WHAT BEST CONSTRAINS GALAXY EVOLUTION IN THE LOCAL UNIVERSE?

DAVID W. HOGG
Center for Cosmology and Particle Physics, Department of Physics,
New York University, 4 Washington Pl., New York, NY 10003, USA.
`david.hogg@nyu.edu`

Abstract. After a polemical introduction about the proper activity of an astrophysicist facing a dominant theoretical model and many Tb of highly informative data, I review a few recent results on the properties of galaxies in the nearby (redshift one-tenth) Universe that directly bear on physical cosmology. In one example, I show that there are a number of ways of measuring, or strongly constraining, massive galaxy–galaxy major merger rates, which are predicted with limited uncertainties in the current generation of models. In another, I show that we can go beyond "correlations" between individual galaxy properties and "environment". Our results show that it is galaxy star-formation histories—not their morphologies—that are sensitive to environmental density. I look forward to a future, perhaps not that far away, in which these results guide a fundamental modification to our theoretical assumptions, though I fear that the dominant paradigm may not require subversion.

1 Introduction

1.1 There's a lot of information out there

Figure 1 shows a (sorry, black-and-white) image of M51 taken from the Sloan Digital Sky Survey [50] imaging data. It positively *teems* with *information*. You can see not just a bulge and spiral arms, but (in the color version) that star forming regions and dusty regions are systematically displaced within the spiral arms, that the spiral galaxy is interacting with a smaller bulge-dominated galaxy, that there are tidal tails and loops at very low surface brightnesses, that the spiral structure is related to the interaction, and even that the bulge-dominated galaxy contains an active nucleus. One of the ironies of my life is that these fantastically beautiful, detailed, and informative pictures are almost *completely useless* for constraining physical models for the formation and evolution of galaxies in a cosmological context. We simply do not have theoretical models—simulations—that produce anything that looks even remotely like Figure 1. As you read the contributions to this volume, many of which speak to the great success of our cosmological model and the tremendous advances made in the last ten years in understanding galaxy evolution, I would like you to keep in mind this important point: *These great theories do not make M51.* Of course, that is not to say they won't, someday.

If that is one of the ironies of my life, then one of the *crusades* of my life is to promote an understanding of the tremendous amount of astronomical information available to us, and the best ways to harness that information in the service of physics. By "information" here I don't just mean the touchy–feely thing you get from reading this volume, but I mean the scalar quantity, measured in bits, first worked out for the world by Shannon [41] (in what must be

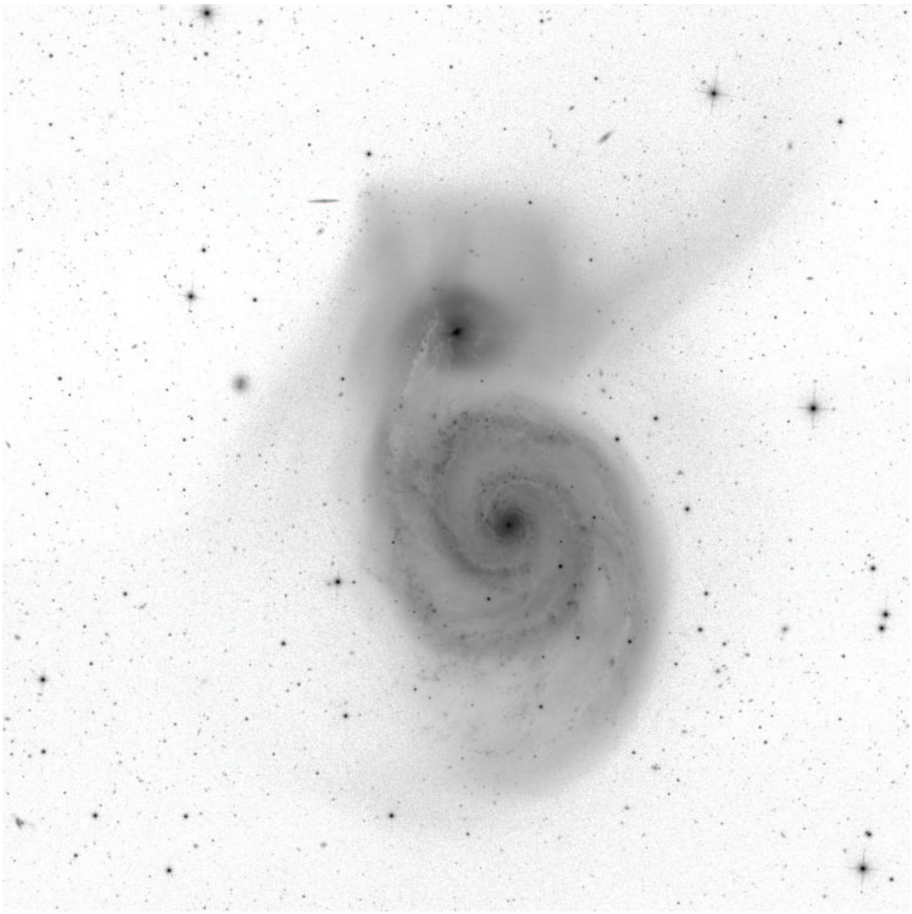

Figure 1: A black-and-white representation of SDSS visible imaging of M51. Note the tidal tails, relationship of the spiral structure to the interaction, and the regularity of the dust. This looks much better in color, of course.

one of the most important developments in the entire history of science). I'll come back to this later, as it bears on questions we will encounter, such as "should I measure the age–mass relation or the color–magnitude relation?", and "should I extract the mean color as a function of environment, or the mean environment as a function of color?".

Philosophically, The primary goal of observational astrophysics—as distinct from, say, pure astronomy—ought to be to *rule out* theoretical physical models. Note that all important physical experiments have the property that they ruled out—or could have ruled out—one of the fundamental, dominant theories of their day. In physical cosmology, we have a fundamental, dominant theory for the growth of structure: Cold Dark Matter (CDM). Our role as observers is not to *bolster* this model, or find ad-hoc parameters we can add to the model that *make it consistent* with the data. Our role is to find experiments that have the power, even in the face of uncertainties (about, *eg*, how galaxies form), to rule out or substantially modify the fundamental assumptions of this theory. If an experiment does not have the power to rule out the theory, then it can hardly be said to provide substantial support for it if its results end up being in agreement!

That said, CDM may indeed turn out to be correct. It is *certainly* correct on large scales, indeed with its successful prediction of the angular spectrum of CMB inhomogeneities (yes, it was truly a prediction, a parameterized prediction, and the results *do* live in that parameterized but nonetheless highly constraining prediction space), the amplitude of the galaxy power spectrum, and the baryon feature now detected at low redshift, CDM is one of the best-tested theories in all of the physical sciences. But *these tests are all on large scales* (> 10 Mpc). On scales the size of galaxies and galaxy clusters, CDM is *not* well tested because on these scales the star-formation, black-hole-accretion, thermodynamic, and dissipative-evolution uncertainties come in. So lets harness the enormous power (the information) of our observations, and the rapidly improving power of our modeling, to make rock-hard tests of CDM on galaxy scales.

1.2 Why work at a redshift of one-tenth?

The principal disadvantage of working at low redshift is that the redshift *range* is small. Although we *do* see evolution (in, *eg*, the galaxy luminosity function, or mean galaxy specific star-formation rates) *within* low-redshift samples like SDSS [50] and 2MASS [42], we can measure evolution directly with much greater precision by comparing samples at very different redshifts.

On the other hand, the advantages of working at low redshift are many, and they all relate to the enormous amount of information we have about this epoch in the Universe. With SDSS, 2MASS, and GALEX [29] we have large enough solid angle and high enough sensitivity to see a large fraction of the galaxy population in a large fraction of the total volume of the Universe out to a redshift of one-tenth. In this review, I will focus on things we have learned from the SDSS data, simply because that is what I know best.

The SDSS is enormously *over-designed.* The spectra in the SDSS have far higher signal-to-noise than is necessary to obtain a redshift; indeed we measure not just redshifts but star-formation rates [23, 37], stellar population mixtures [37], and dust attenuations (both those affecting the lines [47] and those affecting the stars [23]).

But that's not all! The imaging in which these galaxies were selected for spectroscopy is far deeper than necessary for object selection; a typical Main Sample [44] galaxy target is detected in the SDSS imaging at a S/N of many hundreds. This means that we have very high-precision galaxy colors and magnitudes, sizes, concentrations, and surface brightnesses [21, 7].

Finally, and importantly for what follows, the large contiguous area of the SDSS (and other surveys) permits the study of clustering and galaxy environments; it also allows us to identify large, gravitationally bound systems, such as groups, clusters, and superclusters. Survey edges do not technically *prevent* the measurement of clustering and environments, but in practice, they make it difficult. So with the huge "volume-to-boundary" ratio of the SDSS, we have good-quality measures of galaxy environments, the mean galaxy density, and the correlation function on all scales out to the scale of large-scale homogeneity [20, 16] (these studies also demonstrate that the photometric calibration of the survey is very stable on large angular scales).

2 First example: galaxy–galaxy merger rate

Recall our goal: *Rule out the dominant theory.* And recall our approach: *Use the information in the observations.* What do these mean in practice? They mean that we should find the most robust predictions in the dominant theory, and find the "most informative" observational test. In principle, *any* observation is a test of theory, in that, in principle, it can be predicted. However, some predictions rely strongly on uncertain physics, including things like star-formation triggering and efficiency, mechanical stellar feedback, metallicity history, ionization history, and black-hole accretion feedback. Each of these processes affect the observed properties of galaxies at a level the magnitude (and sometimes even the sign!) of which is extremely uncertain.

The necessity of, in effect, "marginalizing" over these uncertainties removes almost all of the information in most observational data (or, equivalently, removes almost all of the precision of the prediction). So we want to find the observational tests that depend *least* on the unknown physical effects, and depend *most* on the fundamental physical model we want to attack.

In CDM, collapsed objects grow by accretion and merging. This process can be predicted with great precision for the dark sector, both analytically and with numerical simulations. Baryons only affect these predictions on the smallest scales, smaller than the "virial radii" of collapsed objects at the current day. So as long as we can find galaxies that are reliably associated with high-mass dark halos, we can use the statistics of their merging to very directly test the merging activity that CDM *requires*. The principal uncertainties are related to the relationship between galaxies and the halos that host them; this uncertainty is serious but *(a)* it is a limited, clearly circumscribed problem, and *(b)* there are many independent methods for investigating it, especially in the many approaches of those working on the Halo Occupation Distribution model for galaxy clustering [40, 52, 1, 6].

2.1 Close pairs

If we know anything about the relationship between galaxies and dark-matter halos, we know that the luminous red galaxies (LRGs) [14] in SDSS (super-L^* red, dead, early-type galaxies) live in very massive halos. Indeed, their abundance and clustering [51] very strongly constrain this relationship. In addition, they contain very simple and very similar (object-to-object) stellar populations [15], they have uniform mass-to-light ratios, and they don't tend to have difficult (for automated software) morphological features such as bars, HII regions, dust lanes, or spiral arms. LRGs are the natural galaxies to use for the study of merger rates, because they are "easy" to observe, and (relatively) easy to predict.

The only thing not "easy" about observing LRGs is that they are rare, so you need an enormous volume. The SDSS special LRG spectroscopic sample [14] fills $\sim 0.7\ h^{-3}\,\mathrm{Gpc}^3$ and contains, even after conservative cuts, 5×10^4 LRGs [20, 16].

We are looking for close pairs; the usual technical problem is that there can be chance superpositions. In principle we could use only close pairs where both have SDSS redshifts that agree (within a reasonable tolerance) but in practice we cannot (easily) because in 3/4 of the SDSS solid angle no two objects both get spectra if they are closer than a hardware-enforced limit of 55 arcsec (the other 1/4 of the survey area is covered by "overlap regions" of the spectroscopic fields and this repeat coverage essentially eliminates this constraint). We have excellent survey geometry and calibration, though, so we don't merely count close pairs, we cross-correlate the SDSS spectroscopic LRG sample with LRG targets in the imaging (whether or not they got spectra), in the two-dimensional plane of the sky, but scaled, for each spectroscopic LRG, to physical distances using the redshift of that galaxy (and in cross-correlating, we statistically remove chance interlopers at the "subtract the mean density" step). We then spherically deproject this projected $w_{\mathrm{p}}(r_{\mathrm{p}})$ to the three-dimensional, real-space correlation function $\xi(r)$ shown in Figure 2. Details are presented elsewhere [31].

I find this result remarkable in that the correlation function is so close to a power law over so many orders of magnitude, spanning the range ($\sim$ 100 Mpc) where the structure formation is linear and shows the baryon feature, through scales ($\sim$ 10 Mpc) where nonlinearities become important into the range ($\sim$ 1 Mpc) at which galaxy pairs tend to be within the same virialized, gravitationally bound system, all the way down to galaxy-sized scales ($\sim$ 10 kpc) at which dynamical friction, tidal forces, and dissipation must matter, at some level. But those are subjects for an entirely *different* review.

Because of dynamical friction (among other things), sufficiently close pairs of similar-mass

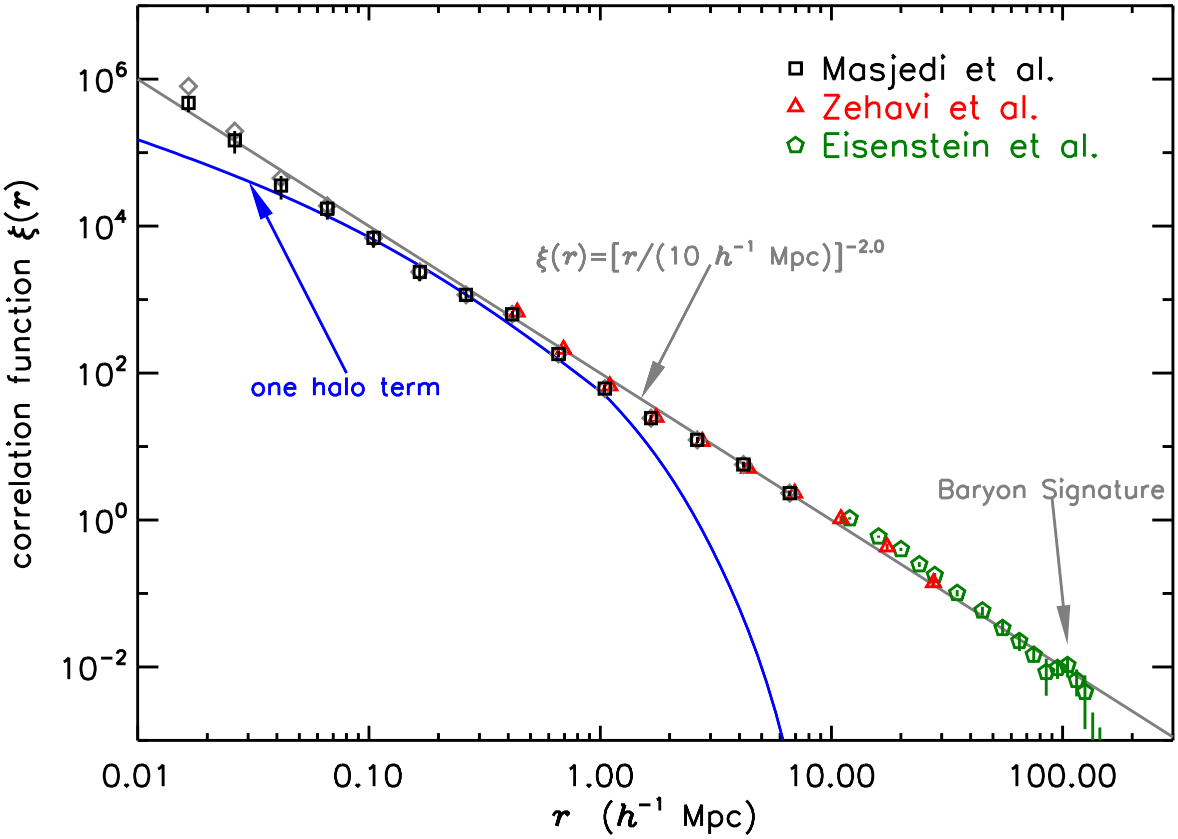

Figure 2: Three-space correlation function for LRGs, deprojected from an estimate of the projected function $w_{\mathrm{p}}(r_{\mathrm{p}})$ found by cross-correlating the spectroscopic and imaging samples, weighting to correct for the incompleteness due to SDSS spectrograph constraints or "fiber collisions" (from [31]). Also shown are a hand-drawn power-law, a NFW profile [33], and LRG correlation functions at intermediate scales [51] and large scales [16], the latter being a redshift-space function and therefore biased high on intermediate scales.

LRGs (almost all of our pairs are similar in mass, because the LRG luminosity selection is so strong) will merge in a dynamical time. In this picture, we can treat the correlation function $\xi(r)$ at the smallest scales as a quasi–steady-state inflow leading to the mergers of pairs of LRGs; we can straightforwardly turn the measured $\xi(r)$ into a *merger rate*. Really we put a *limit* on the merger rate, because not *all* close pairs will merge in a dynamical time; they might take much longer to merge, although they *cannot* take much less.

There is some length scale r_{f} (the exact value is not important, see below) inside of which dynamical friction is so effective that pairs at this separation merge in a dynamical time t_{dyn}, where the dynamical time is

$$t_{\mathrm{dyn}} \approx \frac{2\pi\, r_{\mathrm{f}}}{\sigma_v} \quad , \tag{1}$$

and σ_v is some characteristic gravitational velocity for the LRG in question. The average number N_{f} of neighbor LRGs within distance r_{f} of any "target" LRG is

$$N_{\mathrm{f}} \approx 4\pi\, r_{\mathrm{f}}^3\, n_{\mathrm{LRG}}\, \xi(r_{\mathrm{f}}) \quad , \tag{2}$$

where $n_{\rm LRG}$ is the space density of LRGs, and I have implicitly used that $\xi(r) \sim r^{-2}$ and $\xi(r_{\rm f}) \gg 1$. The LRG merger rate $\Gamma_{\rm LRG}$ is, therefore

$$\Gamma_{\rm LRG} < \frac{N_{\rm f}}{t_{\rm dyn}} \approx 2\, r_{\rm f}^2\, \sigma_v\, n_{\rm LRG}\, \xi(r_{\rm f}) \quad . \tag{3}$$

Since Figure 2 shows something very close to $\xi(r) \sim r^{-2}$, this expression for $\Gamma_{\rm LRG}$ does not depend strongly on the choice of $r_{\rm f}$, which is indeed poorly known.

This merger rate (LRG mergers per LRG) can be written as

$$\Gamma_{\rm LRG} < \frac{1}{160~{\rm Gyr}} \left[\frac{r_{\rm f}^2\, \xi(r_{\rm f})}{(10~{\rm kpc})^2\, 10^6}\right] \left[\frac{\sigma_v}{300~{\rm km\,s^{-1}}}\right] \left[\frac{n_{\rm LRG}}{10^{-4}~{\rm Mpc^{-3}}}\right] \quad , \tag{4}$$

ie, each LRG has less than a one percent probability of merging with another LRG each Gyr.

A few points deserve emphasis: *(a)* Unlike most of the results in the literature, there are no worries here about chance superpositions contaminating the results; we have deprojected to the three-space $\xi(r)$. We have the volume and the numbers to make that deprojection believeable. *(b)* This is a correlation function, not a "pair fraction," so the number density has been divided out. The pair fractions commonly discussed in the literature have the "instability" that they are proportional to the abundance (of secondaries) and that abundance can be an extremely strong function of magnitude, color, and surface-brightness cuts. Indeed, some of the discrepancies in the literature about pair fraction and its evolution may be caused by these instabilities. *(c)* This rate is a *strict upper limit*; two galaxies in a pair simply cannot in a time much shorter than the local dynamical time.

Our LRG–LRG rate limit (<1 percent per Gyr) is low, much lower than the rate at which LRG-hosting dark-matter halos merge with one another. On the other hand, the galaxies *in* halos merge much less frequently than the halos themselves. One frustrating aspect of the literature is that although there are some theoretical predictions [32], none of them is presented in quite the way you need to do a direct comparison. But we are very optimistic that this rate can be directly compared to models, for all the reasons given above (LRGs are simple and massive and in massive halos). We are also optimistic that we can repeat this experiment for a wide range of galaxy masses and mass ratios.

2.2 Post-starburst galaxies

Any measurement of a merger rate involves a measurement of an *abundance* of candidates and a *timescale* on which they will merge, are merging, or have merged. Often the timescale is quite uncertain, but there is one place where it is well known: in the passive evolution of starbursts.

There are two independent arguments—beyond the simple argument that stars probably form in disks and mergers are required to destroy those disks—that post-starburst galaxies lie along an "evolutionary sequence" connecting disk-dominated galaxies (spirals and late-type galaxies and irregulars) to bulge-dominated galaxies (lenticulars and ellipticals) via merging (that triggers the burst and disrupts the disks). The first is that bulge-dominated chemical abundance patterns differ from late-type patterns, and the difference can be explained if star-formation terminates with a large, rapid burst that exhausts all remaining fuel for star formation [49]. The second is that the photometric properties of post-starburst galaxies are such that their passive fading will cause them to rapidly evolve *not* to the properties of disks, but to the properties of bulges. We demonstrated this latter point with a sample of post-starburst galaxies taken from the SDSS [37] (see Figure 3).

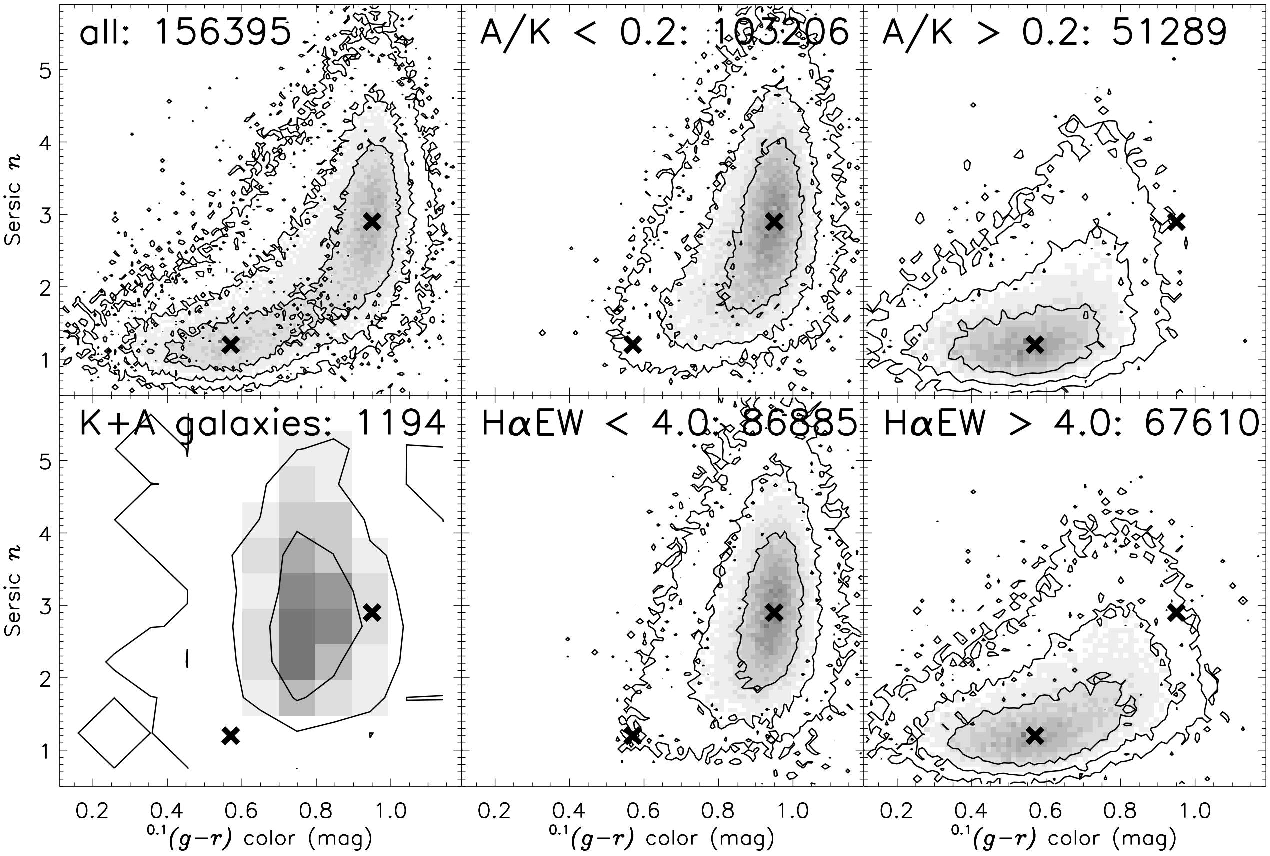

Figure 3: The distribution of all galaxies, star-forming galaxies, dead galaxies, and post-starburst (K+A) galaxies in the space of color and concentration (from [37]). The top left panel shows all galaxies, with crosses superimposed to guide the eye to the early-type (red, concentrated) and late-type (blue, exponential) galaxy populations, which separate nicely in this plane. The four panels on the right show the same for various cuts on A-star-based and OB-star-based measures of the recent-past star-formation history. The lower left panel shows the same for the post-starburst population. Note how they are concentrated and blue, as they will “fade into” the early-type location.

By making quasi-steady-state models of the distribution of A-star excesses in these galaxies, we can determine a *rate* density of their formation; we get $\sim 10^{-4}\,h^3\ \mathrm{Mpc}^{-3}\,\mathrm{Gyr}^{-1}$, or a bulge-dominated galaxy population that is growing by ~ 1 percent per Gyr.

What is this the rate *of*? Since some mergers are no-doubt “dry” (such as the LRG–LRG mergers discussed above), we can see this as a lower limit to the merger rate. However, since some of these starbursts might be triggered by very small accretion events or even tidal or internal perturbations, we can see this as an upper limit. With so many uncertainties, why do I bring this up in a review of *robust* tests of fundamental physics? Because the data are highly *informative,* providing a relatively precise rate density. Our work now is concentrated on constraining the nature of the triggering events; if we can work this out we might have a very strong constraint on merger activity in the local Universe.

2.3 Tidal features

The most straightforward method for measuring the merger rate is to simply look through an atlas of galaxies and identify galaxies that are clearly undergoing, or recovering from, a merger. There are many signatures, of varying veracity, including tidal arms, “S”-shaped distortions, shells, bars, warps, boxiness, and minor-axis dust lanes. I am very sympathetic to this procedure, and it pains me that it has not been executed definitively in the huge datasets we now have available to us in the nearby Universe.

That said, the morphological method, while straightforward, does not necessarily give you the best measure of the merger rate. Recall again that we need an *abundance* and a *timescale.* Identifying morphological signatures, and deciding on clear boundaries between “merger candidate” and “not” is more of an art than a science at this stage. In addition, the timescales over which these features are visible are not known, and they depend on the types of features, the events that raised them, and the fraction of bolometric output that comes from young stellar populations or dust-enshrouded star-formation. That said, there has been some remarkable quantitative work on the ages of particular morphological features, and reconstruction of individual galaxy–galaxy merger events [24, 39, 45]. There have also been some attempts to make automated classification [11].

My problems with the “morphological method” for determing the merger rate are many: *(a)* Features of importance can have a wide range of surface brightnesses; some appear in the UV (*eg*, the Antennae) and some appear in the near-IR (*eg*, the Whirlpool in Figure 1). *(b)* If you look hard enough, everything is “disturbed.” Indeed it has been discussed at this meeting that at “faint enough” surface brightnesses, essentially *all* early-type galaxies show tidal features [48]. Since (above) I constrained the major-merger rate to be low, either these features are raised by very small (and more numerous) accretion events or else they last an extremely long time. *(c)* Not all asymmetries [11] or “shell-like” features [28] are clear evidence for merging, as both can be created by tidal perturbations, small accretion events, localized star-formation episodes, and possibly cooling flows from halo plasma. *(d)* Even the features that *are* clearly signs of mergers can last—observably—for anything from the lifetime of an O star (a small fraction of a dynamical time) to many dynamical times, depending on star-formation activity and the “coldness” of the progenitors.

Fundamentally, until these issues are cleared up, I am bearish on constraining CDM with morphological studies.

3 Second example: galaxy environments

Different parts of the Universe, with different local densities, evolve with different ratios of gravitational-collapse, star-formation, and galaxy-dynamical timescales.

3.1 Environment measures

Unfortunately, there are no *good* measures of local density! Counting galaxies in close volumes is subject to small-number Poisson statistics (or worse). In principle, tesselations or transverse distances to Nth-nearest neighbor give precise numbers. However, these kinds of measures have ill-defined length-scale and are subject to large jumps when sample definition (*eg*, magnitude cut) is changed.

We have shown that all the information about environment we see is captured by environment measures on ~ 1 Mpc scales [9], which happens (not coincidentally, I imagine) to be on the order of the virial scale. This has recently been controversial [2], but I think the point is settled now.

3.2 Star-formation history and environment

It has been observed for some time now that specific star-formation rates are different, on average, in different environments [25, 19, 3, 30, 27, 17, 7]. In CDM, the most *massive* halos are in the most dense environments, so shouldn't we expect *total stellar mass* to be an *even stronger* function of environment? As we emphasize here and below, we don't have to think about these questions separately—or theoretically:

Figure 4 shows the mean environment as a function of color and absolute magnitude. It separates the current star-formation rate (color) and the total time-averaged star-formation rate (luminosity), and shows that it is the former, not the latter, that relates most closely to environment around L^*. Only the very most massive galaxies show a strong relationship between total stellar mass and environment.

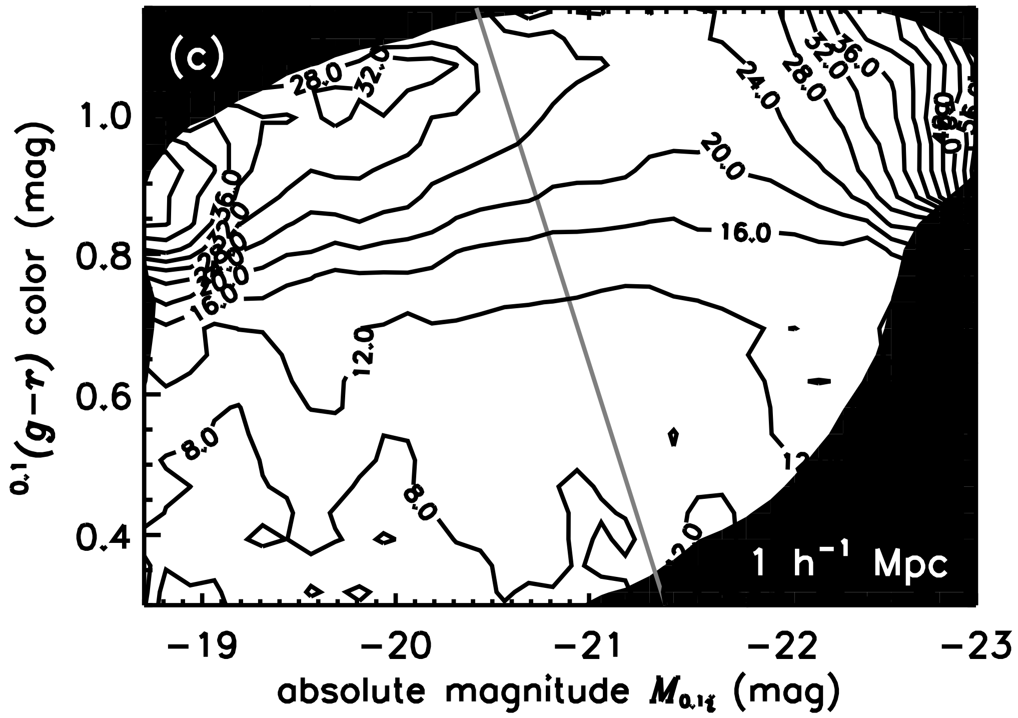

Figure 4: The mean overdensity (measured in a gaussian *spherical* de-projected window [13] of radius 1 Mpc) of the environments of galaxies as a function of galaxy color and luminosity (from [22]). The fact that the contours are close to horizontal over most of the plot shows that it is color, not luminosity, that is most closely related to environmental density for most of the luminosity range; luminosity only "matters" at high luminosity.

In current models, the relationship shown in Figure 4 is simply explained in terms of satellite and central galaxies in groups [5]; at most luminosities, most galaxies are central, not satellites, but at fixed luminosity, the satellites tend to be redder and in higher density environments (because at early times they fell into halos massive enough to contain more than one galaxy of their luminosity).

You might ask, in Figure 4, why we averaged the *environment* as a function of galaxy properties, and not the other way around. I noted that environmental density can never be measured at high signal-to-noise. Color and magnitude, on the other hand, are measured in SDSS data with percent precision (though perhaps not percent *accuracy*). It would be insane to average away that precision, and split the sample by a noisy environmental measure (thereby effectively smoothing any important relationships by the uncertainties in the environmental measures). Most investigators feel that it is somewhat more "natural" to look at galaxy properties as a function of environment (*ie*, the wrong way around), but my point is that my way is *more informative* (objectively), and any theoretical model can predict the relationship either way!

You might also ask, in Figure 4, why we showed the color and luminosity dependences of environmental density, and not the "age" and "stellar mass" dependencies. This issue is somewhat subtle, and it depends on your goals, but I note the following: The color-magnitude diagram of galaxies contains a lot of Shannon *information* in its narrow red sequence and less-narrow blue sequence [7]. Any honest transformation to age and mass involves modeling uncertainties that will (and do [10, 34]) "smooth" the diagram, destroying information. Obviously, comparisons with theoretical models involve making these uncertain transformations. The question is, who should make them, the theorist or the observer? Personally, I respect the information in the data; it is anathema to transform and lose this richness.

3.3 There is no morphology–density relation (we can find)

In some sense, the most important "fact" of galaxy evolution is the observation that *(a)* the morphological mix is different in different environments [12, 35]. At the same time, *(b)* galaxy morphology is related to star-formation history (and a lot of other stuff) [46, 26, 38]. I just showed that *(c)* star-formation history is related to environmental density. Are all three relations on an *equal footing*, or is one the product of the other two? We find ourselves in the position of epidemiologists, trying to understand the causes of asthma in the face of strong relationships between race, socio-economic status, diet, and living conditions. Fortunately we have great data, and a lot of it, more and better than the epidemiologists (I am sad to say).

In Figure 5 I show the variation of concentration (a measure of bulge/total ratio and therefore a morphology surrogate [43, 18]) on environment (clustocentric distance, a high-precision but variable-scale environment measure), in narrow color slices. In Figure 6 I show the variation of color on environment in narrow concentration slices. These Figures look very different: Color depends on environment independently of morphology, morphology *does* not independently of color. Of course I am over-stating the result by calling this "morphology"; take it as you wish, but it is clear that color and concentration are *not* on an equal footing when we ask what they can tell us about environment.

In current thinking about the *reasons* for the morphology–density relation (things like ram-pressure stripping, mergers, late accretion, tidal perturbations), these results are very difficult to understand. What physical processes can "tell" galaxy star-formation rates about their environments and "tell" morphologies to keep track of star-formation rates but not do much to the morphologies independently? I think the conclusion has to be that the processes that set morphology (or, really, concentraion or bulge/total ratio) are somehow *internal* to the galaxies.

4 Future challenges

4.1 Short-term

For the theorists, I have two requests: *(a)* Explain or predict our LRG–LRG merger rate and more merger rates to come. *(b)* I didn't discuss this, but convincingly explain the luminosity

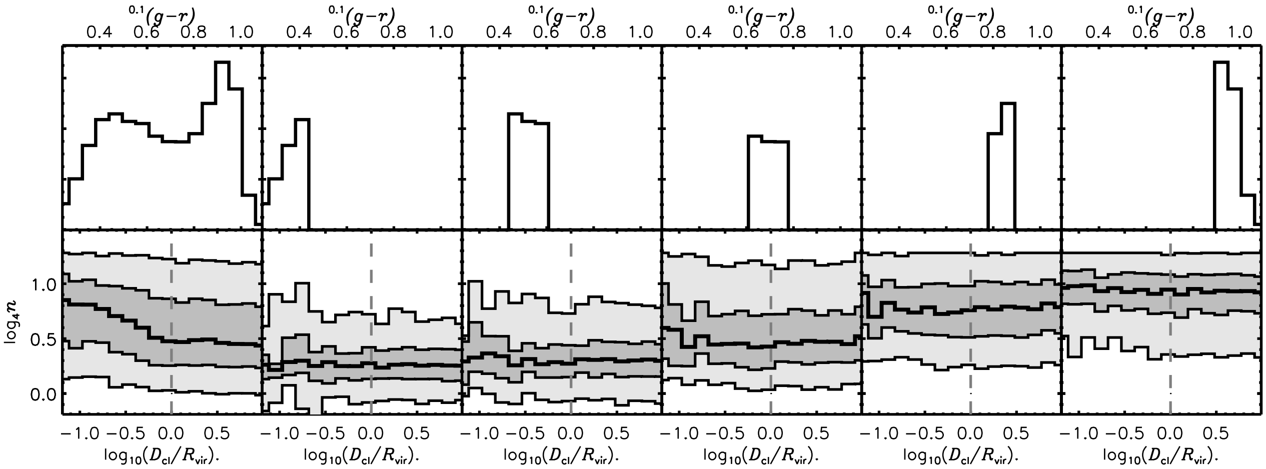

Figure 5: The top left panel shows the color distribution for a subsample of the SDSS Main Sample galaxies with redshifts $0.015 < z < 0.068$. The bottom left panel shows the dependence of weighted quantiles of concentration (Sérsic index) on environment (clustocentric distance; details in [36]). The quantiles shown are 5 percent, 25 percent (first quartile), 50 percent (median), 75 percent (third quartile), and 95 percent. Note the strong dependence of concentration on evironment in the bottom left panel. In each of the next pairs of panels, the top shows the color distribution for a color-selected 20-percent quantile of the data, and the bottom shows the dependence of concentration on evironment within that color quantile. Note that there is *very little dependence of concentration on environment at a given color.* (Figure adapted from [36].)

function, especially its remarkable difference from the mass function of collapsed objects. After years of work, this has still not been done [4].

For the observers, I have two more: *(c)* Separate the kinematically cold and kinematically hot parts of galaxies, without relying on the almost totally unjustified radial profiles used in bulge/disk decomposition. God didn't tell us that disks are exponential, and he *certainly* didn't tell us that bulges are de Vaucouleurs. Your efforts will be rewarded, because the most sophisticated modern theoretical models *do* distinguish these components; they make non-trivial predictions. *(d)* Respond to my objections (above) and measure the merger rate morphologically.

4.2 Long-term

Long term, we need to have simulations that *look like galaxies* on all the scales and at all the wavelengths we can observe. Recall Figure 1. Perhaps when these simulations exist we will understand what sets galaxy morphologies.

If I am to set myself a *very* long-term goal so absurdly difficult it will take much longer than my own natural lifetime, a goal that will make use of literally *all* of the information in the observations, and a goal that is the extrapolation to $t = \infty$ of the research program I have started here, it would be this: Construct a chain of *constrained realizations* of the entire observable Universe, a chain (as in a Markov chain or equivalent) that can be used to marginalize over *all* unknowns, both in the data and in the theory. Each "link" in the chain would be a simulation or model of the entire observable Universe, consistent with every existing observation (*ie*, it would get all the *phases* right, not just the amplitudes and dynamics).

It sounds ridiculous, but with the next (or next to next) generation of multi-scale simulations, it might be possible. Of course there would be huge degeneracies, both because our obser-

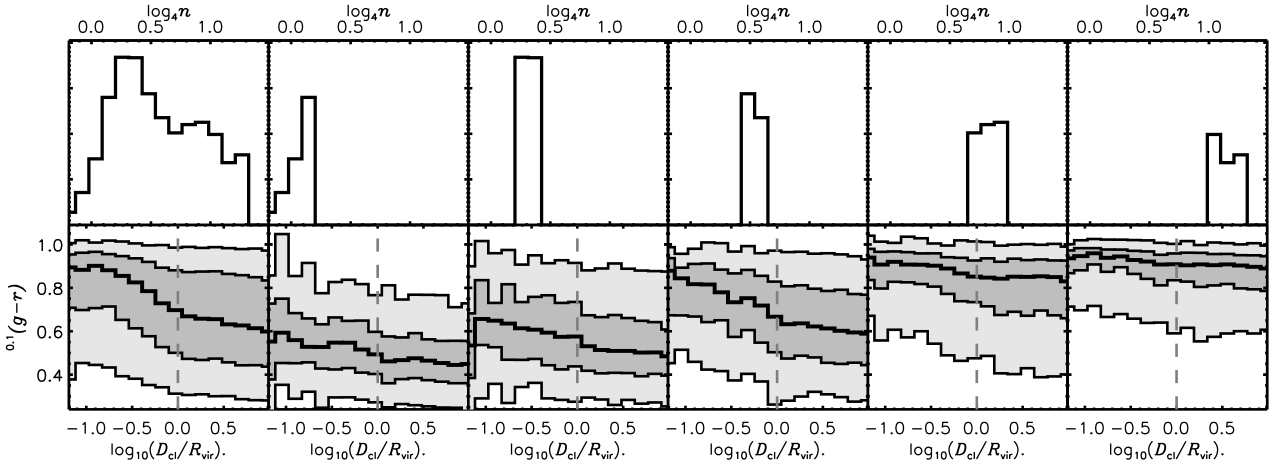

Figure 6: The same as Figure 5 but with color and concentration reversed; *ie*, this figure shows the dependence of color on environment within concentration-selected quantiles. It shows that color *does* depend on evironment, even at fixed concentration (adapted from [36]).

vations are incomplete and because large parts of the physical model are uncertain (especially those parts relating to small-scale baryonic physics). Each new observation or improvement in theoretical understanding would rule out parts of the chain, and permit finer sampling in the (now smaller) allowed regions of parameter space. This chain of realizations could be used to tell us the initial conditions, evolutionary history (and a "neighborhood" around those still consistent with the data and theory) for every *actual* galaxy we have observed, and provide a clear guide for further observational testing.

Indeed, those predicted observables—things to observe but not yet observed—that show enormous variation among the realizations in the chain will be the decisive observations for the future. Those predicted observables that show the least variation will be the *most fundamental* predictions of the physical cosmology. A chain of constrained realizations is a tool for the *objective* pursuit of the objectives I set out in the introduction; it could provide us with a quantitative ranking of the importance of each possible observational experiment, a ranking by its relative capability of *ruling out the dominant physical theory.* If we can't rule it out, then we are just "constructing" it "socially".

Acknowledgements. All of this work was done in very close collaboration with Mike Blanton at NYU, to whom I owe enormous thanks; in addition, Andreas Berlind, Richard Cool, Daniel Eisenstein, Morad Masjedi, Jim Peebles, Alex Quintero, Sam Roweis, and Beth Willman contributed to various parts. This research was supported by NASA (NAG5-11669) and the NSF (AST-0428465), made use of the NASA Astrophysics Data System, and employed open-source software written and maintained by David Schlegel, Doug Finkbeiner, Mike Blanton, Nikhil Padmanabhan, myself, and others. All the data shown in this review have been released publicly by Mike Blanton and myself as part of the NYU-VAGC [8].

References

[1] K. Abazajian, Z. Zheng, I. Zehavi, D. H. Weinberg, J. A. Frieman, A. A. Berlind, M. R. Blanton, N. A. Bahcall, J. Brinkmann, D. P. Schneider, and M. Tegmark. Cosmology and

the Halo Occupation Distribution from Small-Scale Galaxy Clustering in the Sloan Digital Sky Survey. ApJ, 625:613–620, June 2005.

[2] M. Balogh et al. Galaxy ecology: groups and low-density environments in the SDSS and 2dFGRS. MNRAS, 348:1355–1372, March 2004.

[3] M. L. Balogh, D. Christlein, A. I. Zabludoff, and D. Zaritsky. The environmental dependence of the infrared luminosity and stellar mass functions. ApJ, 557:117–125, August 2001.

[4] A. J. Benson, R. G. Bower, C. S. Frenk, C. G. Lacey, C. M. Baugh, and S. Cole. What shapes the luminosity function of galaxies? ApJ, 599:38–49, December 2003.

[5] A. A. Berlind, M. R. Blanton, D. W. Hogg, D. H. Weinberg, R. Davé, D. J. Eisenstein, and N. Katz. Interpreting the relationship between galaxy luminosity, color and environment. ApJ, 629:625–632, 2005.

[6] A. A. Berlind, J. A. Frieman, D. H. Weinberg, M. R. Blanton, M. Warren, K. Abazajian, R. Scranton, D. W. Hogg, and R. Scoccimarro. Galaxy Groups and Clusters in the SDSS Redshift Survey I: Identification, Catalog, and the Multiplicity Function. ApJ, 2006.

[7] M. R. Blanton et al. The broadband optical properties of galaxies with redshifts $0.02 < z < 0.22$. ApJ, 594:186–207, September 2003.

[8] M. R. Blanton et al. New York University Value-Added Galaxy Catalog: A galaxy catalog based on new public surveys. AJ, 129:2562–2578, June 2005.

[9] M. R. Blanton et al. The scale-dependence of relative galaxy bias: Encouragement for the halo model description. ApJ, submitted (astro-ph/0411037), 2005.

[10] J. Brinchmann, S. Charlot, S. D. M. White, C. Tremonti, G. Kauffmann, T. Heckman, and J. Brinkmann. The physical properties of star-forming galaxies in the low-redshift Universe. MNRAS, 351:1151–1179, July 2004.

[11] C. J. Conselice, M. A. Bershady, M. Dickinson, and C. Papovich. A direct measurement of major galaxy mergers at $z <\sim 3$. AJ, 126:1183–1207, September 2003.

[12] A. Dressler. Galaxy morphology in rich clusters—Implications for the formation and evolution of galaxies. ApJ, 236:351–365, 1980.

[13] D. J. Eisenstein. Deprojecting densities from angular cross-correlations. ApJ, 586:718–722, April 2003.

[14] D. J. Eisenstein et al. Spectroscopic target selection for the Sloan Digital Sky Survey: The Luminous Red Galaxy Sample. AJ, 122:2267–2280, November 2001.

[15] D. J. Eisenstein et al. Average spectra of massive galaxies in the SDSS. ApJ, 585:694–713, March 2003.

[16] D. J. Eisenstein et al. Detection of the baryon acoustic peak in the large-scale correlation function of SDSS Luminous Red Galaxies. ApJ, 633:560–574, 2005.

[17] Percy Gomez et al. Galaxy star formation as a function of environment in the Early Data Release of the Sloan Digital Sky Survey. ApJ, 584:210–227, February 2003.

[18] T. Goto, C. Yamauchi, Y. Fujita, S. Okamura, M. Sekiguchi, I. Smail, M. Bernardi, and P. L. Gomez. The morphology-density relation in the Sloan Digital Sky Survey. MNRAS, 346:601–??, 2003.

[19] Y. Hashimoto, A. Oemler, H. Lin, and D. L. Tucker. ?? ApJ, 499:589–??, 1998.

[20] D. W. Hogg, D. J. Eisenstein, M. R. Blanton, N. A. Bahcall, J. Brinkmann, J. E. Gunn, and D. P. Schneider. Cosmic homogeneity demonstrated with luminous red galaxies. ApJ, 624:54–58, May 2005.

[21] D. W. Hogg et al. The luminosity density of red galaxies. AJ, 124:646–651, 2002.

[22] David W. Hogg et al. The overdensities of galaxy environments as a function of luminosity and color. ApJ, 585:L5–L9, March 2003.

[23] G. Kauffmann et al. Stellar masses and star formation histories for 10^5 galaxies from the Sloan Digital Sky Survey. MNRAS, 341:33–53, May 2003.

[24] M. Kaufman, E. Brinks, B. G. Elmegreen, D. M. Elmegreen, M. Klarić , C. Struck, M. Thomasson, and S. Vogel. The interacting galaxies NGC 5394/5395: A post-ocular galaxy and its ring/spiral companion. AJ, 118:1577–1608, October 1999.

[25] R. C. Kennicutt. On the evolution of the spiral galaxies in the Virgo cluster. AJ, 88:483–488, April 1983.

[26] J. Kormendy and S. Djorgovski. Surface photometry and the structure of elliptical galaxies. ARA&A, 27:235–277, 1989.

[27] I. Lewis et al. The 2dF Galaxy Redshift Survey: The environmental dependence of galaxy star formation rates near clusters. MNRAS, 334:673–683, August 2002.

[28] D. F. Malin and D. Carter. A catalog of elliptical galaxies with shells. ApJ, 274:534–540, November 1983.

[29] D. C. Martin et al. The Galaxy Evolution Explorer: A space ultraviolet survey mission. ApJ, 619:L1–L6, January 2005.

[30] H. J. Martínez, A. Zandivarez, M. Domínguez, M. E. Merchán, and D. G. Lambas. Galaxy groups in the 2dF Galaxy Redshift Survey: Effects of environment on star formation. MNRAS, 333:L31–LL34, June 2002.

[31] Morad Masjedi et al. The very small-scale clustering and merger rate of luminous red galaxies. ApJ, submitted, 2006.

[32] C. Murali, N. Katz, L. Hernquist, D. H. Weinberg, and R. Davé. The Growth of Galaxies in Cosmological Simulations of Structure Formation. ApJ, 571:1–14, May 2002.

[33] J. F. Navarro, C. S. Frenk, and S. D. M. White. A Universal Density Profile from Hierarchical Clustering. ApJ, 490:493–+, December 1997.

[34] N. Padmanabhan, U. Seljak, M. A. Strauss, M. R. Blanton, G. Kauffmann, D. J. Schlegel, C. Tremonti, N. A. Bahcall, M. Bernardi, J. Brinkmann, M. Fukugita, and Ž. Ivezić. Stellar and dynamical masses of ellipticals in the Sloan Digital Sky Survey. *New Astronomy*, 9:329–342, June 2004.

[35] M. Postman and M. Geller. The morphology-density relation—The group connection. ApJ, 281:95–99, June 1984.

[36] Alejandro D. Quintero, Andreas A. Berlind, Michael R. Blanton, and David W. Hogg. What galaxies know about their nearest cluster. ApJ, submitted, 2006.

[37] Alejandro D. Quintero et al. Selection and photometric properties of K+A galaxies. ApJ, 602:190–199, 2004.

[38] M. S. Roberts and M. P. Haynes. Physical parameters along the Hubble sequence. ARA&A, 32:115–152, 1994.

[39] H. Salo and E. Laurikainen. N-body model for M51 I. Multiple encounter versus single passage? MNRAS, 319:377–392, December 2000.

[40] R. Scoccimarro, R. K. Sheth, L. Hui, and B. Jain. How Many Galaxies Fit in a Halo? Constraints on Galaxy Formation Efficiency from Spatial Clustering. ApJ, 546:20–34, January 2001.

[41] Claude Elwood Shannon. *The mathematical theory of communication.* University of Illinois Press, Urbana, IL, 1949.

[42] M. F. Skrutskie et al. The Two Micron All Sky Survey (2MASS): Overview and status. In *ASSL Vol. 210: The Impact of Large Scale Near-IR Sky Surveys*, page 25, 1997.

[43] I. Strateva, Željko Ivezić, G. R. Knapp, V. K. Narayanan, M. A. Strauss, J. E. Gunn, R. H. Lupton, D. Schlegel, N. A. Bahcall, J. Brinkmann, R. J. Brunner, T. ;. Budavári, I. ;. Csabai, F. J. Castander, M. Doi, M. Fukugita, Z. Győry, M. Hamabe, G. Hennessy, T. Ichikawa, P. Z. Kunszt, D. Q. Lamb, T. A. McKay, S. Okamura, J. Racusin, M. Sekiguchi, D. P. Schneider, K. Shimasaku, and D. York. Color Separation of Galaxy Types in the Sloan Digital Sky Survey Imaging Data. AJ, 122:1861–1874, October 2001.

[44] M. A. Strauss et al. Spectroscopic target selection in the Sloan Digital Sky Survey: The Main Galaxy Sample. AJ, 124:1810–1824, September 2002.

[45] C. Struck and B. J. Smith. Models of the morphology, kinematics, and star formation history of the prototypical collisional starburst system NGC 7714/7715 = ARP 284. ApJ, 589:157–178, May 2003.

[46] G. A. Tammann, A. Yahil, and A. Sandage. The velocity field of bright nearby galaxies. II—Luminosity functions for various Hubble types and luminosity classes—The peculiar motion of the local group relative to the Virgo cluster. ApJ, 234:775–784, December 1979.

[47] C. A. Tremonti et al. The original of the mass-metallicity relation: insights from 53,000 star-forming galaxies in the SDSS. ApJ*, in press*, (astro-ph/0405537), 2004.

[48] P. G. van Dokkum. The recent and continuing assembly of field ellipticals by red mergers. AJ, in press (astro-ph/0506661), 2005.

[49] G. Worthey. Abundance Ratio Trends and Nucleosynthesis in Elliptical Galaxies and Spheroids. PASP, 110:888–899, August 1998.

[50] D. York et al. The Sloan Digital Sky Survey: Technical summary. AJ, 120:1579–1587, 2000.

[51] I. Zehavi, D. J. Eisenstein, R. C. Nichol, M. R. Blanton, D. W. Hogg, J. Brinkmann, J. Loveday, A. Meiksin, D. P. Schneider, and M. Tegmark. The Intermediate-Scale Clustering of Luminous Red Galaxies. ApJ, 621:22–31, March 2005.

[52] I. Zehavi et al. On departures from a power law in the galaxy correlation function. ApJ, 608:16–24, June 2004.