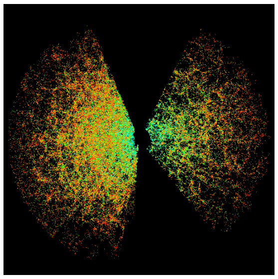
Galaxy Clustering in the Completed SDSS Redshift Survey: The Dependence on Color and Luminosity

Idit Zehavi (CWRU), Zheng Zheng (Yale), David Weinberg (OSU), Michael Blanton (NYU), et al.

(*Astrophysical Journal*, 2011, in press; arXiv:1005.2413)



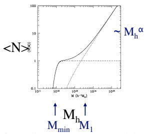
Distribution of galaxies in the SDSS redshift survey (spanning here $0.02<z<0.22$), illustrating the large-scale structure we are quantifying. Galaxies are color coded here by their *g-r* color. We analyze volume-limited subsamples of well defined luminosity ranges. We measure the correlation function using the Landy-Szalay estimator and we obtain the full error covariance matrix using jackknife resampling.

We study the luminosity and color dependence of the galaxy two-point correlation function in the largest-existing redshift survey, the main galaxy sample of the Sloan Digital Sky Survey (SDSS) Seventh Data Release. We measure the projected correlation function $w_p(r_p)$ of volume-limited samples extracted from the parent sample of ~700,000 galaxies over 8000 deg^2 extending up to a redshift of 0.25. The amplitude of $w_p(r_p)$ grows slowly with luminosity for $L<L_*$ with a sharper increase at higher luminosities. At fixed luminosity, redder galaxies exhibit a higher amplitude and steeper correlation function, a steady trend that runs through the “blue cloud” and “green valley” and continues across the “red sequence”. The cross-correlation of red and blue galaxies is close to the geometric mean of their auto-correlations, dropping slightly below it on small scales. The correlation amplitude of blue galaxies increases steadily with luminosity, but the luminosity dependence of red galaxies is more complex, with the faint red galaxies exhibiting the strongest clustering on small scales.

We interpret these results using halo occupation distribution (HOD) models, assuming a concordance cosmology. Most of the observed trends can be naturally understood in this framework. The growth of $w_p(r_p)$ with luminosity reflects an overall shift in the mass scale of their host dark matter halos, in particular an increase in the minimum host halo mass $M_{\rm min}$. The mass at which a halo has, on average, one satellite galaxy brighter than L is $M_1 \approx 17\ M_{\rm min}(L)$ over most of the luminosity range. The growth and steepening of $w_p(r_p)$ for redder galaxies reflects the increasing fraction of galaxies that are satellite systems in high-mass halos instead of central galaxies in low-mass halos, a trend that is especially marked at low luminosities. This provides insight on the relation between galaxies and dark matter halos and informative tests for galaxy formation theories.

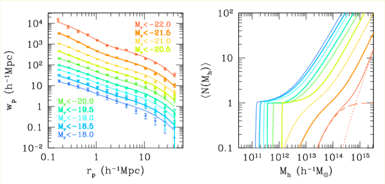
HOD Modeling

The HOD formalism is a useful framework to describe galaxy biasing by specifying how galaxies populate the dark-matter halos. The key ingredient is the halo occupation function, the mean number of galaxies as a function of halo mass. For a sample of galaxies brighter than a given luminosity, the mean occupation function can be well described by a smoothed step function for the *central* galaxy and a power-law number of *satellites* increasing with halo mass.



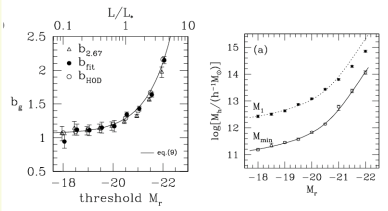
The HOD formalism naturally explains the distinct shape of correlation function as the transition from inter-halo to intra-halo galaxy pairs. For an assumed cosmology, one can find the best-fit HOD model to the observations. In a sense, the HOD formalism casts galaxy clustering results into a more physically informative form that is easier to relate to galaxy formation theories.

Dependence on Luminosity



Projected galaxy correlation functions, $w_p(r_p)$, for luminosity threshold samples (left panel; samples are staggered by 0.25 dex for clarity). The curves are the best-fit HOD models to these measurements. The measured w_p’s are well fit by these models (better than power-law fits). A clear dependence on luminosity is observed, where the more luminous galaxies are more clustered. The individual correlation functions have similar slopes with increasing amplitudes, indicating a roughly “scale-invariant” bias.

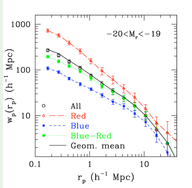
The right panel shows the corresponding halo occupation functions $\langle N(M_h)\rangle$, color-coded in the same way. The occupation functions shift to the right, toward more massive halos, as the luminosity threshold increases. The separation of central and satellite galaxies is shown for the brightest sample, as the dashed and dotted curves. The three brightest samples require smooth central-galaxy cutoff profiles to match the clustering data, while for the fainter samples we have chosen models with sharp cutoffs that are very close to the best-fits.



These $w_p(r_p)$ measurements can be translated to bias factors as a function of luminosity (left panel), calculated in different ways: using a fixed separation of r_p=2.67 h^{-1} Mpc; from their ratio to the dark matter w_p over a range of separations; and from the HOD model fits. The amplitude grows slowly with luminosity below L_* and increases sharply above it, with a large-scale bias factor of $b(>L)x(\sigma_8/0.8)=1.06+0.21(L/L_*)^{1.12}$.

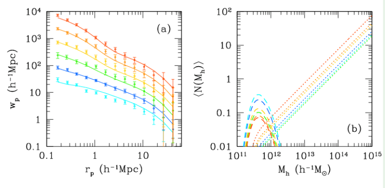
The right panel shows the characteristic halo mass scales for hosting a central galaxy, $M_{\rm min}$, and satellites, M_1. Both halo mass scales increase with luminosity, with a steeper dependence for bright galaxies (as these tend to reside in clusters, so require a more massive halo to host them). The solid curve is a simple fit to the $M_{\rm min}$ values, and the dotted curve is a scaled-up version of it showing that $M_1 \approx 17\ M_{\rm min}$ over most of the range. This implies that a halo hosting two galaxies above a luminosity threshold has to be, on average, at least 17 times more massive as a halo hosting one such galaxy. Halos in the “hosting gap” tend to host more luminous central galaxies rather than multiple galaxies.

Dependence on Color



Projected correlation function of the full volume-limited sample corresponding to $-20<M_r<-19$ and of the “red sequence” and “blue cloud” galaxies in this sample (based on a *g-r* tilted color cut parallel to the red sequence). The red galaxies exhibit a steeper and stronger clustering, while the blue galaxies have a shallower and lower correlation function. These trends are similar for all luminosity samples, but the differences in clustering are stronger with decreasing luminosity.

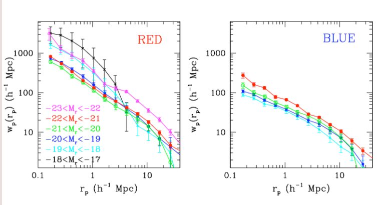
The green symbols show the projected *cross-correlation* function of the red and blue galaxies. On large scales, as expected, we find that it agrees with the geometrical mean of the red and blue auto-correlations (solid black line). On small scales, it falls below the geometric mean, possibly indicating a slight segregation in the mixing of galaxy populations within halos.



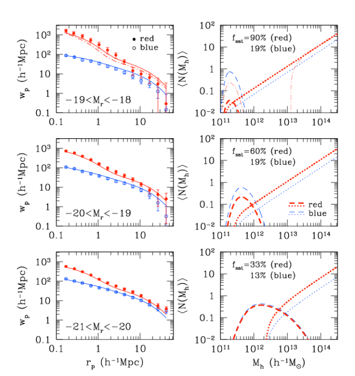
We farther divide the galaxies into bluest, bluer, redder, and reddest *g-r* color subsamples, using cuts parallel to the red sequence in the color-magnitude diagram. Additionally, we study an intermediate “green valley” galaxy population and a “red sequence” subsample along its cusp. The left panel shows the projected correlation functions for these samples (staggered by 0.25 dex) and their best-fit HOD model predictions. We find a continuous color trend with both the amplitude and slope increasing for redder galaxies.

The corresponding central and satellite halo occupation functions are shown on the right, using a simplified HOD model. The steady color trend is explained by a change in the relative contribution of central and satellite galaxies. The redder the color, the more satellites there are relatively, resulting in a larger typical halo mass and stronger clustering (the satellite fraction increases from ~15% to ~75% for these samples). This is a very different trend than the luminosity dependence , which is explained by an overall shift toward higher halo masses with luminosity and a decrease in the satellite fraction. In contrast, the trend with color at fixed luminosity is consistent with a fixed halo mass for central galaxies and a steady increase of satellite fraction with redder color.

Joint Dependence on Luminosity and Color



We now examine the luminosity dependence of clustering within the red (left) and blue (right) populations separately. The overall differences in clustering strength and slope between red and blue are apparent. The projected correlation functions of the blue galaxies are all roughly parallel with a slow but steady increase of amplitude with luminosity. The luminosity dependence of red galaxies is more complex. In particular, the faint red galaxies (cyan and black curves) exhibit very strong clustering on small scales, indicating that these galaxies likely reside in high-density environments.



HOD fitting to the projected correlation functions of red and blue galaxies in three luminosity bins. Left-hand panels show the $w_p(r_p)$ measurements (the same as in the figure above) and best-fitting models. The occupation functions for central and satellite galaxies are shown on the right. The simplified model explains the rather complex color-luminosity trends fairly well. Most blue galaxies are central and the satellite fraction for red galaxies is higher and increases toward low luminosities. The strong small-scale clustering of faint red galaxies is reproduced and implies that most of these galaxies are satellites in massive halos. (The dot-dashed lines in the upper panels indicate an alternative fit for the faint red population, where the satellite cutoff is modified.)