Making Mock Galaxy Catalogs with ADDGALS

Michael T. Busha & Risa H. Wechsler
KIPAC, Stanford University, 452 Lomita Mall. Stanford, CA 94305, USA

We present the ADDGALS algorithm for making mock photometric galaxy catalogs. This is a statistical method that attempts to connect galaxy magnitudes to the local dark matter density in an N-body simulation such that the luminosity function and luminosity dependent 2-pt function are correctly modeled. Colors from a training set are then added such that the color-environment relation is reproduced. Comparisons are made to observational data, showing that we reproduce both the large scale and inter-cluster distribution of colors and magnitudes. We also discuss ongoing applications of the mocks, including attempts to calibrate optical cluster finders and photometric redshift estimators.

1 Introduction

Because images can be taken rapidly, large-scale photometric surveys have become a driving force in cosmology by providing deep, detailed maps of a large portion of the universe. Such surveys allow us probe cosmology through, among other things, studies of galaxy clusters, baryon acoustic oscillations (BAO), and weak lensing. However, achieving competitive percent level constraints on cosmological parameters requires the ability to characterize and constrain systematic uncertainties, many of which depend on both the behavior of the analysis tools and the underlying dark matter observable relation. One way to address this is to connect the observed galaxy distribution with the predicted dark matter distribution, whose cosmological dependences can be accurately simulated. Mock galaxy catalogs play an essential role in accomplishing this by letting us run analysis tools on a data set whose underlying cosmological model is known.

Currently there are a number of methods for making mock galaxy catalogs from N-body simulations, the most common of which include modeling the halo occupation (HOD)[1], semi-analytic modeling (SAM)[4], and matching the subhalo distribution with the observed galaxy population by abundance[7]. However, the latter two methods require high-resolution simulations that resolve either subhalos or the halo mass accretion history, and the former requires that you resolve all halos expected to host a galaxy down to some magnitude limit. For upcoming surveys such as DES, which will map 5,000 sq deg to $z \sim 1.5$, this requires N-body simulations with $> 5\times 10^{10}$ particles for a HOD and substantially more for the other methods. By comparison, the largest published simulation contains only 10^{10} particles. We present an alternative algorithm, ADDGALS (Adding Density Determined GAlaxies to Lightcone Simulations), that can produce a comparable mock using more than an order of magnitude or more fewer particles.

2 The ADDGALS Algorithm

ADDGALS is an algorithm for "painting" galaxies onto dark matter particles in an N-body simulation by matching galaxy luminosities with local dark matter densities, not dark matter halos.

It is designed to work on either a single snapshot or lightcone output. The later is particularly useful because it allows for direct comparison with galaxy surveys. In addition to an N-body simulation, the algorithm takes as inputs a galaxy luminosity function, a luminosity-dependent correlation function, and a distribution of galaxy colors given luminosity and environment.

2.1 Galaxy Luminosities

The assignment of galaxy luminosities assumes that the primary physical processes setting a galaxies luminosity operate at a scale around M_*. We therefore smooth our simulation at the lagrangian scale $1.8 \times 10^{13}\mathrm{h}^{-1}\mathrm{M}_\odot$, which allows the algorithm to work on simulations with particle mass resolution as poor as $\sim 10^{12}\mathrm{h}^{-1}\mathrm{M}_\odot$.

To paint galaxies onto particles, we first generate a list of galaxy luminosities by extracting them from a luminosity function,

$$\phi(M)dM = 0.4\ln(10)\phi_* 10^{-0.4(M-M_*)(\alpha+1)} \exp[-10^{-0.4(M-M_*)}]dM. \quad (1)$$

We use $\phi_* = 0.0149$, $M_* - 5\log h = -20.44$, and $\alpha = -1.05$, fit to the observed SDSS r-band luminosity function at $z = 0.1$[2]. These galaxies are then assigned to dark matter particles in the simulation by connecting galaxy luminosity to the local dark matter density using R_δ, the radius around a particle containing $1.8 \times 10^{13}\mathrm{h}^{-1}\mathrm{M}_\odot$ of dark matter. This relation is motivated by measurements of SAMs[4] and subhalo distributions[7] in high-resolution simulations. Figure 1 shows the distribution if R_δ as a function of galaxy magnitude for a SAM. We characterize this as a log-normal plus a gaussian for the "central" and "field" peaks. This gives the probability for a galaxy to have R_δ given L as

$$P(R_\delta|L/L_*) = (1 - p(L/L_*))e^{-(\ln(R_\delta)-\mu_c(L/L_*))^2/2\sigma_c(L/L_*)^2}/R\sqrt{2\pi}\sigma_c(L/L_*)$$
$$+p(L/L_*)e^{(R_\delta-\mu_f(L/L_*))^2/2\sigma_f(L/L_*)^2}/\sqrt{2\pi}\sigma_f(L/L_*). \quad (2)$$

Here, μ_c, σ_c, μ_f, and σ_f define the log-normal and the gaussian while $p(L/L_*)$ determines the relative height of the peaks. Galaxies are attached to simulation particles using $P(R_\delta|L/L_*)$ to select a dark matter overdensity and then randomly choosing a particle with that density. When putting galaxies into lightcone simulations, magnitudes are passively evolved by the relation $M(z) = M(z = 0.1) - Q(z - 0.1)$ with $Q = 1.3$[3].

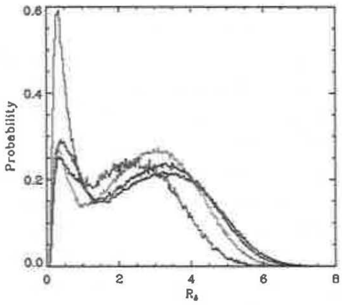


Figure 1: Probability distribution of local dark matter overdensity, R_δ, around semi-analytic galaxies in the Millennium Simulation[8,4]. Colors represent galaxies with different magnitudes: $M_r = -22$ (red), $M_r = -21$ (green), $M_r = -20$ (blue), $M_r = -19$ (black).

The parameters in equation 2 are determined using information about the correlation function. Taking the observed luminosity-dependent 2-point function from SDSS[10], we run an MCMC over a parameterization for the evolution of μ_c, σ_c, μ_f, σ_f, and p with magnitude until we match the observations. Thus, the algorithm is designed to reproduce both the luminosity function and magnitude dependent 2-point function in the r band.

2.2 BCGs

While the above method works for most galaxies, comparisons with SDSS shows that the BCG population is incorrect. First, because galaxies are inserted without direct knowledge the halos distribution, there is no guarantee that a bright galaxy will be placed at the center of each halo. Additionally, even the brightest galaxy in a cluster-sized halo is typically dimmer than observed BCGs for halos of similar mass. To correct this, we add BCGs to halo centers separately, using an observed luminosity-mass relation,

$$\langle L_{BCG} \rangle = L_0 \frac{(M_{200}/M_c)^{\alpha}}{[1 + (M_{200}/M_c)^{bk}]^{1/k}}. \quad (3)$$

Here $L_0 = 2.8 \times 10^9$, $M_c = 3.7 \times 10^9$, $\alpha = 29.78$, $b = 29.5$, and $k = 0.0255$[9,11]. We also add a 15% scatter in log-luminosity at fixed mass to match observations[5].

2.3 Galaxy Colors

While sections 2.1 and 2.2 created a distribution galaxies with r-band magnitudes, in order to model photometric surveys we must also include colors. To do this, we take a training set of SEDs from the SDSS catalog and map them to our simulated galaxies so that we match the color-environment relation. We measure the distance to the 5th nearest galaxy for our mock catalog and randomly select a SDSS galaxy with similar galaxy overdensity and M_r and assign its SED to our galaxy. The SED is then k-corrected to the appropriate redshift and filters are applied to return colors.

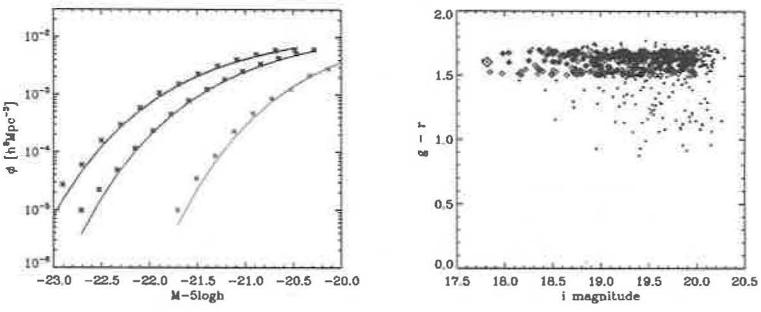
3 Comparisons with Data

Figure 2(a) compares our g, r, and i band luminosity functions with the best fit Schechter functions from the SDSS sample at z = 0.1[2]. There is excellent agreement in the r band (which was an input constraint), although there is a slight excess at the bright end, $M_r < -22.2$, due to the addition of BCGs. We see similarly good agreement in the g and i bands. It is important to emphasize that information from these bands was not directly input into the catalogs. This agreement demonstrates that our method for mapping SEDs to galaxies reproduces the observed large scale distribution of colors.

We have also studied local color trends by looking at cluster members. Figure 2(b) shows the distribution of galaxies in color-magnitude space around a rich cluster identified by the maxBCG[6] cluster finder. This cluster finder works by looking for a suitable BCG candidate at the end of a well-defined red sequence. Such a red sequence is clearly present in Figure 2(b), indicating that we are able to simultaneously reproduce local color structure.

4 Applications of the Catalogs

As noted in the introduction, the primary purpose of ADDGALS is to create a realistic catalog with a galaxy distribution to which photometric data from large surveys can be compared. In particular, we are working to prepare for the upcoming Dark Energy Survey (DES), a 5,000 sq deg survey in the *ugrizY* bands that will constrain dark energy using studies of supernovae, BAO, weak lensing, and clusters. Many systematics need to be addressed and understood before



(a) (b)

Figure 2: (a) The g (green), r (red), and i (blue) band luminosity functions for our mock catalog (points) and best fit Schechter functions to the SDSS data[2] (lines). (b) The color-magnitude distribution for ADDGALS generated galaxies in a cluster identified by maxBCG[6]. Black points show the distribution of all galaxies around the halo, while red diamonds show the identified red cluster members. The blue diamond marks the identified BCG.

competitive cosmological constraints can be extracted. These mocks are being used to, among other things, help calibrate optical cluster finders and photometric redshift estimators.

In order for clusters to provide competitive cosmological constraints, issues such as the purity, completeness, and the mass-observable relation of cluster finders need to be understood. We are in the process of directly comparing ~ 7 different cluster finders, including red sequence, voronoi tessellation, and matched filter methods. By running all of the cluster finders on the same mock, we are able to directly compare the results of these methods with each other and with the underlying dark matter distribution. Additionally, the results are being used to optimize the cluster finders.

The catalogs are also serving as a testing platform for a number of photometric redshift estimators. We have been working to understand the accuracy of template, neural network, and boosted decision tree methods using these catalogs. This is a difficult task because DES will be pushing to high-z environments where little spectroscopic information exists. In particular, we need to be able to characterize the amount of scatter and rate of catastrophic failures for each photo-z estimator. Failure to do so will result in both an increase in systematic errors and possible bias when extracting cosmological parameters from photometric surveys.

References

1. Berlind, A., Weinberg, D., H., ApJ **575**, 587 (2002)
2. Blanton, M., R., et al. , ApJ **592**, 819 (2003).
3. Brown, M., J., I., et al. ApJ **654**, 858 (2007).
4. De Lucia, G., & Blaizot, J., MNRAS **375**, 2 (2007)
5. Hansen, S., M., Sheldon, E., S., Wechsler, R., H., & Koester, B., P., arXiv/0710.3780, *submitted to ApJ.*
6. Koester, B., P., et al. , ApJ **660**, 221 (2007).
7. Kravtsov, A., V., et al. , ApJ **609**, 35 (2004).
8. Springel, V., et al. , Nature **435**, 629 (2005).
9. Vale, A., & Ostriker, J., P., MNRAS **371**, 1173 (2006).
10. Zehavi, I., et al. , ApJ **630**, 1 (2005).
11. Zheng, Z., Coil, A., & Zehavi, I., ApJ **667**, 760 (2007).