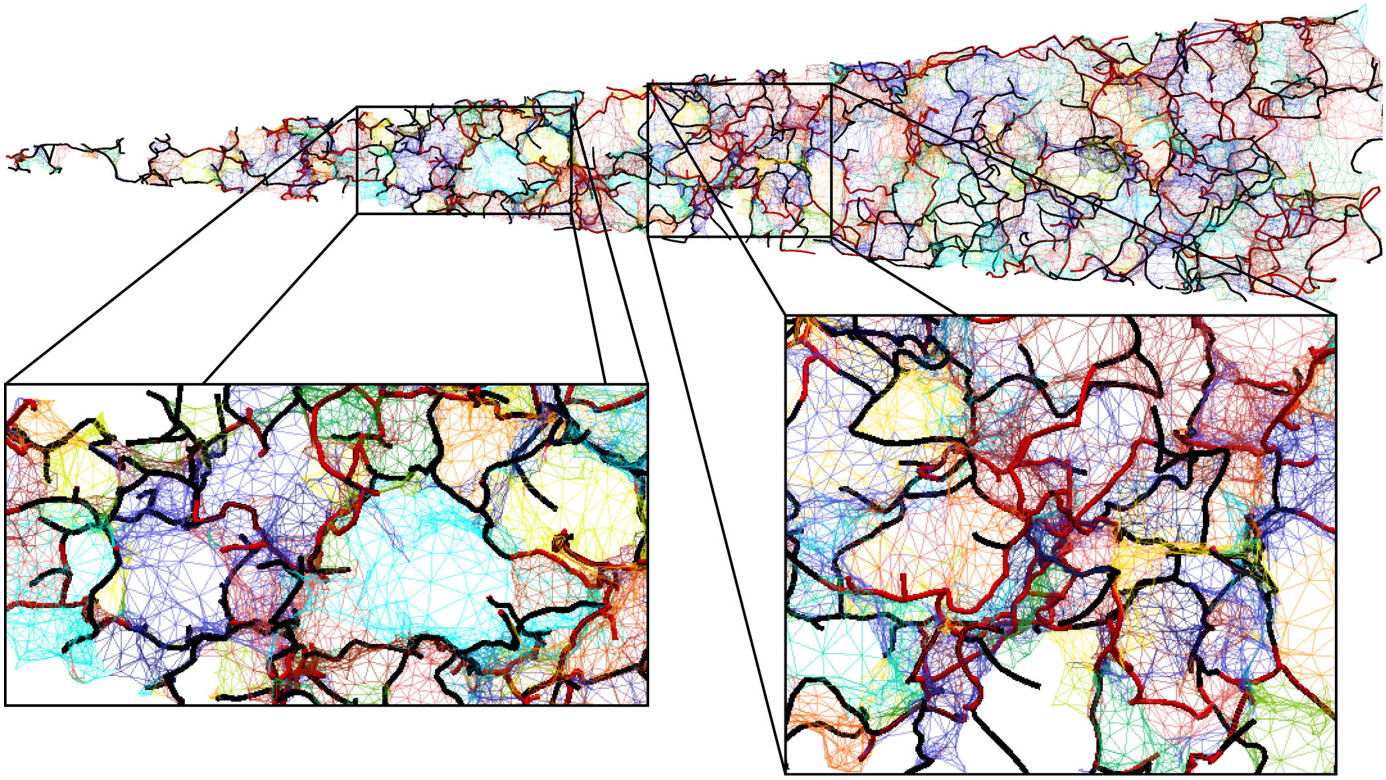


Figure 2. Illustration of the walls and filaments in the G12 field. For the sake of clarity and for the illustrative purposes, only the cosmic web features detected above a persistence threshold of 5σ are shown. Filaments are coloured in black, with the most persistent ones ($>6\sigma$) plotted in red, while walls are colour coded randomly. Note how DisPerSE is capable of recovering the important features of the underlying cosmic field by identifying its (topologically) most-robust features. In particular, it extracts filaments as a set of connected segments, which outskirt walls, themselves circumventing voids.

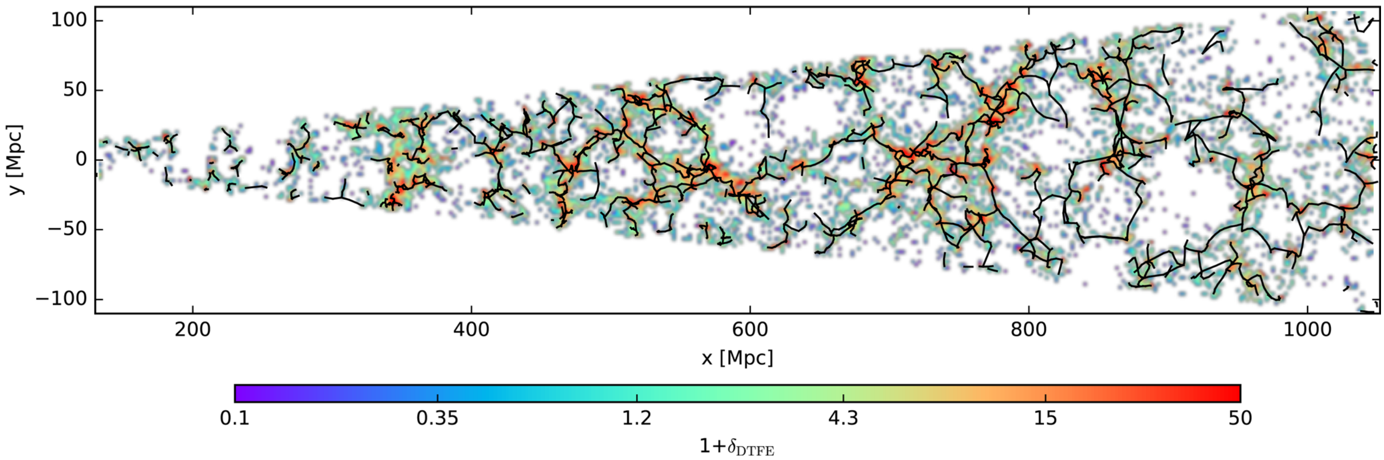


Figure 3. Illustration of the filamentary network (black lines) extracted with the DisPerSE code within the $\pm 1.2^\circ$ of the central declination of the G12 field. The persistence threshold with which the filamentary network and the associated structures, used in this work and shown here, are extracted is 3σ. Also shown is the density contrast of the underlying galaxy distribution, measured with the small-scale adaptive DTFE estimator (see the text) and averaged over cells of $2.3 \times 2.3\,\mathrm{Mpc}^2$ (white colour is used for empty cells). In spite of the projection effects, the visual inspection reveals that filaments follow the ridges of the density field which connect the peaks together.

sampling the larger the uncertainty on the location of the individual components of the cosmic web. To account for the variation of the sampling throughout the survey, unless stated differently, all the distances are normalized by the redshift-dependent mean inter-galaxy separation $\langle D_z \rangle$, defined as $\langle D_z \rangle \equiv n(z)^{-1/3}$, where $n(z)$ represents the number density of galaxies at a given redshift z. For the combined three fields of GAMA survey, $\langle D_z \rangle$ varies from 3.5 to 7.7 Mpc across the redshift range $0.03 \le z \le 0.25$, with a mean value of $\sim$5.6 Mpc.

4 GALAXY PROPERTIES WITHIN THE COSMIC WEB

In this section, the dependence of various galaxy properties, such as stellar mass, $u - r$ colour, sSFR and type, with respect to their location within the cosmic web is analysed. First, the impact of the nodes, representing the largest density peaks, is investigated. Next, by excluding these regions, galaxy properties are studied within the intermediate density regions near the filaments. Finally, the analysis is extended to the walls.

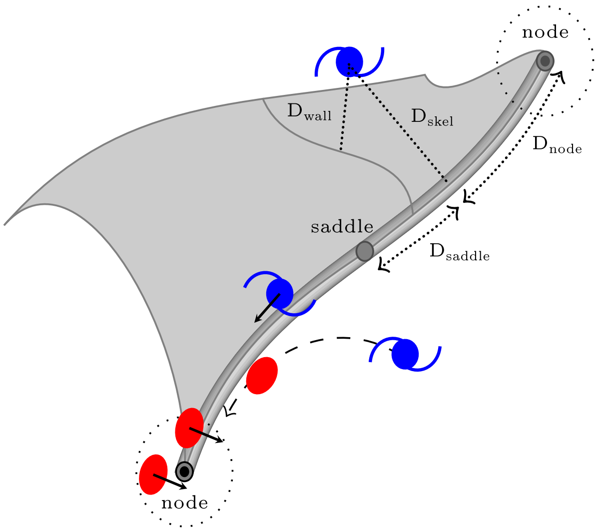


Figure 4. Schematic view of the 'cosmic web' metric in which the analysis is performed. The position of a galaxy within the cosmic web is parametrized by its distance to the closest filament, $D_{\rm skel}$, and its distance to the closest wall, $D_{\rm wall}$. $D_{\rm node}$ and $D_{\rm saddle}$ represent the distances from the impact point to the node and saddle along the corresponding filament, respectively.

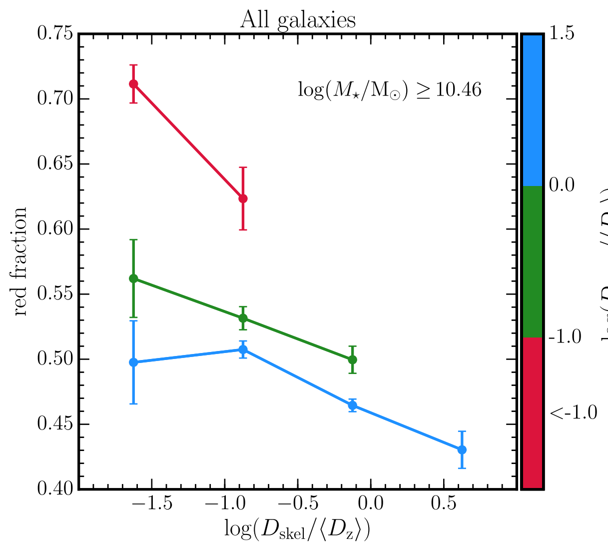


Figure 5. Red fraction of galaxies (the number of quiescent galaxies over the entire population) as a function of $D_{\rm skel}$ for three different bins of $D_{\rm node}$ as indicated by the colour. Both distances are normalized by the redshift-dependent mean inter-galaxy separation $\langle D_z\rangle$. Only galaxies with $\log(M_\star/\mathrm{M}_\odot) \geq 10.46$ are considered. Star-forming and quiescent populations are matched in mass (see Section 4.2.1). The error bars are calculated from 100 bootstrap samples. The fraction of red galaxies is found to increase with decreasing distances both to the closest filament $D_{\rm skel}$ and to the node of this $D_{\rm node}$. Recalling that $\langle D_z\rangle \sim 5.6$ Mpc, the fraction of passive galaxies increases at given $D_{\rm skel}$ by ∼20 per cent from several tens of Mpc away from the nodes (blue line) to less than ∼0.5 Mpc (red line). At fixed $D_{\rm node}$, the increase of the red fraction with decreasing distance to filaments is milder, of ∼10 per cent, regardless of the distance to the node.

4.1 The role of nodes via the red fractions

Let us start by analysing the combined impact of nodes and filaments on galaxies through the study of the red fractions. The red fraction, defined as the number of passive galaxies with respect to the entire population, is analysed as a function of the distance to the nearest filament, $D_{\rm skel}$ and the distance to its associated node, $D_{\rm node}$.

This analysis is restricted to galaxies more massive than $\log(M_\star/\mathrm{M}_\odot) \geq 10.46$, as imposed by the mass limit completeness of the passive population (see Section 2). The stellar mass distributions of the passive and star-forming populations are not identical, with the passive galaxies dominating the high mass end. Therefore, to prevent biases in the measured gradients introduced by such differences, the mass-matched samples are used. The detailed description of the mass-matching technique can be found in Appendix A1.

In Fig. 5 the red fraction of galaxies is shown as a function of $D_{\rm skel}$ in three different bins of $D_{\rm node}$. While the fraction of passive galaxies is found to increase with decreasing distances to both the filaments and nodes, the dominant effect is the distance to the nodes. At fixed $D_{\rm skel}$, the fraction of passive galaxies sharply increases with decreasing distance to the nodes. Recalling that the mean inter-galaxy separation $\langle D_z\rangle \sim 5.6$ Mpc, a 20–30 per cent increase in the fraction of passive galaxies is observed from several Mpc away from the nodes to less than ∼500 kpc. This behaviour is expected since the nodes represent the loci where most of the groups and clusters reside and reflect the well-known colour-density (e.g. Blanton et al. 2003; Baldry et al. 2006; Bamford et al. 2009) and star formation-density (e.g. Lewis et al. 2002; Kauffmann et al. 2004) relations. However, the gradual increase suggests that some physical processes already operate before the galaxies reach the virial radius of massive haloes. At fixed $D_{\rm node}$, the fraction of passive galaxies increases with decreasing distance to filaments, but this increase is milder compared to that with respect to nodes: an increase of ∼10 per cent is observed regardless of the distance to the nodes. These regions with intermediate densities appear to be a place where the transformation of galaxies takes place as emphasized in the next section.

4.2 The role of filaments

In order to infer the role played by filaments alone in the transformation of galactic properties, the impact of nodes, the high-density regions has to be mitigated. By construction, nodes are at the intersection of filaments: they drive the well-known galaxy type-density as well as stellar mass-density relations. To account for this bias, Gay et al. (2010) and Malavasi et al. (2017) adopted a method where a given physical property or distance of each galaxy was down-weighted by its local density. Laigle et al. (2017) adopted a more stringent approach by rejecting all galaxies that are too close to the nodes. This method allows us to minimize the impact of nodes, avoiding the difficult-to-quantify uncertainty of the residual contribution of the density weighting scheme. We therefore adopt the latter approach. As shown in Appendix B1, this is achieved by rejecting all galaxies below a distance of 3.5 Mpc from a node.

4.2.1 Stellar mass gradients

Fig. 6 shows the normalized probability distribution functions (PDFs) of the distance to the nearest filament $D_{\rm skel}$ in three stellar mass bins for the entire population and star-forming galaxies alone (top left-hand and right-hand panels, respectively). The medians of the PDFs, shown by vertical lines, are listed together with the corresponding error bars in Table 1. The significance of the observed trends is assessed by computing the residuals between the distributions in units of σ (bottom panels), defined as $\Delta_{1-2}/\sqrt{\sigma_1^2+\sigma_2^2}$, where Δ_{1-2} is the difference between the PDFs of the populations 1 and 2, and σ_1 and σ_2 are the corresponding standard deviations.

For the entire population (left-hand panels), differences between the PDFs of the three stellar mass bins are observed: the most

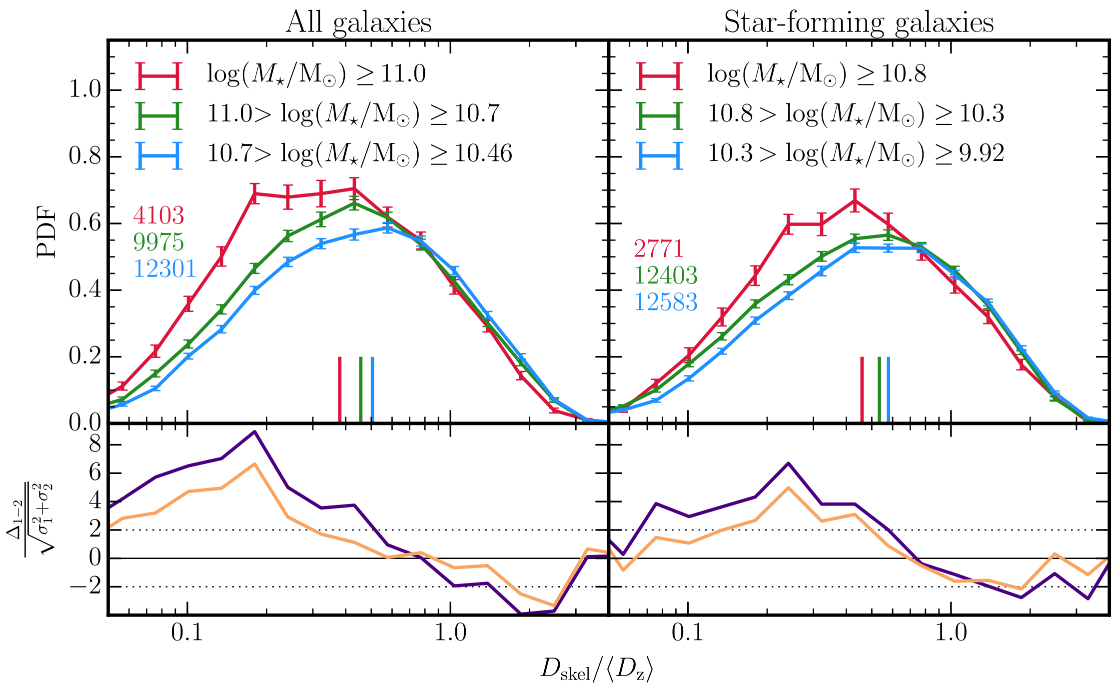


Figure 6. *Top row:* Differential distributions of the distances to the nearest filament, $D_{\rm skel}$ (normalized by $\langle D_z\rangle$, the redshift-dependent mean inter-galaxy separation) for the entire galaxy population (left-hand panel) and star-forming galaxies alone (right-hand panel) in three different stellar mass bins. Note that these bins are different for the two populations: this is due to the stellar mass completeness limit that is different (see Section 2). To highlight an effect specific to the filaments, the contribution of node is minimized (see the text for details). The vertical lines indicate the medians of the distributions and their values together with associated error bars are listed in Table 1. The numbers of galaxies in different considered bins are indicated in each panel. The error bars are calculated from 100 bootstrap samples. There is a mass segregation of galaxies with respect to filaments of the entire as well as star-forming population: more massive galaxies tend to be preferentially located closer to the filaments compared to their lower-mass counterparts. *Bottom row:* Residuals in units of σ between the two most extreme mass bins (purple line; $10.7 > \log(M_\star/{\rm M}_\odot) \geq 10.46$ and $\log(M_\star/{\rm M}_\odot) \geq 11.0$ on the left-hand panel and $10.3 > \log(M_\star/{\rm M}_\odot) \geq 9.92$ and $\log(M_\star/{\rm M}_\odot) \geq 10.8$ on the right-hand panel), and between the high and intermediate mass bins (orange solid line; $\log(M_\star/{\rm M}_\odot) \geq 11.0$ and $11.0 > \log(M_\star/{\rm M}_\odot) \geq 10.7$ on the left-hand panel and $\log(M_\star/{\rm M}_\odot) \geq 10.8$ and $10.8 > \log(M_\star/{\rm M}_\odot) \geq 10.3$ on the right-hand panel).

Table 1. Medians for the PDFs displayed in Figs 6–10.

	Selection[a]	Bin	Median[b] $D_{\rm skel}/\langle D_z\rangle$	Median[b] $D_{\rm wall}/\langle D_z\rangle$
Mass[c]	All galaxies	$\log(M_\star/{\rm M}_\odot) \geq 11$	0.379 ± 0.009	0.334 ± 0.005
		$11 > \log(M_\star/{\rm M}_\odot) \geq 10.7$	0.456 ± 0.007	0.381 ± 0.004
		$10.7 > \log(M_\star/{\rm M}_\odot) \geq 10.46$	0.505 ± 0.006	0.403 ± 0.004
	SF galaxies	$\log(M_\star/{\rm M}_\odot) \geq 11$	0.459 ± 0.012	0.385 ± 0.011
		$11 > \log(M_\star/{\rm M}_\odot) \geq 10.4$	0.534 ± 0.007	0.429 ± 0.006
		$10.4 > \log(M_\star/{\rm M}_\odot) \geq 9.92$	0.578 ± 0.007	0.453 ± 0.007
Type[d]	SF versus passive[e]	Star-forming	0.504 ± 0.008	0.411 ± 0.006
		Passive	0.462 ± 0.007	0.376 ± 0.006

Notes. [a]Panels of Figs 6–10.
[b]Medians of distributions as indicated in Figs 6–10 by vertical lines; errors represent half width at half-maximum of the bootstrap distribution, i.e. the distribution of medians from each of 100 bootstrap samples, fitted by a Gaussian curve.
[c]Figs 6 and 9.
[d]Figs 7 and 10.
[e]Only galaxies with stellar masses $\log(M_\star/{\rm M}_\odot) \geq 10.46$ are considered.

massive galaxies ($\log(M_\star/{\rm M}_\odot) \geq 11$) are located closer to the filaments than the intermediate population ($11 > \log(M_\star/{\rm M}_\odot) \geq 10.7$), while the population with the lowest stellar masses ($10.7 > \log(M_\star/{\rm M}_\odot) \geq 10.46$) is found furthest away from the filaments. The significances of the difference between the most massive and the two lowest stellar mass bins are shown in the bottom panel. Between the most extreme stellar mass bins (purple line), the difference exceeds 4σ close to the filament and 2σ at larger distances. It is slightly less significant between the intermediate and lowest stella mass bins (orange line), but still in excess of 2σ close to the filament. The differences between the PDFs can be also quantified in terms of their medians, where the differences between the highest and lowest stellar mass bins are significant at an $\sim 10\sigma$ level (see Table 1). These results confirm previous claims of a mass segregation with respect to filaments, where the most massive galaxies are located near the core of the filaments, while

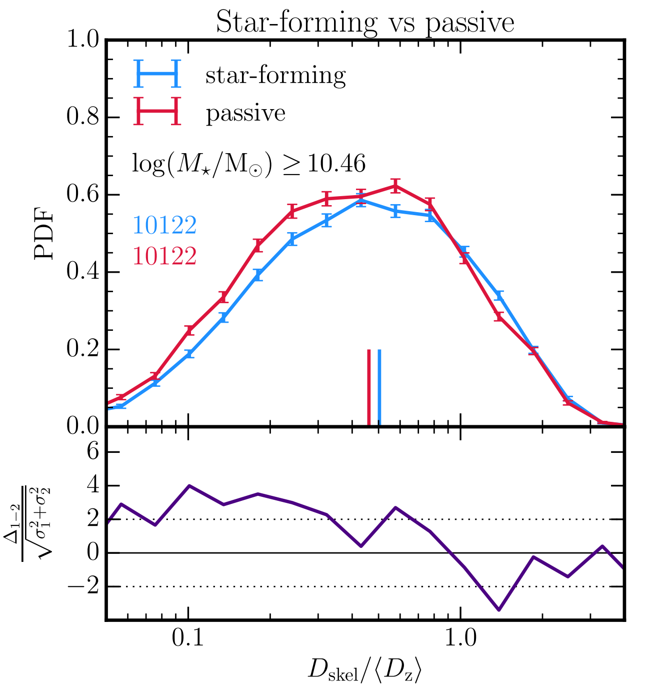


Figure 7. *Top:* Differential distributions of the distances to the nearest filament, $D_{\rm skel}$ (normalized by $\langle D_z\rangle$, the redshift-dependent mean inter-galaxy separation) for star-forming and quiescent galaxies that have been matched in mass (see the text for details). To highlight an effect specific to the filaments, the contribution of node is minimized (see the text for details). The vertical lines indicate the medians of the distributions and their values, together with associated error bars, are listed in Table 1. The numbers of galaxies in different considered bins are indicated in each panel. The error bars are calculated from 100 bootstrap samples. Galaxies are found to segregate, relative to filaments, according to their type: quiescent galaxies tend to be preferentially located closer to the filaments compared to their star-forming counterparts. *Bottom:* Residuals in units of σ between the star-forming and passive galaxies.

the less massive ones tend to reside preferentially on their outskirts (Laigle et al. 2017; Malavasi et al. 2017). As the impact of the nodes has been minimized, it is therefore established that this stellar mass gradient is driven by the filaments themselves and not by the densest regions of the cosmic web.

The mass segregation is also found among the star-forming population alone (right-hand panels), such that more massive star-forming galaxies tend to be closer to the geometric core of the filament than their less massive counterparts. Note that the mass bins for star-forming galaxies differ from mass bins used for the entire population. The completeness stellar mass limit allows us to decrease the lowest mass bin to $\log(M_\star/\mathrm{M}_\odot) = 9.92$ when considering the star-forming galaxies alone (see Section 2). The significance of these stellar mass gradients between the extreme stellar mass bins exceeds 4σ near the filaments, while the difference of the medians reaches an $\sim 8\sigma$ level (see Table 1).

4.2.2 *Type gradients*

Let us now investigate the impact of the filamentary network on the type/colour of galaxies. To do so, galaxies are split by type between star-forming and passive galaxies based on the dust corrected $u - r$ colour as discussed in Section 2.1. As for the analysis of the red fraction (Section 4.1), the sample is restricted to galaxies with $\log(M_\star/\mathrm{M}_\odot) \geq 10.46$ and the star-forming and passive populations are matched in stellar mass. Fig. 7 shows the PDFs of

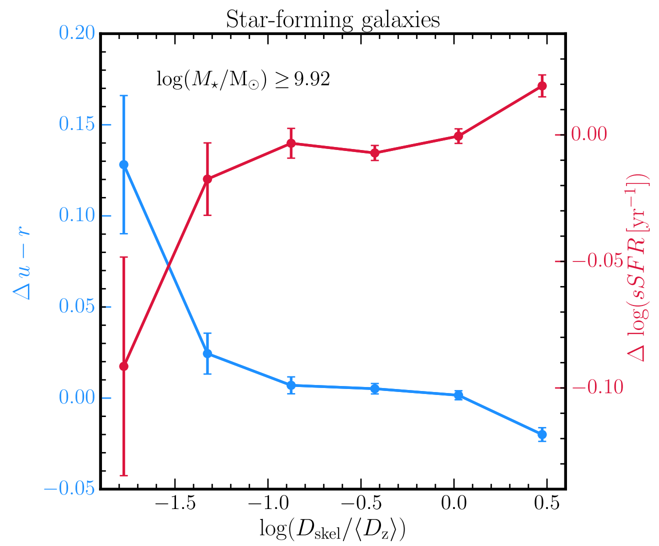


Figure 8. $u - r$ colour (blue line) and sSFR (red line) of star-forming galaxies as a function of $D_{\rm skel}$. The y-axes indicate the amount by which $u - r$ colour and sSFR differ from the median values at given mass (see the text for details). Only galaxies with $\log(M_\star/\mathrm{M}_\odot) \geq 9.92$ and far-away from nodes (at $D_{\rm node} > 3.5$ Mpc) are considered. Star-forming galaxies tend to have higher $u - r$ colour (tend to be redder) and lower sSFR when they get closer to the filaments than their more distant counterparts.

the normalized distances $D_{\rm skel}$ within the mass-matched samples of star-forming and passive populations, which by construction have the same number of galaxies. Galaxies are found to segregate according to their type such that passive galaxies tend to reside in regions located closer to the core of filaments than their star-forming counterparts. The significance of the type gradients between the two populations exceeds 3σ near filaments while the difference between the medians reaches an $\sim 4\sigma$ level (see Table 1).

4.2.3 *Star formation activity gradients*

To explore whether the impact of filaments on the star formation activity of galaxies can be detected beyond the red fractions and type segregation reported above, the focus is now on the star-forming population alone through the study of their (dust corrected) $u - r$ colour and sSFR.

Both these quantities are known to evolve with stellar mass which itself varies within the cosmic web (see above). To remove this mass dependence, the offsets of $u - r$ colour and sSFR, $\Delta u - r$ and ΔsSFR, respectively, from the median values of all star-forming galaxies at a given mass are computed for each galaxy. Fig. 8 shows the medians of $\Delta u - r$ and ΔsSFR as a function of $D_{\rm skel}$. Both quantities are found to carry the imprint of the large-scale environment. At large distances from the filaments ($D_{\rm skel} \geq 5$ Mpc), star-forming galaxies are found to be more active than the average. At intermediate distances ($0.5 \leq D_{\rm skel} \leq 5$ Mpc), star formation activity of star-forming galaxies does not seem to evolve with the distance to the filaments, while in the close vicinity of the filaments ($D_{\rm skel} \leq 0.5$ Mpc), they show signs of a decrease in star formation efficiency (redder colour and lower sSFR). The significance of these results will be discussed in Section 7.

4.3 The role of walls in mass and type gradients

Let us now investigate the impact of walls on galaxy properties. Figs 9 and 10 show the PDFs of the distances to the closest wall $D_{\rm wall}$ for the same selections as in Figs 6 and 7, respectively. The distances

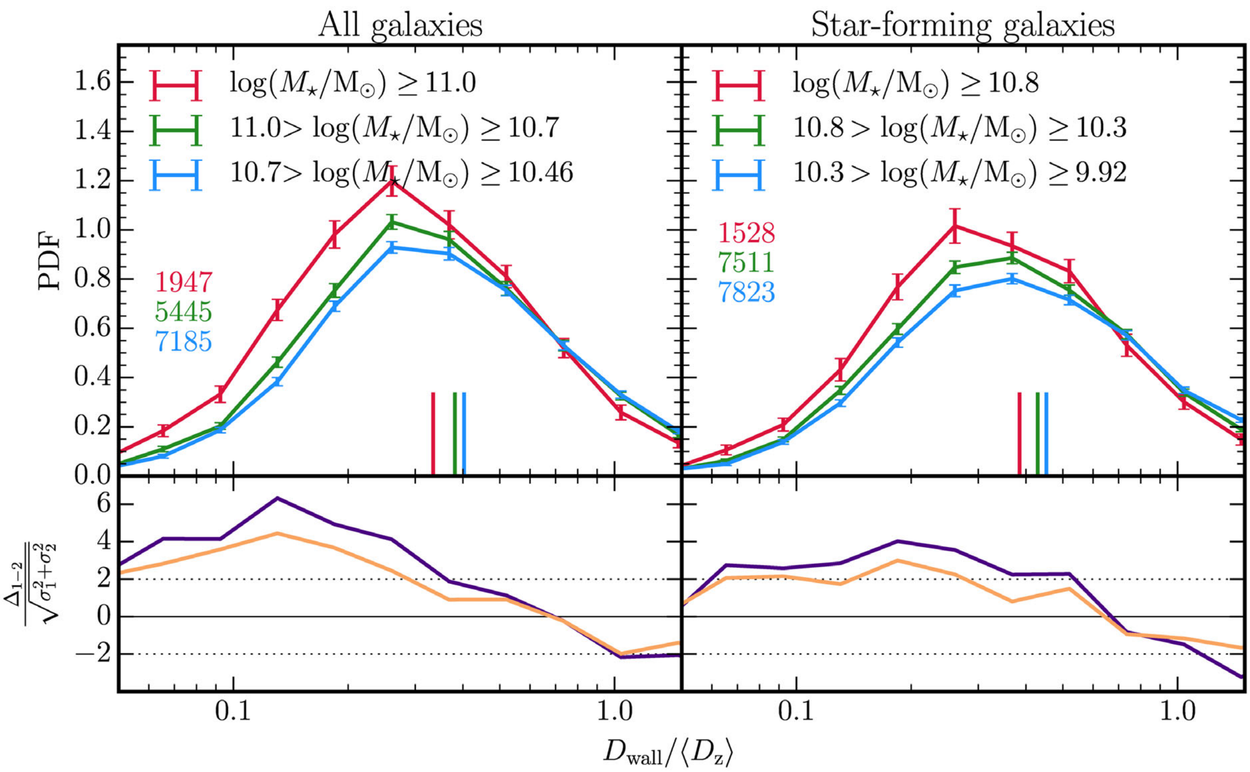


Figure 9. *Top row:* As in Fig. 6, but for the distances to the nearest wall, D_wall. To minimize the contribution of nodes and filaments to the measured signal, galaxies located closer to a node than 3.5 Mpc and closer to a filament than 2.5 Mpc are removed form the analysis. There is a mass segregation of galaxies with respect to walls of the entire as well as star-forming population: more massive galaxies tend to be preferentially located closer to the filaments compared to their lower-mass counterparts. *Bottom row:* Residuals in units of σ as in Fig. 6.

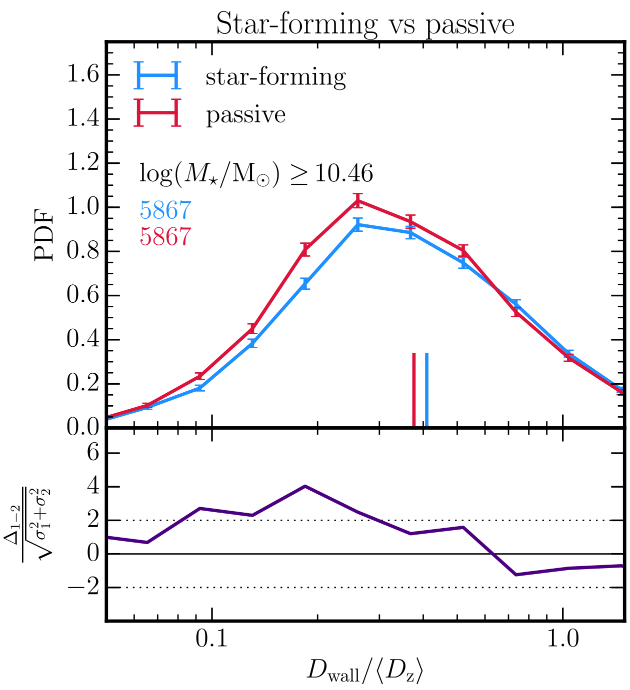


Figure 10. *Top row:* As in Fig. 7, but for the distances to the nearest wall, D_wall. To minimize the contribution of nodes and filaments to the measured signal, galaxies located closer to a node than 3.5 Mpc and closer to a filament than 2.5 Mpc are removed from the analysis. Galaxies are found to segregate, with respect to walls, according to their type: quiescent galaxies tend to be preferentially located closer to the walls compared to their star-forming counterparts. *Bottom row:* Residuals in units of σ as in Fig. 7.

are again normalized by the redshift-dependent mean inter-galaxy separation $\langle D_z\rangle$. The values of medians with corresponding error bars are listed in Table 1. As for filaments, one seeks signatures induced by a particular environment solely, walls in this case. Given that filaments are located at the intersections between walls, in addition to the contamination by nodes, which is of concern for filaments, one has to make sure that the contribution of filaments themselves is minimized as well. Following the method adopted in Section 4.2.1, Appendix B2 shows that this can be achieved by removing from the analysis galaxies having distances to the nodes smaller than 3.5 Mpc and distances to the closest filaments less than 2.5 Mpc.

The derived trends are qualitatively similar to those measured with respect to filaments. Massive galaxies are located closer to walls compared to their low-mass counterparts; star-forming galaxies preferentially reside in the outer regions of walls; and mass segregation is present also among star-forming population of galaxies with more massive star-forming galaxies having smaller distances to the walls than low-mass counterparts. Since these walls typically embed smaller-scale filaments, the net effect of transverse gradients perpendicular to these filaments should add up to transverse gradients perpendicular to walls.

The significance of the measured trends, in terms of the residuals between medians (see Table 1), is above 3σ for all considered gradients, slightly lower than for the gradients towards filaments. The deviations of $\sim 10\sigma$ and $\sim 5\sigma$ are detected between the highest and lowest stellar mass bins among the whole and star-forming population alone, respectively, while between the star-forming and passive galaxies it reaches $\sim 4\sigma$, as in the case of gradients towards filaments.

Table 2. Medians for the PDFs displayed in Fig. 11.

Selection[a]	Bin	Median[b]	
		$D_{\rm skel}$ (Mpc)	$D_{\rm wall}$ (Mpc)
	$\log(M_\star/{\rm M}_\odot) \geq 10.8$	1.34 ± 0.09	0.79 ± 0.04
Mass	$10.8 > \log(M_\star/{\rm M}_\odot) \geq 10.4$	1.73 ± 0.08	1.14 ± 0.03
	$10.4 > \log(M_\star/{\rm M}_\odot) \geq 10$	1.97 ± 0.04	1.22 ± 0.02
	$-10.8 > \log(sSFR/{\rm yr}^{-1})$	1.46 ± 0.07	1.02 ± 0.03
sSFR[c]	$-10.4 > \log(sSFR/{\rm yr}^{-1}) \geq -10.8$	1.88 ± 0.06	1.18 ± 0.03
	$\log(sSFR/{\rm yr}^{-1}) \geq -10.4$	2.0 ± 0.04	1.18 ± 0.02

Notes. [a]Panels of Fig. 11.
[b]Medians of distributions as indicated in Fig. 11 by vertical lines; errors are computed as in Table 1.
[c]Only galaxies with stellar masses $\log(M_\star/{\rm M}_\odot) \geq 10$ are considered.

5 COMPARISON WITH THE HORIZON-AGN SIMULATION

In this section, a qualitative support for the results on the mass and star-formation activity segregation is provided via the analysis of the large-scale cosmological hydrodynamical simulation HORIZON-AGN (Dubois et al. 2014). Note that the main purpose of such an analysis is to provide a reference measurement of gradients in the context of a large-scale 'full physics' experiment. The construction of the GAMA-like mock catalogue is not performed because the geometry of HORIZON-AGN does not allow us to recover the entire GAMA volume and the flux-limited sample requires a precise modelling of fluxes in different bands.

A brief summary of some of the main features of the simulation can be found in Appendix D. Here, the results on the mass and sSFR gradients towards filaments and walls are presented. The HORIZON-AGN simulation is analysed at low redshift ($z \sim 0.1$), comparable to the mean redshift studied in this paper, and the same analysis is performed as in the GAMA data. The filamentary network and associated structures are extracted by running the DisPerSE code with the persistence threshold of 3σ.

Fig. 11 shows the mass (left-hand panels) and sSFR (right-hand panels) gradients towards filaments (figure a) and walls (figure b) as measured in the HORIZON-AGN simulation. The impact of the nodes and filaments on the measured signal is minimized by removing from the analysis galaxies that are closer to the node than 3.5 Mpc and closer to the filament than 1 Mpc. The detailed description of the method used to identify these cuts in distances can be found in Appendix B1. Consistently with the measurements in GAMA, galaxies in HORIZON-AGN are found to segregate by stellar mass, with more massive galaxies being preferentially closer to both the filaments and walls than their low-mass counterparts. Similarly, the presence of the sSFR gradient, whereby less star-forming galaxies tend to be closer to the cores of filaments and walls than their more star-forming counterparts, is in qualitative agreement with the type/colour gradients detected in the GAMA survey. Note that the three bins of sSFR are used to separate out the highly star-forming galaxies, with $\log({\rm sSFR/yr}^{-1}) \geq -10.4$, from passive ones, with $\log({\rm sSFR/yr}^{-1}) < -10.8$, in order to compare with the type gradients in the observations. In the simulation, sSFR is a more reliable parameter for type than for the colour.

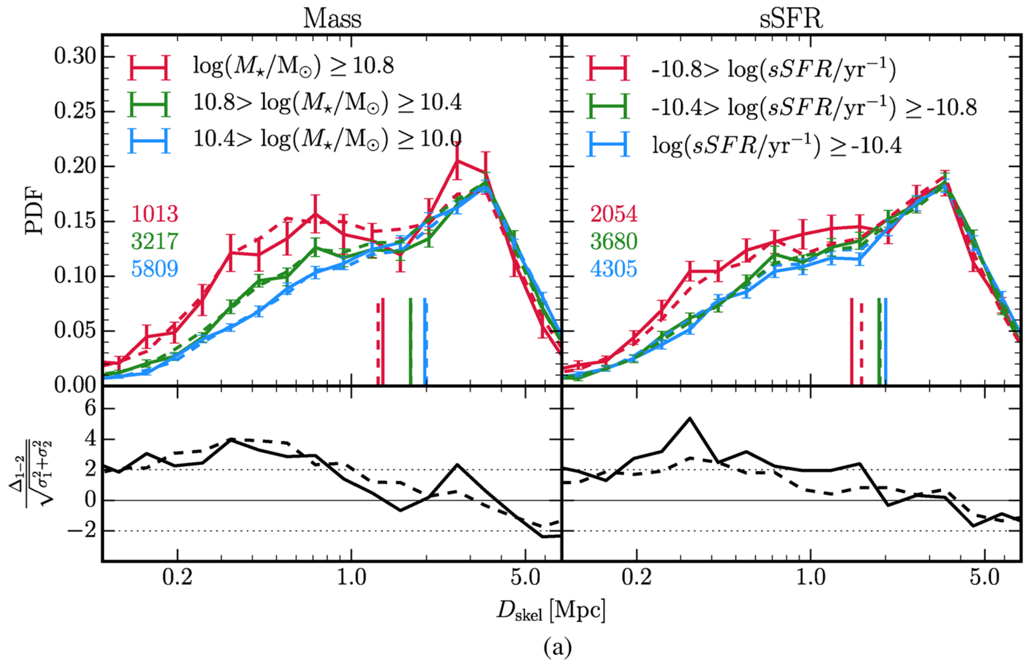
The significance of the trends is measured, as previously, in terms of the residuals between medians (see Table 1). For the gradients towards filaments, the difference of $\gtrsim 6\sigma$ is found between the most extreme, both mass and sSFR, bins, while it drops to $\sim$2–3σ between the intermediate and lowest bins. For the gradients towards walls, the deviation between the most extreme bins is $\sim$10 and 4σ for mass and sSFR bins, respectively, while there is only a little to no difference between intermediate and lowest stellar mass and sSFR bins, respectively. The gradients are slightly less significant than in the GAMA measurements, most likely due to the low numbers of galaxies per individual bins in HORIZON-AGN, but qualitatively similar as in GAMA.

6 THE RELATIVE IMPACT OF DENSITY

Let us now address the following questions: what is the specific role of the geometry of the large-scale environment in establishing mass and type/colour large-scale gradients? Are these gradients driven solely by density, or does the large-scale anisotropy of the cosmic web provide a specific signature?

A key ingredient in answering these questions is the choice of the scale at which the density is inferred. The properties of galaxies at a given redshift are naturally a signature of their past light-cone. This light-cone in turn correlates with the galaxy's environment: the larger the scale is, the longer the look-back time one must consider, the more integrated the net effect of this environment. This past environment accounts for the total accreted mass of the galaxy, but may also impact the geometry of the accretion history and more generally other galactic properties such as its star formation efficiency, its colour or its spin. At small scales, the density correlates with the most recent and stochastic processes, while going to larger scales allows taking the integrated hence smoother history of galaxies into account. Since this study is concerned about the statistical impact of the large-scale structure on galaxies, it is natural to consider scales large enough to average out local recent events they may have encountered, such as binary interactions, mergers and outflows. Therefore in the discussion below, the density is computed at the scale of 8 Mpc, the 'smallest' scale at which the effect of the anisotropic large-scale tides can be detected.

In practice, in order to try to disentangle the effect of density from that of the anisotropic large-scale tides, the following reshuffling method (e.g. Malavasi et al. 2017) is adopted. For mass gradients, 10 equipopulated density bins are constructed and in each of them the stellar masses of galaxies are randomly permuted. By construction, the underlying mass-density relation is preserved, but this procedure randomizes the relation between the stellar mass and the distance to the filament or the wall. For the type/colour gradients, in each of 10 equipopulated density bins, 10 equipopulated stellar mass bins are constructed. Within each of such bins, $u - r$ colour of galaxies are randomly permuted. Thus by construction, this preserves the underlying colour-(mass)-density relation, but breaks the relation between the colour/type and the distance to the particular environment, the filament or wall.



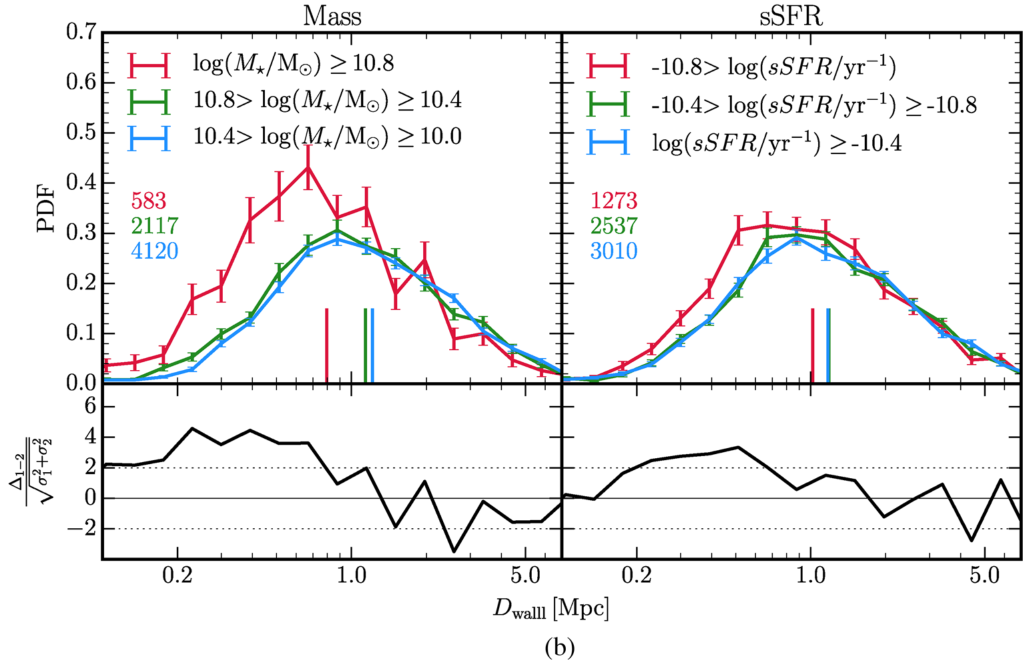
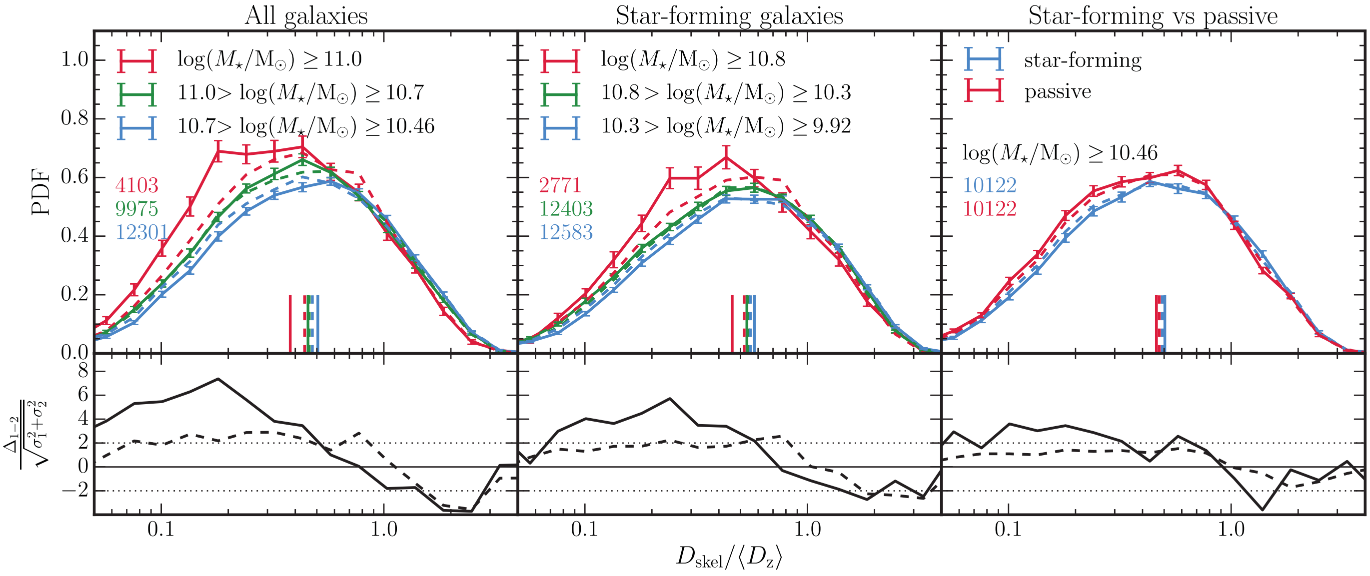


Figure 11. *Top rows:* Differential distributions of the distances as a function of stellar mass (left-hand panels) and sSFR (right-hand panels) for galaxies in HORIZON-AGN. To minimize the contribution of nodes and filaments to the measured signal, galaxies located closer to a node than 3.5 Mpc and closer to a filament than 1 Mpc are removed from the analysis. The vertical lines indicate the medians of the distributions (see Table 2 for the numerical values). Numbers of galaxies in different considered bins are indicated in each panel. There is mass and sSFR segregation of galaxies with respect to both filaments and walls: more massive and less star-forming galaxies tend to be preferentially located closer to the cores of filaments and walls compared to their lower-mass and more star-forming counterparts, respectively. These results are in qualitative agreement with the measurements in GAMA. *Bottom rows:* Residuals in units of σ between the two most extreme mass and sSFR bins, $\log(M_\star/\mathrm{M}_\odot) \geq 10.8$ and $10.4 > \log(M_\star/\mathrm{M}_\odot) \geq 10$ on the left-hand panel and $-10.8 > \log(\mathrm{sSFR/yr}^{-1})$ and $\log(\mathrm{sSFR/yr}^{-1}) \geq -10.4$ on the right-hand panel, respectively. (a) Differential distributions of the distances to the nearest filament, $D_{\rm skel}$. (b) Differential distributions of the distances to the nearest wall, $D_{\rm wall}$.

In order to account for the variation of the density through the survey, the density contrast, defined as $1 + \delta = n/n(z)$, where $n(z)$ corresponds to the mean redshift-dependent number density, is used in logarithmic bins. The number density n is computed in the Gaussian kernel and every time five reshuffled samples are constructed.

In Fig. 12(a), the mass and type gradients towards filaments, as measured in GAMA and previously shown in Figs 6 and 7, are compared with the outcome of the reshuffling technique. The original signal is found to be substantially reduced after the reshuffling of masses and colours of galaxies. For the mass gradients, the deviation between the most extreme bins before reshuffling exceeds 3σ, while after the reshuffling, the signal gets reduced, with typical deviations of $\sim 1\sigma$. The original signal for the type/colour gradients is weaker than in the case of the mass gradients, however it is similarly nearly cancelled out once the reshuffling method is



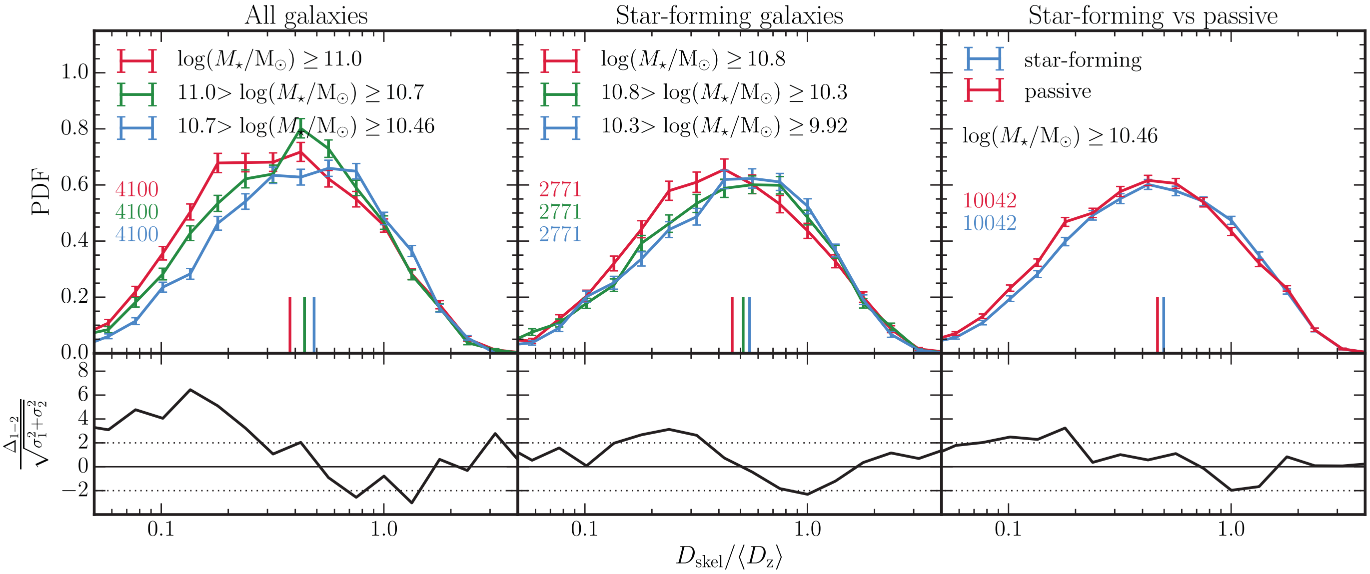


Figure 12. *Top rows:* Differential distributions of the normalized distances to the nearest filament, D_{skel} as a function of stellar mass of the entire galaxy population (left-hand panels), for star-forming galaxies only (middle panels) and as a function of galaxy's type (right-hand panels) with reshuffling (Figure a) and with density-matched samples (Figure b). In Figure (a), the distributions before applying the reshuffling method (solid lines) are compared to the results after the reshuffling (dashed lines). Figure (b) illustrates the distributions for the galaxy samples that are matched so that their density distributions are the same (see the text for details on the matching). The density estimators used in both the reshuffling and density matching corresponds to the (large-scale) density computed in the Gaussian kernel at the scale of 8 Mpc. As previously, the contribution of nodes to the measured signal is minimized. The numerical values of medians, shown as vertical lines, are listed in Table 3. The two methods yield qualitatively similar result: on the one hand when the large-scale density is used in reshuffling, the signal is reduced (dashed lines, Figure a) suggesting that the measured gradients (solid lines, Figure a) are not driven by the density at this scale, on the other hand, the gradients are measured within the samples that are matched in density at large scale. *Bottom rows:* Residuals in units of σ between the highest and lowest mass bins (left-hand and middle panels) and between the star-forming and passive galaxies (right-hand panels). (a) Reshuffling. (b) Density matching.

applied. The values of medians of the distributions after the reshuffling can be found in Table 3. Qualitatively similar behaviour is obtained for the gradients towards walls (not shown here). The analysis in HORIZON-AGN provides a qualitative support for these results. In Appendix D2, Fig. D1(a), the same reshuffling method is applied to simulated galaxies. The density used for this test is computed in the Gaussian kernel at 5 Mpc. This scale corresponds to the $\sim$1.5$\times$ mean inter-galaxy separation in HORIZON-AGN, consistently with the GAMA data.

Alternatively, to assess the impact of the density on the measured gradients within the cosmic web, one may want to use density matching. The purpose of this method is to construct mass- and colour-density matched samples, whereby galaxies with different masses and/or colours have similar density distributions, in order to make sure that the measured properties are not driven by their differences (see Appendix A2 for details on the matching technique). As shown in Fig. 12(b), the main result on the density-matching technique leads to the same conclusions as the reshuffling method.

Table 3. Medians for the PDFs displayed in Fig. 12: large-scale density

	Selection[a]	Bin	Median[b] Original[c]	$D_{\rm skel}/\langle D_z\rangle$ reshuffling[d]	Matching[e]
Masses	All galaxies	$\log(M_\star/\mathrm{M}_\odot) \geq 11$	0.379 ± 0.009	0.441 ± 0.009	0.379 ± 0.01
		$11 > \log(M_\star/\mathrm{M}_\odot) \geq 10.7$	0.456 ± 0.007	0.463 ± 0.006	0.44 ± 0.009
		$10.7 > \log(M_\star/\mathrm{M}_\odot) \geq 10.46$	0.505 ± 0.007	0.475 ± 0.006	0.486 ± 0.01
	SF galaxies	$\log(M_\star/\mathrm{M}_\odot) \geq 11$	0.459 ± 0.01	0.541 ± 0.015	0.459 ± 0.011
		$11 > \log(M_\star/\mathrm{M}_\odot) \geq 10.4$	0.534 ± 0.007	0.543 ± 0.007	0.514 ± 0.012
		$10.4 > \log(M_\star/\mathrm{M}_\odot) \geq 9.92$	0.578 ± 0.007	0.552 ± 0.007	0.549 ± 0.012
Types	SF versus passive[f]	Star-forming	0.503 ± 0.007	0.491 ± 0.007	0.498 ± 0.007
		Passive	0.462 ± 0.007	0.476 ± 0.007	0.467 ± 0.006

Notes. [a]Panels of Fig. 12.
[b]Medians of distributions as indicated in Fig. 12 by vertical lines; errors are computed as in Table 1.
[c]As in Table 1 for $D_{\rm skel}/\langle D_z\rangle$.
[d]Reshuffling is done in bins of density computed at 8 Mpc (see the text for details).
[e]Medians for the density-matched sample, where the density considered is computed at 8 Mpc.
[f]Only galaxies with stellar masses $\log(M_\star/\mathrm{M}_\odot) \geq 10.46$ are considered.

After matching galaxy populations in the large-scale density, mass and type gradients towards filaments and walls are still detected, suggesting that beyond the density, large-scale structures of the cosmic web do impact these galactic properties.

7 DISCUSSION

Let us first discuss the observational findings of the previous section in the framework of existing work (Section 7.1) and then focus on a recent extension of anisotropic excursion set which is developed in the companion paper (Section 7.2). The latter will allow us to explain why colour gradients prevail at fixed density.

7.1 Cosmic web metric: expected impact on galaxy evolution

In the current framework for galaxy formation, in which galaxies reside in extended dark matter haloes, it is quite natural to split the environment into the *local* environment, defined by the dark matter halo and the *global* large-scale anisotropic environment, encompassing the scale beyond the halo's virial radius. The anisotropy of the cosmic web is already a direct manifestation of the generic anisotropic nature of gravitational collapse on larger scales. It provides the embedding in which dark haloes and galaxies grow via accretion, which will act upon them via the combined effect of tides, the channeling of gas along preferred directions and angular momentum advection on to forming galaxies.

The observations and simulations presented in Sections 4, 5 and 6 provide a general support for this scenario. While rich clusters and massive groups are known to be environments which induce major galaxy transformations, the red fraction analysis presented in Section 4.1 (Fig. 5) reveals that the fraction of passive galaxies in the filaments starts to increase several Mpc away from the nodes and peaks in the nodes. This gradual increase suggests that some 'pre-processing' already happens before the galaxies reach the virial radius of massive haloes and fall into groups or clusters (e.g. Porter et al. 2008; Martínez et al. 2016). The above-mentioned morphological transformation of elliptical-to-spiral ratio when getting closer to the filaments (see also Kuutma et al. 2017) can be interpreted as the result of mergers transforming spirals into passive elliptical galaxies along the filaments when migrating towards nodes as suggested by theory and simulations (Codis et al. 2012; Dubois et al. 2014). These findings show that filamentary regions, corresponding to intermediate densities, are important environments for galaxy transformation. This is also confirmed by the segregation found in Sections 4.2 (Figs 6 and 7). More massive and/or passive galaxies are found closer to the core of filaments than their less massive and/or star-forming counterparts. These differential mass gradients persist among the star-forming population alone. In addition to mass segregation, star-forming galaxies show a gradual evolution in their star formation activity (see Fig. 8). They are bluer than average at large distances from filaments ($D_{\rm skel} \gtrsim 5$ Mpc), in a 'steady state' with no apparent evolution in star formation activity at intermediate distances ($0.5 \leq D_{\rm skel} \leq 5$ Mpc) and they show signs of decreased star formation efficiency near the core of the filaments ($D_{\rm skel} \lesssim 0.5$ Mpc). These results are in line with the picture where on the one hand more massive/passive galaxies lay in the core of filaments and merge while drifting towards the nodes of the cosmic web. On the other hand, the low mass/star-forming galaxies tend to be preferentially located in the outskirts of filaments, a vorticity-rich regions (Laigle et al. 2015), where galaxies acquire both their angular momentum (leading to a spin parallel to the filaments) and their stellar mass via essentially smooth accretion (Dubois et al. 2012b; Welker et al. 2017, also relying on HORIZON-AGN). The steady state of star-formation in these regions can reflect the right balance between the consumption and refuelling of the gas reservoir by the cold gas controlled by their surrounding filamentary structure (as shown by Codis et al. 2015, following Pichon et al. 2011, the outskirts of filaments are the loci of most efficient helicoidal infall of cold gas). This may not be true anymore when galaxies fall in the core of the filaments. The decline of star formation activity can, in part, be due to the higher merger rate but also due to a quenching process such as strangulation, where the supply of cold gas is halted (Peng, Maiolino & Cochrane 2015). It could also find its origin in the cosmic web detachment (Aragon-Calvo et al. 2016), where the turbulent regions inside filaments prevent galaxies to stay connected to their filamentary flows and thus to replenish their gas reservoir.

7.2 Link with excursion set theory

The distinct transverse gradients found for mass, density and type or colour may also be understood within the framework of conditional excursion set theory. Qualitatively, the spatial variation of the (traceless part of the) tidal tensor in the vicinity of filaments will delay infall on to galaxies, which will impact differentially galactic colour (at fixed mass), provided accretion can be reasonably converted into star formation efficiency.

7.2.1 Connecting gradients to constrained excursion set

The companion paper (Musso et al. 2017) revisits excursion set theory subject to conditioning the excursion to the vicinity of a filament. In a nutshell, the main idea of excursion set theory is to compute the statistical properties of the initial (over)density – a stochastic variable – enclosed within spheres of radius R, the scale which, through the spherical collapse model, can be related to the final mass of the object (should the density within the sphere pass the threshold for collapse). Increasing the radius of the sphere provides us with a proxy for 'evolution' (larger sphere, larger mass, smaller variance, later formation time) *and* a measure of the impact of the environment (different sensitivity to tides for different, larger, spheres). The expectations associated with this stochastic variable can be re-computed subject to the tides imposed by larger scale structures, which are best captured by the geometry of a filament-saddle point, $\mathcal{S}$, providing the local natural 'metric' for a filament (Codis et al. 2015). These large-scale tides will induce distinct weighting in the conditional PDF(δ, $\partial_R\delta|\mathcal{S}$) for the overdensity δ, and its successive derivatives with respect to scale, $\partial_R\delta$ etc. (so as to focus on collapsed accreting regions). Indeed, the saddle will shift not only the mean expectation of the PDFs but also importantly their co-variances (see Musso et al. 2017, for details). The derived expected (dark matter) mean density $\rho(r, \theta, \phi)$, Press-Schechter mass $M(r, \theta, \phi)$ and typical accretion rate $\dot{M}(r, \theta, \phi)$ then become explicit *distinct* functions of distance r and relative orientation to the closest (oriented) saddle point. Within this model, it follows that the orientation of the mass, density and accretion rate gradients differ. The misalignment arises because the various fields weight differently the constrained tides, which will physically e.g. delay infall, and technically involve different moments of the aforementioned conditional PDF (see Appendix E for more quantitative information on contour misalignment). This is shown in Fig. 13, which displays a typical longitudinal cross-section of those three maps in the frame of the saddle (with the filament along the Oz axis) in Lagrangian space.[5]

This line of argument explains environmentally driven differential gradients, yet there is still a stretch to connect it to the observed gradients. While there is no obvious consensus on the detailed effect of large-scale (dark matter) accretion on to the colour or star formation of galaxies at fixed mass and density, one can expect that the stronger the accretion, the stronger the AGN feedback, the stronger the quenching. Should this (reasonable) scaling hold true, the net effect in terms of gradients would be that colour gradients differ from mass and density ones. This is qualitatively consistent with the findings of this paper.

[5] This companion paper does not capture the strongly non-linear process of dynamical friction of sub clumps within dark haloes, nor strong deviations from spherical collapse. We refer to Hahn et al. (2009), which captures the effect on satellite galaxies, and to Ludlow, Borzyszkowski & Porciani (2014), Castorina et al. (2016) and Borzyszkowski et al. (2017) which study the effect of the local shear on haloes forming in filamentary structures. This requires adopting a threshold for collapse that depends explicitly on the local shear. The shear-dependent part of the critical density (and its derivative) correlates with the shear of the saddle, and introduces an additional anisotropic effect on top of the change of mean values and variances of density and slope.

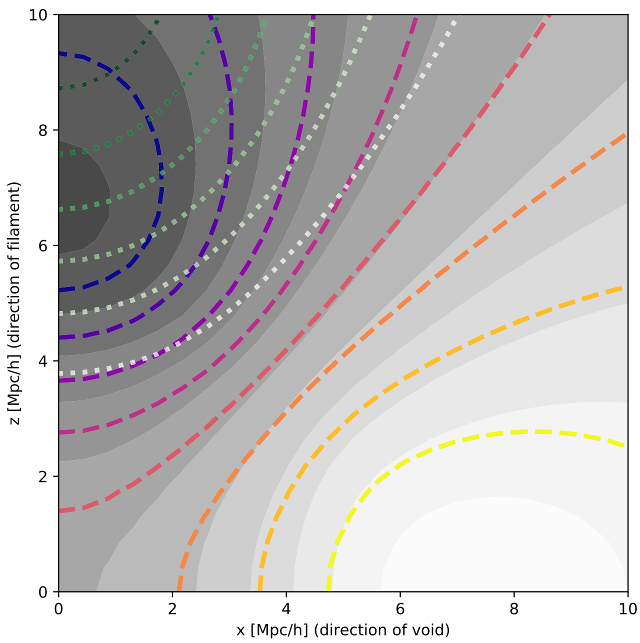


Figure 13. Isocontours of constant typical redshift $z = 0$ mean density (filled contours), mass (dotted lines) and accretion rate (dashed lines) in the frame of a filament (along the Oz axis) in Lagrangian space (initial conditions) from low (light colours) to high values (dark colours). The saddle is at coordinate (0,0) while the induced peak and void are at coordinates (0,$\pm$7) and ($\pm$8,0) h^{-1}Mpc, respectively. As argued in the main text, this figure shows that the contours, hence the gradients of the three fields, are not parallel (the contours cross). The choice of scale sets the units on the x- and z-axis (chosen here to be 5 h^{-1}Mpc, while the mass and accretion rates are computed for a local smoothing of 0.5 h^{-1}Mpc). At lower redshift/smaller scales, one expects the non-linear convergence of the flow towards the filament to bring those contours together, aligning the gradients (see Fig. 14).

7.2.2 Gradient alignments on smaller non-linear scales

The above-presented Lagrangian theory clearly applies only on sufficiently large scales so that dynamical evolution has not driven the large-scale flow too far from its initial configuration. On smaller scales, one would expect the same line of argument to operate in the frame set by the saddle smoothed on the corresponding scale, but with one extra caveat: the increased level of non-linearity will have compressed the local filament transversally and stretched it longitudinally, following the generic kinetic flow measured in N-body simulation (e.g. Sousbie et al. 2008a), or predicted at the level of the Zel'dovich approximation (Codis et al. 2015).

Consequently, the contours of constant dark matter density ρ, typical dark halo mass M and typical relative accretion rate $\dot{M}/M$ in the frame of the saddle shown in Fig. 13 will be driven more parallel to each other, hence the difference in the orientation of the density, mass and accretion gradient will become smaller and smaller as one considers smaller scales, and/or more non-linear dynamics (see Fig. 14). As colour gradient at fixed mass, and mass gradient at fixed density towards filaments originate from this initial misalignment, it should come as no surprise that as one probes smaller scales, such relative gradients disappear. When considering statistical expectations concerned with anisotropy (delayed accretion, acquisition of angular momentum, etc.), the net effect of past interactions should first be considered on the largest significant scale, beyond which the universe becomes isotropic. Conversely, the level of stochasticity should increase significantly on smaller scales, where one must

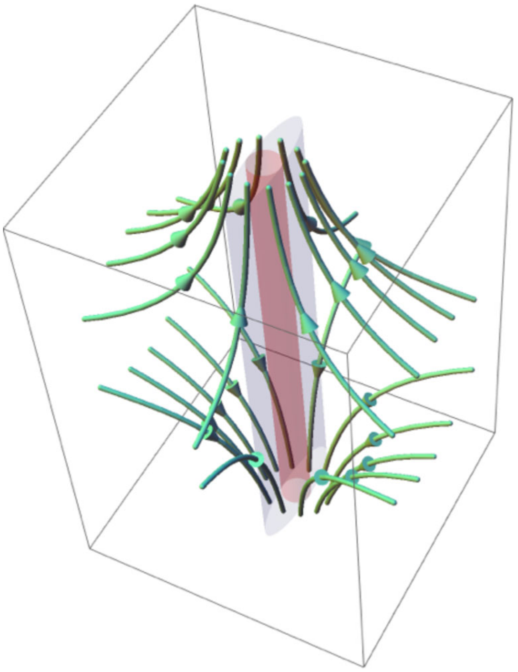


Figure 14. Illustration of the Zel'dovich flow (green arrows) in the vicinity of a filament (red cylinder) embedded in a wall (purple flattened cylinder), with filament saddle at the centre. The non-linear evolution operating more strongly on smaller scales will advect the contours presented in Fig. 13 along the green arrows, bringing them more parallel to each other. Consequently at these smaller scales, the mass and accretion gradients do not differ significantly from the density gradients. See Codis et al. (2015) and Musso et al. (2017) for more details.

account for, e.g. the configuration of the last merger event, or the last fly-by. Such a scenario is indeed supported by our findings in both GAMA and HORIZON-AGN, presented in Appendices C and D2, Figs C1 and D1, respectively, whereby the use of the small-scale density tracer does not allow us to disentangle between the effects of the local density and that of cosmic web, suggesting that at such scale, they are closely correlated through the small-scale processes.

7.2.3 Relationship to wall gradients

When measured relative to the walls, galaxy properties are found to exhibit the same trends as for filaments, in that more massive and/or quiescent galaxies are found closer to the walls than their low mass and/or star-forming counterparts. This result is again in qualitative agreement with the idea of walls being, together with the filaments, the large-scale interference patterns of primordial fluctuations capable of inducing anisotropic boost in overdensity together with the corresponding tides, and consequently imprinting their geometry in the measured properties of galaxies. The gradients measured for walls have the same origin as those inducing the differential gradients near the filament-type saddles, but are sourced by the geometry of the tides near the wall-type saddles (Codis et al. 2015, Appendix B). The main difference between the two saddles lies in the transverse curvatures, which is steeper for wall-type saddles than for filament-type saddles (when considering the mean, eigenvolume weighted, eigenvalues of the curvature tensor with the relevant signatures) leading to weaker differences between the different gradients when considering walls. This is consistent with the findings of Section 4.3.

In closing, note that the (resp. Eulerian and Lagrangian) interpretations presented in Sections 7.1 and 7.2 are complementary, but fall short in explaining in details the origin of quenching. Nevertheless, in view of both observation and theory, the cosmic web metric appears as a natural framework to understand galaxy formation beyond stellar mass and local density.

8 SUMMARY AND CONCLUSIONS

This paper studies the impact of the large-scale environment on the properties of galaxies, such as their stellar mass, dust corrected $u - r$ colour and sSFR. The discrete persistent structure extractor (DisPerSE) was used to identify the peaks, filaments and walls in the large-scale distribution of galaxies as captured by the GAMA survey. The principal findings are the following.

(i) *Mass segregation.* Galaxies are found to segregate by stellar mass, such that more massive galaxies are preferentially located closer to the cores of filaments than their lower mass counterparts. This mass segregation persists among the star-forming population. Similar mass gradients are seen with respect to walls in that galaxies with higher stellar mass tend to be found closer to the walls compared to galaxies with lower mass and persisting even when star-forming population of galaxies is considered alone.

(ii) *Type/colour segregation.* Galaxies are found to segregate by type/colour, with respect to both filaments and walls, such that passive galaxies are preferentially located closer to the cores of filaments or walls than their star-forming counterparts.

(iii) *Red fractions.* The fraction of passive galaxies increases with both decreasing distance to the filament and to the node, i.e. at fixed distance to the node, the relative number of passive galaxies (with respect to the entire population) increases as the distance to the filament decreases and similarly, at a given distance to the filament, this number increases with decreasing distance to the node.

(iv) *Star formation activity.* Star-forming galaxies are found to carry an imprint of large-scale environment as well. Their dust corrected $u - r$ and sSFR are found to be more enhanced and reduced, respectively, in the vicinity of the filaments compared to their outskirts.

(v) *Consistency with cosmological simulations.* All the found gradients are consistent with the analysis of the HORIZON-AGN 'full physics' hydrodynamical simulation. This agreement suggests that what drives the gradients is captured by the implemented physics.

(vi) *Connection to excursion set theory.* The origin of the distinct gradients can be qualitatively explained via conditional excursion set theory subject to filamentary tides (Musso et al. 2017).

This work has focused on filaments, nodes and in somewhat lesser details on walls. Similar observational results were recently reported at high redshift by using the cosmic web filamentary structures in the VIPERS spectroscopic survey (Malavasi et al. 2017) and while using projected filaments in photometric redshift slices in the COSMOS field (Laigle et al. 2017). These observations are of intrinsic interest as a signature of galactic assembly; they also comfort theoretical expectations which point towards distinct gradients for colour, mass and density with respect to the cosmic web. The tides of the large-scale environment play a significant specific role in the evolution of galaxies, and are imprinted in their integrated physical properties, which vary as a function of scale and distance to the different components of the cosmic web in a manner which is specific to each observable.

These observations motivate a theory which eventually should integrate the anisotropy of the cosmic web as an essential ingredient to (i) describe jointly the dynamics and physics of galaxies, (ii) explain galactic morphological diversity, and (iii) mitigate intrinsic alignment in upcoming lensing dark energy experiments, once a proper modelling of the mapping between galaxies and their haloes (allowing e.g. to convert the DM accretion rate into colour of galaxy) becomes available.

Future large-scale spectrographs on 8 metre class telescopes (MOONS;[6] Cirasuolo et al. 2014; Cirasuolo & MOONS Consortium 2016, PFS;[7] Sugai et al. 2015) or space missions (WFIRST;[8] Spergel et al. 2013, 2015, and Euclid;[9] Laureijs et al. 2011, the deep survey for the latter) will extend the current analysis at higher redshift ($z \geq 1$) with similar samplings, allowing us to explore the role of the environment near the peak of the cosmic star formation history, an epoch where the connectivity between the LSS and galaxies is expected to be even tighter, with ubiquitous cold streams. Tomography of the Lyman-α forest with PFS, MOONS, ELT-HARMONI (Thatte et al. 2010) tracing the intergalactic medium will make the study of the link between galaxies and this large-scale gas reservoir possible.

ACKNOWLEDGEMENTS

The authors thank the anonymous referee for suggestions and comments that helped to improve the presentation of the paper. This research is carried out within the framework of the Spin(e) collaboration (ANR-13-BS05-0005, http://cosmicorigin.org). We thank the members of this collaboration for numerous discussions. The HORIZON-AGN simulation was post processed on the Horizon Cluster hosted by Institut d'Astrophysique de Paris. We thank S. Rouberol for running it smoothly for us. GAMA is a joint European-Australasian project based around a spectroscopic campaign using the Anglo-Australian Telescope. The GAMA input catalogue is based on the data taken from the Sloan Digital Sky Survey and the UKIRT Infrared Deep Sky Survey. Complementary imaging of the GAMA regions is being obtained by a number of independent survey programmes including GALEX MIS, VST KiDS, VISTA VIKING, WISE, Herschel-ATLAS, GMRT and ASKAP providing UV to radio coverage. GAMA is funded by the STFC (UK), the ARC (Australia), the AAO, and the participating institutions. The GAMA web site is http://www.gama-survey.org/. The VISTA VIKING data used in this paper are based on observations made with ESO Telescopes at the La Silla Paranal Observatory under programme ID 179.A-2004. CC acknowledges support through the ILP PhD thesis fellowship. CL is supported by a Beecroft Fellowship.

REFERENCES

Alpaslan M. et al., 2015, MNRAS, 451, 3249
Alpaslan M. et al., 2016, MNRAS, 457, 2287
Andrae R., Jahnke K., 2011, MNRAS, 418, 2014
Aragon-Calvo M. A., Yang L. F., 2014, MNRAS, 440, L46
Aragón-Calvo M. A., Jones B. J. T., van de Weygaert R., van der Hulst J. M., 2007a, A&A, 474, 315
Aragón-Calvo M. A., van de Weygaert R., Jones B. J. T., van der Hulst J. M., 2007b, ApJ, 655, L5
Aragón-Calvo M. A., van de Weygaert R., Jones B. J. T., 2010, MNRAS, 408, 2163
Aragon-Calvo M. A., Neyrinck M. C., Silk J., 2016, preprint (arXiv:1607.07881)
Arnouts S., Cristiani S., Moscardini L., Matarrese S., Lucchin F., Fontana A., Giallongo E., 1999, MNRAS, 310, 540
Aubert D., Pichon C., Colombi S., 2004, MNRAS, 352, 376
Baldry I. K., Balogh M. L., Bower R. G., Glazebrook K., Nichol R. C., Bamford S. P., Budavari T., 2006, MNRAS, 373, 469
Bamford S. P. et al., 2009, MNRAS, 393, 1324
Barnes J., Efstathiou G., 1987, ApJ, 319, 575
Benítez-Llambay A., Navarro J. F., Abadi M. G., Gottlöber S., Yepes G., Hoffman Y., Steinmetz M., 2013, ApJ, 763, L41
Beygu B., Kreckel K., van der Hulst J. M., Jarrett T. H., Peletier R., van de Weygaert R., van Gorkom J. H., Aragon-Calvo M. A., 2016, MNRAS, 458, 394
Blanton M. R. et al., 2003, ApJ, 594, 186
Blumenthal G. R., Faber S. M., Primack J. R., Rees M. J., 1984, Nature, 311, 517
Bond J. R., Myers S. T., 1996, ApJS, 103, 1
Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440
Bond J. R., Kofman L., Pogosyan D., 1996, Nature, 380, 603
Bond N. A., Strauss M. A., Cen R., 2010a, MNRAS, 406, 1609
Bond N. A., Strauss M. A., Cen R., 2010b, MNRAS, 409, 156
Borzyszkowski M., Porciani C., Romano-Diaz E., Garaldi E., 2017, MNRAS, 469, 594
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Castorina E., Paranjape A., Hahn O., Sheth R. K., 2016, preprint (arXiv:1611.03619)
Cautun M. C., van de Weygaert R., 2011, Astrophysics Source Code Library, record ascl:1105.003
Cervantes-Sodi B., Hernandez X., Park C., 2010, MNRAS, 402, 1807
Chen Y.-C. et al., 2017, MNRAS, 466, 1880
Cirasuolo M. et al., 2014, Proc. SPIE, 9147, 91470N
Cirasuolo M., MOONS Consortium 2016, in Skillen I., Barcells M., Trager S., eds, ASP Conf. Ser. Vol. 507, Multi-Object Spectroscopy in the Next Decade: Big Questions, Large Surveys, and Wide Fields. Astron. Soc. Pac., San Francisco, p. 109
Codis S., Pichon C., Devriendt J., Slyz A., Pogosyan D., Dubois Y., Sousbie T., 2012, MNRAS, 427, 3320
Codis S., Pichon C., Pogosyan D., 2015, MNRAS, 452, 3369
Colless M., Dalton G., Maddox S., Sutherland W., Norberg P., Cole S., Bland-Hawthorn J., 2001, MNRAS, 328, 1039
Colombi S., Pogosyan D., Souradeep T., 2000, Phys. Rev. Lett., 85, 5515
Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371
De Lapparent V., Geller M. J., Huchra J. P., 1986, ApJ, 302, L1
Doroshkevich A. G., 1970, Astrofizika, 6, 581
Dressler A., 1980, ApJ, 236, 351
Dressler A. et al., 1997, ApJ, 490, 577
Driver S. P. et al., 2009, Astron. Geophys., 50, 5.12
Driver S. P., Hill D. T., Kelvin L. S., Robotham A. S. G., Liske J., Norberg P., Baldry I. K., 2011, MNRAS, 413, 971
Driver S. P. et al., 2016, MNRAS, 455, 3911
Dubois Y., Devriendt J., Slyz A., Teyssier R., 2012a, MNRAS, 420, 2662
Dubois Y., Pichon C., Haehnelt M., Kimm T., Slyz A., Devriendt J., Pogosyan D., 2012b, MNRAS, 423, 3616
Dubois Y. et al., 2014, MNRAS, 444, 1453
Eardley E., Peacock J. A., McNaught-Roberts T., Heymans C., Norberg P., Alpaslan M., Baldry, 2015, MNRAS, 448, 3665
Efstathiou G., Jones B. J. T., 1979, MNRAS, 186, 133
Efstathiou G., Frenk C. S., White S. D. M., Davis M., 1988, MNRAS, 235, 715
Eke V. R. et al., 2004, MNRAS, 348, 866
Forero-Romero J. E., Hoffman Y., Gottlöber S., Klypin A., Yepes G., 2009, MNRAS, 396, 1815
Gay C., Pichon C., Le Borgne D., Teyssier R., Sousbie T., Devriendt J., 2010, MNRAS, 404, 1801
Geller M. J., Huchra J. P., 1989, Science, 246, 897

[6] Multi-Object Optical and Near-infrared Spectrograph.
[7] Prime Focus Spectrograph; http://pfs.ipmu.jp/
[8] Wide-Field Infrared Survey Telescope; http://wfirst.gsfc.nasa.gov
[9] http://sci.esa.int/euclid/, http://www.euclid-ec.org

Goto T., Yamauchi C., Fujita Y., Okamura S., Sekiguchi M., Smail I., Bernardi M., Gomez P. L., 2003, MNRAS, 346, 601
Guo Q., Tempel E., Libeskind N. I., 2015, ApJ, 800, 112
Haardt F., Madau P., 1996, ApJ, 461, 20
Hahn O., Porciani C., Carollo C. M., Dekel A., 2007a, MNRAS, 375, 489
Hahn O., Carollo C. M., Porciani C., Dekel A., 2007b, MNRAS, 381, 41
Hahn O., Porciani C., Dekel A., Carollo C. M., 2009, MNRAS, 398, 1742
Hahn O., Angulo R. E., Abel T., 2015, MNRAS, 454, 3920
Hashimoto Y., Oemler Jr A., Lin H., Tucker D. L., 1998, ApJ, 499, 589
Heavens A., Peacock J., 1988, MNRAS, 232, 339
Hopkins A. M. et al., 2013, MNRAS, 430, 2047
Ilbert O. et al., 2006, A&A, 457, 841
Jackson J. C., 1972, MNRAS, 156, 1P
Jõeveer M., Einasto J., Tago E., 1978, MNRAS, 185, 357
Jones B. J. T., Efstathiou G., 1979, MNRAS, 189, 27
Jones B. J. T., van de Weygaert R., Aragón-Calvo M. A., 2010, MNRAS, 408, 897
Kaiser N., 1984, ApJ, 284, L9
Kaiser N., 1987, MNRAS, 227, 1
Kauffmann G., White S. D. M., Heckman T. M., Ménard B., Brinchmann J., Charlot S., Tremonti C., Brinkmann J., 2004, MNRAS, 353, 713
Kleiner D., Pimbblet K. A., Jones D. H., Koribalski B. S., Serra P., 2017, MNRAS, 466, 4692
Komatsu E. et al., 2011, ApJS, 192, 18
Kuutma T., Tamm A., Tempel E., 2017, A&A, 600, L6
Laigle C. et al., 2015, MNRAS, 446, 2744
Laigle C. et al., 2017, preprint (arXiv:1702.08810)
Laureijs R. et al., 2011, preprint (arXiv:1110.3193)
Lee J., 2004, ApJ, 614, L1
Lee J., Erdogdu P., 2007, ApJ, 671, 1248
Lewis I. et al., 2002, MNRAS, 334, 673
Libeskind N. I., Hoffman Y., Knebe A., Steinmetz M., Gottlöber S., Metuki O., Yepes G., 2012, MNRAS, 421, L137
Libeskind N. I. et al., 2017, preprint (arXiv:1705.03021)
Liske J. et al., 2015, MNRAS, 452, 2087
Ludlow A. D., Borzyszkowski M., Porciani C., 2014, MNRAS, 445, 4110
Malavasi N. et al., 2017, MNRAS, 465, 3817
Martínez H. J., Muriel H., Coenda V., 2016, MNRAS, 455, 127
Merson A. I. et al., 2013, MNRAS, 429, 556
Mo H., van den Bosch F. C., White S., 2010, Galaxy Formation and Evolution. Cambridge Univ. Press, Cmabridge
Musso M., Cadiou C., Pichon C., Codis S., Dubois Y., 2017, preprint (arXiv:1709.00834)
Navarro J. F., Abadi M. G., Steinmetz M., 2004, ApJ, 613, L41
Novikov D., Colombi S., Doré O., 2006, MNRAS, 366, 1201
Pahwa I. et al., 2016, MNRAS, 457, 695
Paz D. J., Stasyszyn F., Padilla N. D., 2008, MNRAS, 389, 1127
Peebles P. J. E., 1969, ApJ, 155, 393
Peng Y., Maiolino R., Cochrane R., 2015, Nature, 521, 192
Pichon C., Pogosyan D., Kimm T., Slyz A., Devriendt J., Dubois Y., 2011, MNRAS, 418, 2493
Planck Collaboration XIII, 2016, A&A, 594, A13
Pogosyan D., Bond J. R., Kofman L., Wadsley J., 1996, in American Astronomical Society Meeting Abstracts, p. 1289
Pogosyan D., Pichon C., Gay C., Prunet S., Cardoso J. F., Sousbie T., Colombi S., 2009, MNRAS, 396, 635
Porciani C., Dekel A., Hoffman Y., 2002a, MNRAS, 332, 325
Porciani C., Dekel A., Hoffman Y., 2002b, MNRAS, 332, 339
Porter S. C., Raychaudhury S., Pimbblet K. A., Drinkwater M. J., 2008, MNRAS, 388, 1152
Postman M., Geller M. J., 1984, ApJ, 281, 95
Robotham A. S. G. et al., 2011, MNRAS, 416, 2640
Robotham A. S. G. et al., 2013, MNRAS, 431, 167
Rojas R. R., Vogeley M. S., Hoyle F., Brinkmann J., 2004, ApJ, 617, 50
Schaap W. E., van de Weygaert R., 2000, A&A, 363, L29
Schmidt M., 1959, ApJ, 129, 243
Shi F. et al., 2016, ApJ, 833, 241
Sousbie T., 2011, MNRAS, 414, 350
Sousbie T., Pichon C., Colombi S., Novikov D., Pogosyan D., 2008a, MNRAS, 383, 1655
Sousbie T., Pichon C., Courtois H., Colombi S., Novikov D., 2008b, ApJ, 672, L1
Sousbie T., Pichon C., Kawahara H., 2011, MNRAS, 414, 384
Spergel D. et al., 2013, preprint (arXiv:1305.5422)
Spergel D. et al., 2015, preprint (arXiv:1503.03757)
Stewart K. R., Brooks A. M., Bullock J. S., Maller A. H., Diemand J., Wadsley J., Moustakas L. A., 2013, ApJ, 769, 74
Subba Rao M. U., Aragón-Calvo M. A., Chen H. W., Quashnock J. M., Szalay A. S., York D. G., 2008, New J. Phys., 10, 125015
Sugai H. et al., 2015, J. Astron. Telescopes Instrum. Syst., 1, 035001
Sutherland R. S., Dopita M. A., 1993, ApJS, 88, 253
Tegmark M. et al., 2004, ApJ, 606, 702
Tempel E., Libeskind N. I., 2013, ApJ, 775, L42
Tempel E., Stoica R. S., Saar E., 2013, MNRAS, 428, 1827
Teyssier R., 2002, A&A, 385, 337
Thatte N. et al., 2010, in Proc. SPIE Conf. Ser. Vol. 77352I, Ground-based and Airborne Instrumentation for Astronomy III. SPIE, Bellingham, p. 77352I
Tojeiro R. et al., 2017, MNRAS, 470, 3720
Trowland H. E., Lewis G. F., Bland-Hawthorn J., 2013, ApJ, 762, 72
Trujillo I., Carretero C., Patiri S. G., 2006, ApJ, 640, L111
Tully R. B., Fisher J. R., 1978, in Longair M. S., Einasto J., eds, IAU Symp. Vol. 79, Large Scale Structures in the Universe. Springer, Dordrecht, p. 31
Tweed D., Devriendt J., Blaizot J., Colombi S., Slyz A., 2009, A&A, 506, 647
Welker C., Devriendt J., Dubois Y., Pichon C., Peirani S., 2014, MNRAS, 445, L46
Welker C., Dubois Y., Pichon C., Devriendt J., Chisari E. N., 2015, preprint (arXiv:1512.00400)
Welker C., Dubois Y., Devriendt J., Pichon C., Kaviraj S., Peirani S., 2017, MNRAS, 465, 1241
White S. D. M., 1984, ApJ, 286, 38
White S. D. M., Rees M. J., 1978, MNRAS, 183, 341
Wright A. H. et al., 2016, MNRAS, 460, 765
Zel'dovich Y. B., 1970, A&A, 5, 84

APPENDIX A: MATCHING TECHNIQUE

A1 Mass matching

This Appendix provides details on the mass matching procedure. First the mass distributions of the two populations are cut so that they cover the same stellar mass range, i.e. they have the same minimum and maximum value of stellar mass. Then, in each stellar mass bin, the population with lower number of galaxies is taken as the reference sample and $N_{\rm match}$ samples of galaxies are extracted in the other population, such that their mass distribution is the same as the one of the reference sample. In practice, for each galaxy in the reference sample, the corresponding galaxy of the larger sample is sought among galaxies whose mass difference with respect to the reference mass is smaller than $\Delta M_\star$ in logarithmic space. If there is no galaxy in larger sample satisfying this condition, the galaxy of the reference sample is removed from the analysis. In each of $N_{\rm match}$ samples, every galaxy of the larger sample is considered only once, however repetitions are allowed across all samples. By construction, after applying this procedure, one ends up with $N_{\rm match}$ samples consisting of the same number of star-forming and passive galaxies and having very similar stellar mass distributions.

If not stated differently, 20 mass-matched samples are typically constructed using 10 equipopulated stellar mass bins for each and choosing a value of 0.1 for $\Delta M_\star$ parameter. Varying the values of N_{match}, $\Delta M_\star$ and the number of stellar mass bins within the reasonable range does not alter our conclusions.

A2 Density matching

This Appendix provides details on the density matching procedure. First, let us describe how the mass-density matched samples are constructed. The galaxy sample is first divided into three logarithmic stellar mass bins for which the density matched samples are to be constructed. In each of the 10 equipopulated logarithmic overdensity $(1+\delta)$ bins, the reference sample is identified as that of the previously constructed stellar-mass subsamples with the lowest number of galaxies. Next, for each galaxy in the reference sample, a galaxy is randomly chosen from each of two stellar mass bins having the overdensity closest to the galaxy in the reference sample. In practice, the nominal absolute difference in the $\log(M_\star/\mathrm{M}_\odot)$ values used to match galaxies is 0.1. If no suitable galaxy is found in at least one of the two stellar mass bins, the galaxy of the reference sample is removed from the analysis. This procedure is repeated 10 times, ending up with 10 samples of galaxies having the same overdensity distributions in three different stellar mass bins.

Similarly, to construct type-density matched samples, the entire galaxy sample is first divided into the subsamples of star-forming and passive galaxies. Then, in each of the 10 equipopulated logarithmic overdensity $(1+\delta)$ bins, the reference sample (sample of passive or star-forming galaxies) is identified as the one having the lowest number of galaxies. We continue by randomly choosing a galaxy from the larger sample with an overdensity and stellar mass close to that of the galaxy from the reference sample. In practice, we pair galaxies for which the distance in the two-parameter logarithmic space, defined by the stellar mass and the overdensity, is minimal and smaller than 0.1. The procedure is again repeated 10 times in order to construct 10 samples of star-forming and passive galaxies having their mass and density distributions close to each other.

APPENDIX B: THE IMPACT OF COSMIC BOUNDARIES

It was stated in Sections 4.2.1 and 4.3 that the measured gradients towards filaments (Figs 6 and 7) and walls (Figs 9 and 10) are not simply due to gradients towards nodes in the former and due to gradients towards nodes and filaments in the latter case. This Appendix presents the performed tests that allowed us to reach such a conclusion.

B1 Gradients towards filaments

Let us start by considering the gradients towards filaments. In order to probe these gradients without being substantially contaminated by the contribution from nodes, galaxies that are closer to nodes than 3.5 Mpc are removed from the analysis. The choice of this distance $d_{\mathrm{min}}^{\mathrm{node}}$ is motivated by the compromise between eliminating the most of the gradient coming from nodes while keeping enough objects to have a statistically significant sample. Note that the distance of 3.5 Mpc is greater than the typical size of groups, which is ~ 1.5 Mpc in the redshift range considered in this work, measured as a median (or mean) projected group radius. The value of median (and mean) is insensitive to the definition of the group radius (see Robotham et al. 2011, for various definitions considered). In Fig. B1, the solid lines show the mass gradients towards filaments for the entire sample (left-hand panel) on the one hand and after excluding galaxies with distances to the node $D_{\mathrm{node}} \leq 3.5$ Mpc (right-hand panel).

The contribution of nodes to mass gradients towards filaments is measured by randomizing distances to the filament, D_{skel}, in bins of distances to the node, D_{node}. By construction, gradients towards nodes are preserved. 20 samples are constructed in each of which this reshuffling method is applied in 20 equipopulated logarithmic bins. As shown by the dashed lines in Fig. B1 and values of medians listed in Table F1, the reshuffling cancels the gradients towards filaments for $d_{\mathrm{min}}^{\mathrm{node}} = 3.5$ Mpc.

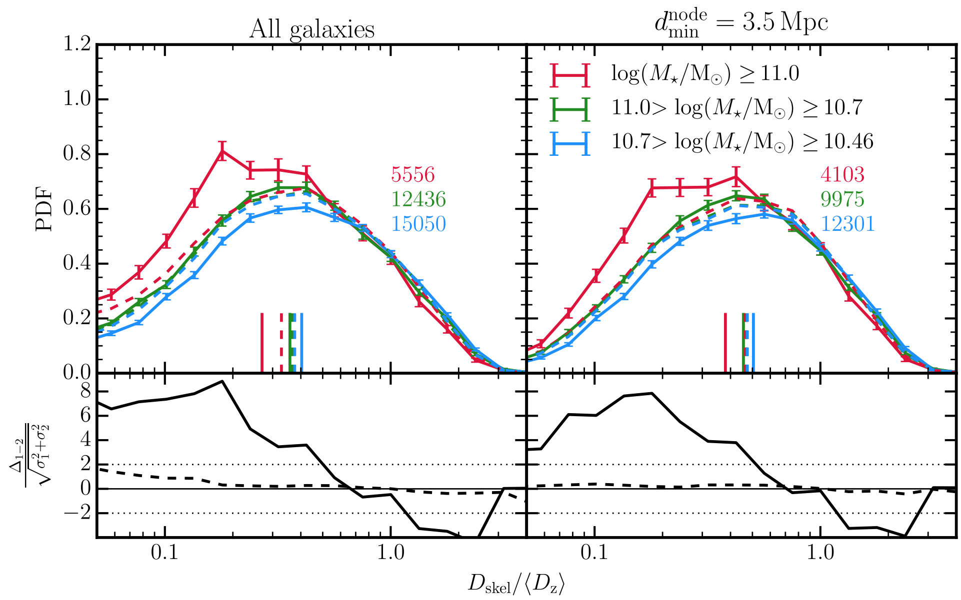


Figure B1. *Top row:* Differential distributions of the normalized distances to the nearest filament, D_{skel}. The solid lines show mass gradients for all galaxies (left-hand panel) and after removing galaxies with distances to the node smaller than 3.5 Mpc (right-hand panel). The dashed lines illustrate mass gradients after the reshuffling of D_{skel} of galaxies in bins of distances to the node D_{node}. The vertical lines indicate the medians of the distributions and their values, together with associated errors, are listed in Table F1. The reshuffling method cancels mass gradients towards filaments once galaxies at distances closer than 3.5 Mpc from nodes are removed. *Bottom row:* Residuals in units of σ between the two most extreme mass bins ($\log(M_\star/\mathrm{M}_\odot) \geq 11.0$ and $10.7 > \log(M_\star/\mathrm{M}_\odot) \geq 10.46$) before (solid lines) and after (dashed lines) the reshuffling.

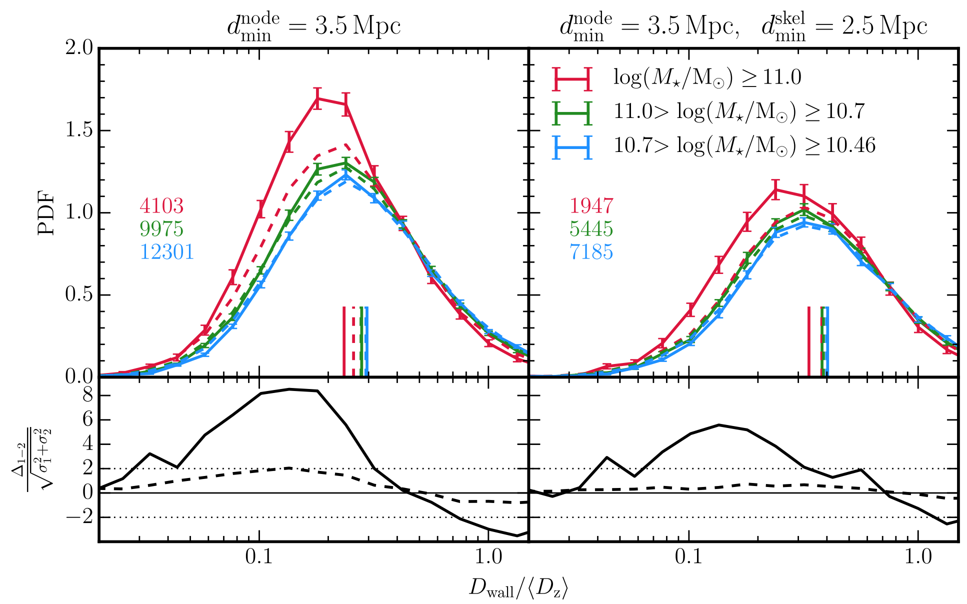


Figure B2. *Top row:* Differential distributions of the normalized distances to the nearest wall, $D_{\rm wall}$. The solid lines show mass gradients after removing galaxies with distances to the node smaller than 3.5 Mpc (left) and after applying an additional criterion on the distance to the filament, such that galaxies with distances to the filament smaller than 2.5 Mpc (right) are removed. The dashed lines illustrate mass gradient after reshuffling of $D_{\rm skel}$ of galaxies in bins of distances to the node $D_{\rm node}$. As shown on the right-hand panel, these are almost completely cancelled after removing sufficiently large regions around nodes and filaments. The vertical lines indicate the medians of the distributions and their values, together with associated errors, are listed in Table F2. *Bottom row:* Residuals are in units of σ as in Fig. B1.

In addition, following Laigle et al. (2017), it can be shown that in the regions sufficiently far away from nodes, gradients towards nodes and those towards filaments are independent. It was checked that the mass gradients towards nodes, present for the entire galaxy sample, are substantially reduced once galaxies for which distances to the node $D_{\rm node} \leq 3.5$ Mpc are excluded. This time, the distances to the node, $D_{\rm node}$, were randomized in bins of distances to the filament, $D_{\rm skel}$, i.e. by construction, gradients towards filaments were preserved. Again, 20 samples were constructed using 20 equipopulated logarithmic bins. After reshuffling, weak gradients at the level of at most 1σ are still present, but note that additional increase in $d_{\rm min}^{\rm node}$ does not reduce them further.

This analysis allows us to conclude that by removing from our sample galaxies that are closer to nodes than 3.5 Mpc, the impact of nodes to the measured gradients towards filaments is minimized, and even if weak gradients towards nodes still exist, these are independent of gradients towards filaments, i.e. gradients towards filaments and gradients towards nodes can be disentangled.

Let us finish this section with two remarks. First, note that distances to the node considered here are 3D euclidian distances. Curvilinear distances along the filaments could have been used instead (as illustrated in Fig. 4). This alternative choice of the distance does not alter our conclusions. Secondly, instead of using distances to the node $D_{\rm node}$, one could have considered distances normalized by the redshift-dependent mean inter-galaxy separation, $D_{\rm node}/\langle D_z \rangle$. These two approaches give consistent results not only qualitatively, but also quantitatively.

B2 Gradients towards walls

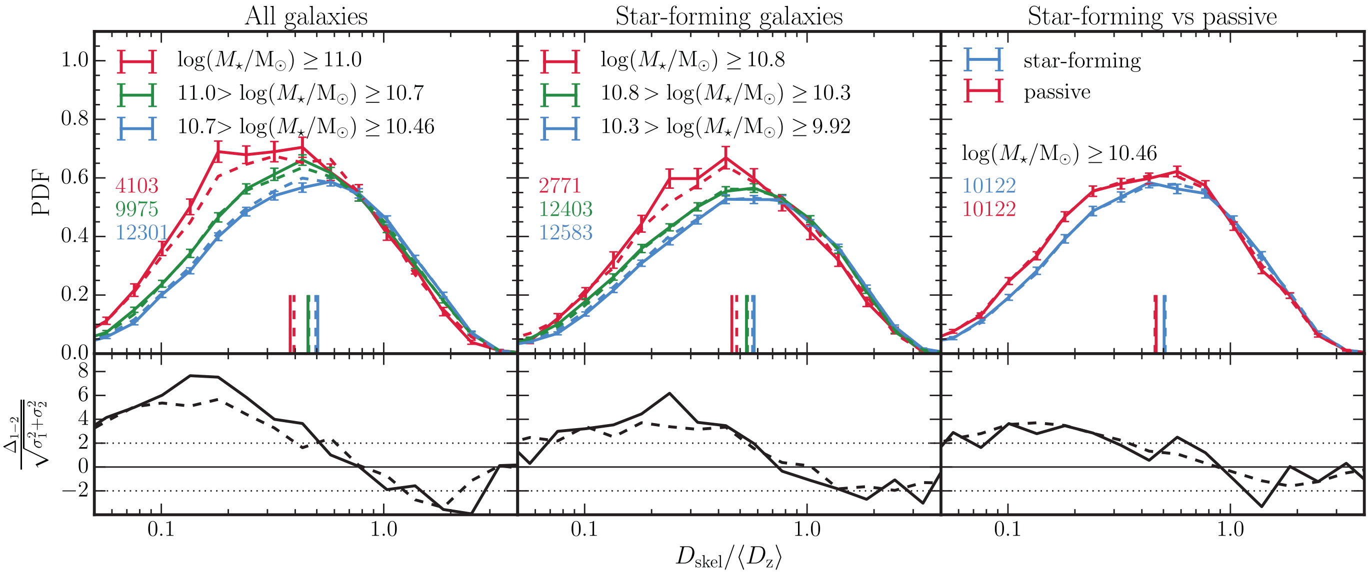
As with filaments, when measuring the gradients towards walls, one should investigate whether the gradient is not dominated by other component of the environments. As filaments are regions where walls intersect, these represent on top of nodes an additional source of contamination for the measured gradients towards walls. Fig. B2 shows the mass gradients towards walls for the galaxy sample outside the zone of influence of nodes parametrized by $d_{\rm min}^{\rm node} = 3.5$ Mpc (left-hand panel) and after applying an additional criterion by excluding galaxies with distances to the closest filament $D_{\rm skel} \leq d_{\rm min}^{\rm skel}$ with $d_{\rm min}^{\rm skel} = 2.5$ Mpc (right-hand panel). The contribution of filaments to the mass gradients towards walls is measured by randomizing distances to the wall, $D_{\rm wall}$, in bins of distances to the filament, $D_{\rm skel}$. By construction, the gradients towards filaments are preserved. Here 20 samples are constructed in each of which the reshuffling method is applied in 20 equipopulated logarithmic bins. As shown by the dashed lines in Fig. B2 and values of medians listed in Table F2, the reshuffling cancels the gradients towards walls for $d_{\rm min}^{\rm skel} = 2.5$ Mpc.

Following the method used in Appendix B1, it was verified (but not shown here) that the mass gradients towards filaments after randomization of the distances $D_{\rm skel}$ in bins of distances to the nearest wall $D_{\rm wall}$ are substantially reduced. Only a very weak mass gradient (at a 1σ level at most) is detected after randomization even for $d_{\rm min}^{\rm skel} = 2.5$ Mpc. Similarly to what was found in Section B1, increasing this parameter does not induce any substantial reduction of the gradient. Thus this distance was chosen as the limit for the exclusion region around filaments.

APPENDIX C: SMALL-SCALE DENSITY-COSMIC WEB RELATION

In this Appendix, the impact of the small-scale density estimator on the mass and type/colour gradients is presented. The density used here is DTFE, i.e. the density computed at the smallest possible scale.[10] As in Section 6, the two methods, the reshuffling and density-matching, are applied.

[10] There is no specific scale associated with the DTFE: it is a local adaptive method which determines the density at each point while preserving its multiscale character.



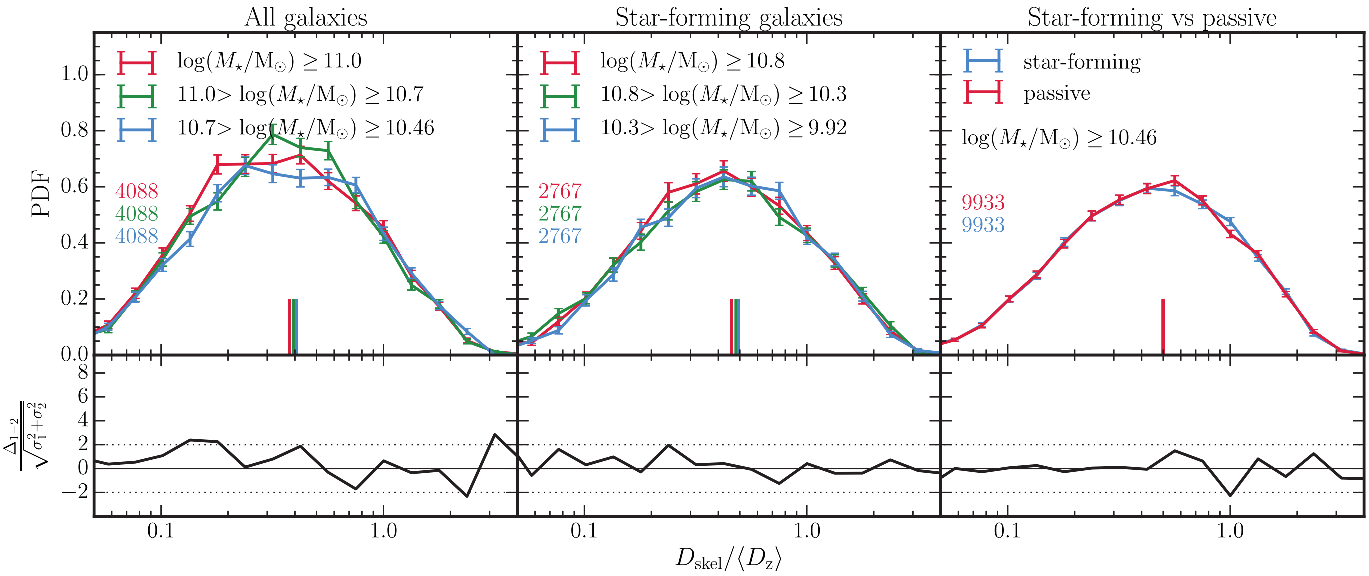


Figure C1. *Top rows:* As in Fig. 12, but using the DTFE density for both methods, reshuffling (Figure a) and density matching (Figure b). The numerical values of medians, shown as vertical lines, are listed in Table F3. When the small-scale density, DTFE in this case, is used in the reshuffling method, the randomized (dashed lines) and original signal (solid lines) are nearly identical. Similarly, all gradients are almost completely erased, as expected. *Bottom rows:* Residuals are in unit of σ as in Fig. 12. (a) Reshuffling. (b) Density matching.

Fig. C1 shows the differential distributions of the distances to the nearest filament, $D_{\rm skel}$ (normalized by $\langle D_z\rangle$, for the same selections as in Fig. 12. The contribution of the nodes to the measured signal is minimized, by removing from the analysis galaxies located closer to a node than 3.5 Mpc. Star-forming and passive galaxies have been matched in mass, as described in Appendix A1. The vertical lines indicate the medians of the distributions, whose values, together with the error bars, are listed in Table F3.

In Figure (a), the mass and type gradients are shown before (solid lines, as in 12) and after (dashed lines) applying the reshuffling of galaxies in the bins of overdensity (1+δ), where the number density corresponds to the DTFE density. The result conforms to the expectations. The reshuffling does not remove the observed mass and type/colour gradients, i.e. the distributions before and after the reshuffling are almost identical, suggesting that at the small scale, traced by DTFE, the density and cosmic web are closely correlated through the small-scale processes.

Figure (b) illustrates the PDFs for samples that have been matched in overdensity (1 + δ), as described in Appendix A2, where the density considered is DTFE. The density-matching technique yields qualitatively similar result than the above used reshuffling in that almost no mass and type gradients are detected when galaxies matched in the DTFE density.

Qualitatively same results are obtained for both methods when applied to the measurements of gradients with respect to the walls (not shown).

APPENDIX D: THE HORIZON-AGN SIMULATION

This Appendix is dedicated to presenting the large-scale cosmological hydrodynamical simulation HORIZON-AGN (Dubois et al. 2014). First, some of the main features of the simulation are briefly summarized. The reshuffling method is then implemented on the simulation, as defined in Section 6, and shown to yield qualitatively similar results to those obtained in GAMA for both large- and small-scale density tracers.

D1 Simulation summary

The detailed description of the HORIZON-AGN simulation[11] can be found in Dubois et al. (2014), here only its brief summary is given. The cosmological parameters used in the simulation correspond to the ΛCDM cosmology with total matter density $\Omega_{\rm m} = 0.272$, dark energy density $\Omega_\Lambda = 0.728$, amplitude of the matter power spectrum $\sigma_8 = 0.81$, baryon density $\Omega_{\rm b} = 0.045$, Hubble constant $H_0 = 70.4\,{\rm km\,s^{-1}\,Mpc^{-1}}$ and $n_s = 0.967$ compatible with the WMAP-7 data (Komatsu et al. 2011).

The simulation was run with the Adaptive Mesh Refinement code RAMSES (Teyssier 2002) in a box of length $L_{\rm box} = 100\,h^{-1}$ Mpc containing 1024^3 dark matter (DM) particles, with a DM mass resolution of $M_{\rm DM,res} = 8 \times 10^7\,{\rm M}_\odot$, and initial gas resolution of $M_{\rm gas,res} = 1 \times 10^7\,{\rm M}_\odot$.

The collisionless DM and stellar components are evolved using a particle-mesh solver. The dynamics of the gaseous component are computed by solving Euler equations on the adaptive grid using a second-order unsplit Godunov scheme.

The refinement is done in a quasi-Lagrangian manner starting from the initial coarse grid down to $\Delta x = 1$ proper kpc (seven levels of refinement) as follows: each AMR cell is refined if the number of DM particles in a cell is more than 8, or if the total baryonic mass in a cell is eight times the initial DM mass resolution. This results in a typical number of 7×10^9 gas resolution elements (leaf cells) in the HORIZON-AGN simulation at $z = 0$.

Heating of the gas from a uniform UV background takes place after redshift $z_{\rm reion} = 10$ following Haardt & Madau (1996). Gas is allowed to cool down to 10^4 K through H and He collisions with a contribution from metals using a Sutherland & Dopita (1993) model.

The conversion of gas into stars occurs in regions with gas density exceeding $\rho_0 = 0.1\,{\rm H\,cm^{-3}}$ following the Schmidt (1959) relation of the form $\dot{\rho}_* = \epsilon_* \rho_{\rm g}/t_{\rm ff}$, where $\dot{\rho}_*$ is the SFR mass density, $\rho_{\rm g}$ the gas mass density, $\epsilon_* = 0.02$ the constant star formation efficiency, and $t_{\rm ff}$ the local free-fall time of the gas.

Feedback from stellar winds, supernovae type Ia and type II are included into the simulation with mass, energy and metal release. HORIZON-AGN simulation takes also into account the formation of black holes (BHs) that can grow by gas accretion at a Bondi–Hoyle–Lyttleton rate capped at the Eddington accretion rate when they form a tight enough binary. The AGN feedback is a combination of two different modes (the so-called quasar and radio mode) in which BHs release energy in the form of heating or jet when the accretion rate is, respectively, above and below 1 per cent of Eddington, with efficiencies tuned to match the BH-galaxy scaling relations at $z = 0$ (see Dubois et al. 2012a, for details).

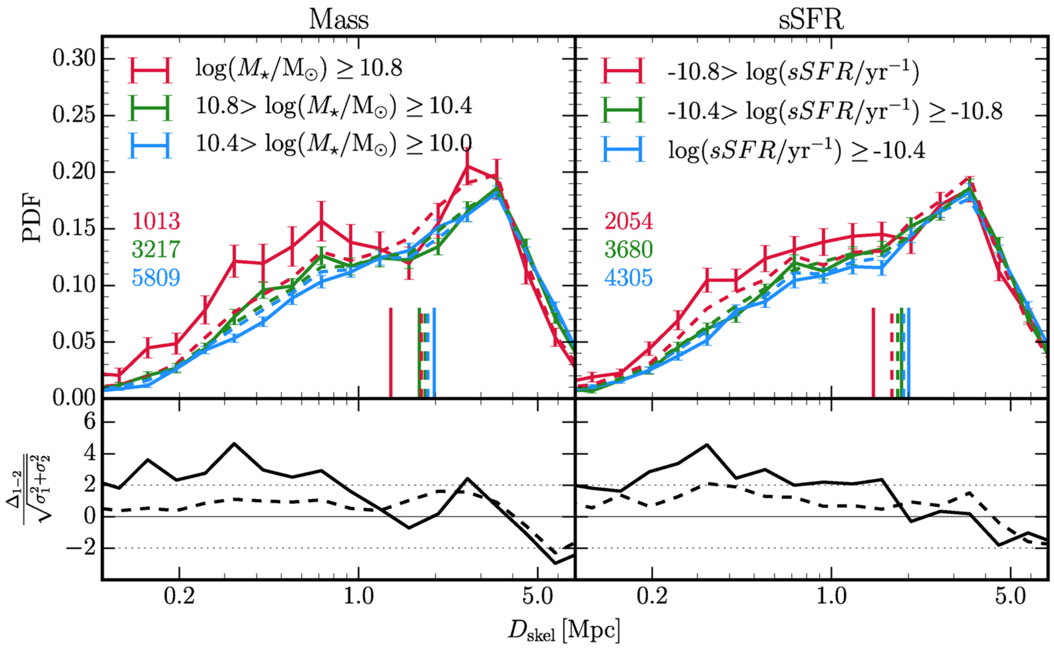
Galaxies are identified using the updated method (Tweed et al. 2009) of the AdaptaHOP halo finder (Aubert et al. 2004) directly operating on the distribution of stellar particles. Only galactic structures with a minimum of $N_{\rm min} = 100$ stellar particles are considered, which typically selects objects with masses larger than $2 \times 10^8\,{\rm M}_\odot$.

D2 Density reshuffling

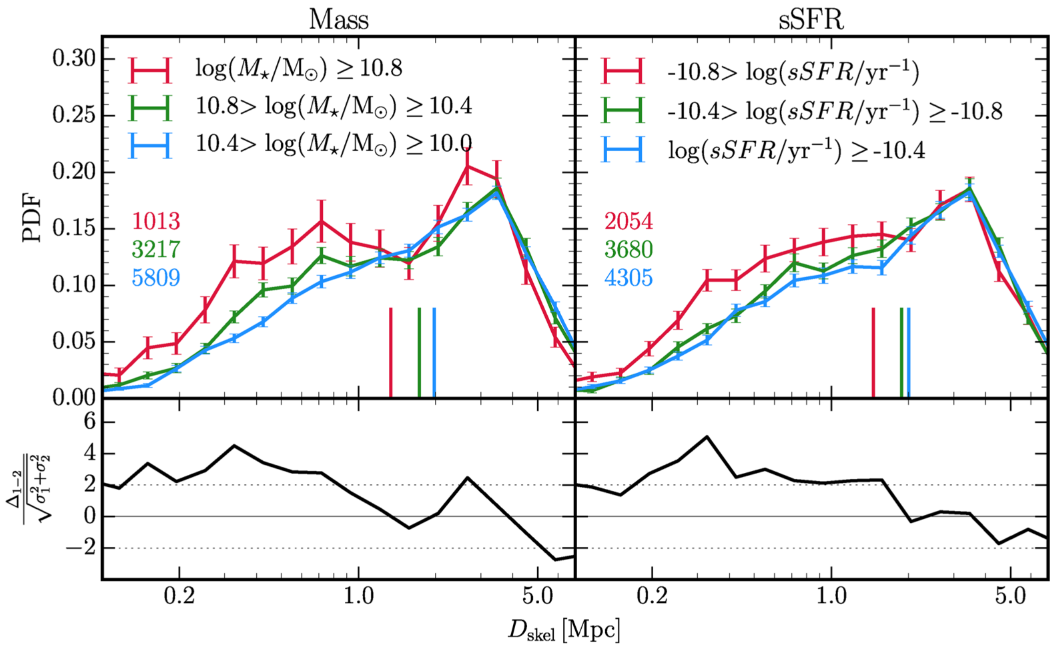
Let us finally present the impact of the reshuffling method, as defined in Section 6, and the choice of the density tracer in the HORIZON-AGN simulation.

Fig. D1 illustrates that the result of reshuffling depends on the scale at which the density is computed. As expected, when using the small-scale density tracer, such as e.g. the DTFE density (Figure a), both mass and sSFR gradients are almost unchanged, while on sufficiently large scales, the gradients tend to cancel out (Figure b). The numerical value of the scale at which this happens is ∼5 Mpc. This is again in a qualitative agreement with the scale required in the GAMA survey, corresponding to the ∼1.5× mean inter-galaxy separation.

[11] http://www.horizon-simulation.org



(a)



(b)

Figure D1. *Top rows:* As in Fig. 11 for the distances to the nearest filament, D_{skel}. The contribution of the nodes is minimized by removing galaxies located within 3.5 Mpc around them from the analysis. The dashed lines correspond to the distributions after the application of the reshuffling method using two different density tracers, a large (Figure a) and small-scale (Figure b) estimators. The numerical values of medians, shown as vertical lines, are listed in Table F4. In qualitative agreement with the results obtained with the observed data, in order to cancel the gradients, density at sufficiently large scale has to be considered. This corresponds to 5 Mpc in the HORIZON-AGN simulation, representing $\sim 1.5\times$ mean inter-galaxy separation, again in agreement with the value found in observations. *Bottom rows:* As in Fig. 11 before (solid lines) and after (dashed lines) the reshuffling. (a) Reshuffling using the density computed in the Gaussian kernel at the scale of 5 Mpc. (b) Reshuffling using the DTFE density.

APPENDIX E: GRADIENT MISALIGNMENTS

In the context of conditional excursion set theory subject to a saddle $\mathcal{S}$ at some finite distance (r, θ, ϕ) from a forming halo, let us consider the Hessian of the potential, $q_{ij} \equiv \partial^2\psi/\partial r_i \partial r_j$, smoothed on the saddle scale R_S and normalized so that $\langle \mathrm{tr}^2(\boldsymbol{q})\rangle = 1$. The anisotropic shear is given by the traceless part $\bar{q}_{ij} \equiv q_{ij} - \delta_{ij}\,\mathrm{tr}\,\boldsymbol{q}/3$, which deforms the region by slowing down or accelerating the collapse along each axis. At finite separation, this traceless shear modifies in an anisotropic way the statistics of the smooth mean density (and of its derivative with respect to scale). The variations are modulated by $\mathcal{Q} = \sum_{i,j} \hat{r}_i \bar{q}_{ij} \hat{r}_j$, with $\hat{r}_i = r_i/r$, i.e. by the relative orientation of the separation vector, $\boldsymbol{r}$ in the frame set by the tidal tensor of the saddle. This extra degree of freedom, $\mathcal{Q}(\theta, \phi)$, provides a supplementary vector space, beyond the radial direction, over which to project the gradients, with statistical weight depending on each specific observable (mass, accretion rate, etc.). These quantities have thus potentially different iso-surfaces from each other and from the local mean density, a genuine signature of the impact of the traceless part of the tidal tensor. Indeed, for each observable, the conditioning on $\mathcal{S}$ introduces a further dependence on the geometry of the environment (the height of the saddle and its anisotropic shear $\bar{q}_{ij}$) and on the position $\mathbf{r}$ of the halo with respect to the saddle point. This dependence arises because the saddle point condition modifies the mean and variance of the stochastic process $(\delta, \partial_R\delta)$ – the height and slope of the excursion set trajectories – in a position-dependent way, making it more or less likely to form haloes of given mass and assembly history within the environment set by $\mathcal{S}$. The expectation of the process becomes anisotropic through $\mathcal{Q}$, and both mean and variance acquire distinct radial dependence through the relevant correlation functions $\xi_{\alpha\beta}$ defined below in equation (E8).

For instance, considering the typical mass, $M_\star$ and accretion rate, $\dot{M}_\star$, at scale R, straightforward trigonometry shows that cross-product of their gradients reads

$$\left(\frac{\partial \dot{M}_\star}{\partial r}\frac{\partial M_\star}{\partial \mathcal{Q}} - \frac{\partial \dot{M}_\star}{\partial \mathcal{Q}}\frac{\partial M_\star}{\partial r}\right)\tilde{\nabla}\mathcal{Q}\,, \tag{E1}$$

where $\tilde{\nabla} = (\partial/\partial\theta, (1/\sin\theta)\partial/\partial\phi)$. The companion paper (Musso et al. 2017) shows that the Taylor expansion in the anisotropy for the angular variation, $\mathcal{Q}$, of $M_\star$ and $\dot{M}_\star$ at fixed distance r from the saddle scale like

$$\Delta M_\star \propto \xi_{20}(r)\mathcal{Q}(\theta,\phi)\,, \tag{E2}$$

and

$$\Delta \dot{M}_\star \propto \left[\xi_{20}'(r) - \frac{\sigma - \boldsymbol{\xi}'\cdot\boldsymbol{\xi}}{\sigma^2 - \boldsymbol{\xi}\cdot\boldsymbol{\xi}}\xi_{20}(r)\right]\mathcal{Q}(\theta,\phi)\,, \tag{E3}$$

in terms of the variance

$$\sigma^2(R) = \int \mathrm{d}k\,\frac{k^2P(k)}{2\pi^2}W^2(kR)\,, \tag{E4}$$

and the radius-dependent vectors

$$\boldsymbol{\xi}(r) \equiv \{\xi_{00}(r), \sqrt{3}\xi_{11}(r)r/R_\star, \sqrt{5}\xi_{20}(r)\}\,, \tag{E5}$$

$$\boldsymbol{\xi}'(r) \equiv \{\xi_{00}'(r), \sqrt{3}\xi_{11}'(r)r/R_\star, \sqrt{5}\xi_{20}'(r)\}\,, \tag{E6}$$

where

$$R_\star^2 \equiv \int \mathrm{d}k\,\frac{P(k)}{2\pi^2}\frac{W^2(kR_\mathcal{S})}{\sigma_\mathcal{S}^2}\,, \tag{E7}$$

with $P(k)$ the underlying power spectrum, $W(k)$ the top hat filter in Fourier space, $\sigma_\mathcal{S} = \sigma(R_\mathcal{S})$, while the finite separation correlation functions, $\xi_{\alpha\beta}(r, R, R_\mathcal{S})$ and $\xi_{\alpha\beta}'(r, R, R_\mathcal{S})$ are defined as

$$\xi_{\alpha\beta} \equiv \int \mathrm{d}k\,\frac{k^2P(k)}{2\pi^2}W(kR)\frac{W(kR_\mathcal{S})}{\sigma_\mathcal{S}}\frac{j_\alpha(kr)}{(kr)^\beta}\,, \tag{E8}$$

$$\xi_{\alpha\beta}' \equiv \int \mathrm{d}k\,\frac{k^2P(k)}{2\pi^2}W'(kR)\frac{W(kR_\mathcal{S})}{\sigma_\mathcal{S}}\frac{j_\alpha(kr)}{(kr)^\beta}\,, \tag{E9}$$

where $j_\alpha(x)$ are the spherical Bessel functions of the first kind and prime denote derivate with respect to σ. Note that equation (E3) clearly highlights the shifted variance, $\sigma^2 - \boldsymbol{\xi}\cdot\boldsymbol{\xi}$, which contributes to the difference between $\Delta M_\star$ and $\Delta\dot{M}_\star$. From equation (E3), since the square bracket is not proportional to ξ_{20} as in equation (E2), it follows that the cross-product in equation (E1) is non-zero, which in turn implies that the contours of mass and accretion rate differ.

APPENDIX F: MEDIANS OF DISTRIBUTIONS

This Appendix gathers tables of medians with corresponding error bars used in previous sections.

Table F1. Medians of $D_{\mathrm{skel}}/\langle D_z\rangle$ for Fig. B1.

Selection[a]	Mass bin	Median[b] $D_{\mathrm{skel}}/\langle D_z\rangle$ Before reshuffling[c]	After reshuffling
All galaxies	$\log(M_\star/\mathrm{M}_\odot) \geq 11$	0.27 ± 0.01	0.33 ± 0.02
	$11 > \log(M_\star/\mathrm{M}_\odot) \geq 10.7$	0.36 ± 0.01	0.37 ± 0.01
	$10.7 > \log(M_\star/\mathrm{M}_\odot) \geq 10.46$	0.40 ± 0.01	0.38 ± 0.01
$d_{\mathrm{min}}^{\mathrm{node}} = 3.5$ Mpc	$\log(M_\star/\mathrm{M}_\odot) \geq 11$	0.38 ± 0.01	0.46 ± 0.02
	$11 > \log(M_\star/\mathrm{M}_\odot) \geq 10.7$	0.46 ± 0.01	0.47 ± 0.01
	$10.7 > \log(M_\star/\mathrm{M}_\odot) \geq 10.46$	0.51 ± 0.01	0.47 ± 0.01

[a]Panels of Fig. B1.
[b]Medians of distributions as indicated in Fig. B1 by vertical lines; errors are computed as in Table 1.
[c]Randomization of D_{skel} in bins of D_{node}.

Table F2. Medians of $D_{\mathrm{wall}}/\langle D_z\rangle$ for Fig. B2.

Selection[a]	Mass bin	Median[b] $D_{\mathrm{wall}}/\langle D_z\rangle$ Before reshuffling[c]	After reshuffling
$d_{\mathrm{min}}^{\mathrm{node}} = 3.5$ Mpc	$\log(M_\star/\mathrm{M}_\odot) \geq 11$	0.234 ± 0.005	0.258 ± 0.011
	$11 > \log(M_\star/\mathrm{M}_\odot) \geq 10.7$	0.279 ± 0.003	0.278 ± 0.005
	$10.7 > \log(M_\star/\mathrm{M}_\odot) \geq 10.46$	0.295 ± 0.003	0.292 ± 0.004
$d_{\mathrm{min}}^{\mathrm{node}} = 3.5$ Mpc, $d_{\mathrm{min}}^{\mathrm{skel}} = 2.5$ Mpc	$\log(M_\star/\mathrm{M}_\odot) \geq 11$	0.334 ± 0.007	0.379 ± 0.028
	$11 > \log(M_\star/\mathrm{M}_\odot) \geq 10.7$	0.381 ± 0.004	0.386 ± 0.011
	$10.7 > \log(M_\star/\mathrm{M}_\odot) \geq 10.46$	0.403 ± 0.004	0.398 ± 0.008

[a]Panels of Fig. B2.
[b]Medians of distributions as indicated in Fig. B2 by vertical lines; errors are computed as in Table 1.
[c]Randomization of D_{wall} in bins of D_{skel}.

Table F3. Medians for the PDFs displayed in Fig. C1: small-scale density

	Selection[a]	Bin	Median[b] Original[c]	$D_{\rm skel}/\langle D_z\rangle$ reshuffling[d]	Matching[e]
Masses	All galaxies	$\log(M_\star/\mathrm{M}_\odot) \geq 11$	0.379 ± 0.009	0.397 ± 0.009	0.378 ± 0.01
		$11 > \log(M_\star/\mathrm{M}_\odot) \geq 10.7$	0.456 ± 0.007	0.459 ± 0.006	0.393 ± 0.009
		$10.7 > \log(M_\star/\mathrm{M}_\odot) \geq 10.46$	0.505 ± 0.006	0.495 ± 0.006	0.406 ± 0.008
	SF galaxies	$\log(M_\star/\mathrm{M}_\odot) \geq 10.8$	0.459 ± 0.012	0.489 ± 0.013	0.458 ± 0.011
		$10.8 > \log(M_\star/\mathrm{M}_\odot) \geq 10.3$	0.534 ± 0.007	0.541 ± 0.008	0.479 ± 0.01
		$10.3 > \log(M_\star/\mathrm{M}_\odot) \geq 9.92$	0.578 ± 0.007	0.567 ± 0.007	0.494 ± 0.006
Types	SF versus passive[f]	Star-forming	0.504 ± 0.008	0.508 ± 0.007	0.495 ± 0.006
		Passive	0.462 ± 0.007	0.458 ± 0.007	0.504 ± 0.006

[a]Panels of Fig. C1.
[b]Medians of distributions as indicated in Fig. C1 by vertical lines; errors are computed as in Table 1.
[c]As in Table 1 for $D_{\rm skel}/\langle D_z\rangle$.
[d]Reshuffling is done in bins of DTFE density (see the main text for more details).
[e]Medians for the density-matched sample, where the density considered is DTFE.
[f]Only galaxies with stellar masses $\log(M_\star/\mathrm{M}_\odot) \geq 10.46$ are considered.

Table F4. Medians for the PDFs displayed in Fig. D1

Selection[a]	Bin	Median[b] $D_{\rm skel}$ [Mpc] Original[c]	after reshuffling[d] DTFE	G5Mpc
Mass	$\log(M_\star/\mathrm{M}_\odot) \geq 10.8$	1.34 ± 0.09	1.26 ± 0.08	1.72 ± 0.1
	$10.8 > \log(M_\star/\mathrm{M}_\odot) \geq 10.4$	1.73 ± 0.08	1.71 ± 0.06	1.82 ± 0.06
	$10.4 > \log(M_\star/\mathrm{M}_\odot) \geq 10$	1.97 ± 0.04	2.0 ± 0.05	1.86 ± 0.04
sSFR	$-10.8 > \log(sSFR/\mathrm{yr})$	1.46 ± 0.07	1.61 ± 0.07	1.74 ± 0.08
	$-10.4 > \log(sSFR/\mathrm{yr}) \geq -10.8$	1.88 ± 0.06	1.89 ± 0.06	1.81 ± 0.06
	$\log(sSFR/\mathrm{yr}) \geq -10.4$	2.0 ± 0.04	1.9 ± 0.05	1.91 ± 0.06

[a]Panels of Fig. D1.
[b]Medians of distributions as indicated in Fig. D1 by vertical lines; errors are computed as in Table 1.
[c]As in Table 2 for $D_{\rm skel}$ (corresponding to the solid lines in Fig. D1).
[d]Reshuffling is done in the bins of the DTFE density and the density computed at the scale of 5 Mpc (corresponding to the dashed lines in Figures a and b, respectively).

This paper has been typeset from a TEX/LATEX file prepared by the author.