Large Synoptic Survey Telescope Galaxies, Dark Matter, and Black Holes: Extragalactic Roadmap

Robertson, Brant[1], Banerji, M.[2], Cooper, Michael[3], Davies, R.[4], Ferguson, Henry C.[5], Kaviraj, S.[6], Lintott, C.[4], Lotz, J.[5], Newman, J.[7], Norman, D.[8], Padilla, N.[9], Schmidt, S.[10], Verma, A.[4], Working Group Participants, Collaboration Members

[1]*University of California, Santa Cruz,* [2]*Cambridge University,* [3]*University of California, Irvine,* [4]*Oxford University,* [5]*Space Telescope Science Institute,* [6]*University of Hertfordshire,* [7]*University of Pittsburgh,* [8]*National Optical Astronomical Observatory,* [9]*Pontifica Universidad Catolica de Chile,* [10]*University of California, Davis,*

Version December 20, 2016

Inquiries about this report or its content can be addressed to Brant Robertson (brant@ucsc.edu).

Abstract.
TBD

Contents

1. Introduction

The Large Synoptic Survey Telescope (LSST) is a wide-field, ground-based telescope, designed to image a substantial fraction of the sky in six optical bands every few nights. It is planned to operate for a decade allowing the stacked images to detect galaxies to redshifts well beyond unity. The LSST and the survey are designed to meet the requirements (Ivezic & the LSST Science Collaboration 2011) of a broad range of science goals in astronomy, astrophysics and cosmology. The LSST was the top-ranked large ground-based initiative in the 2010 National Academy of Sciences decadal survey in astronomy and astrophysics, and is on track to begin the survey early in the next decade.

In 2008, eleven separate quasi-independent science collaborations were formed to focus on a broad range of topics in astronomy and cosmology that the LSST could address. Members of these collaborations have been instrumental in helping to develop the science case for LSST (encapsulated in the LSST Science Book), to refine the concepts for the survey and for the data processing, and to educate other scientists and the public about the promise of this unique observatory.

The Dark Energy Science Collaboration (DESC) has taken the next logical step beyond the science book. They identified they most critical challenges that will need to be overcome to realize LSST's potential for measuring the effects of Dark Energy. They looked at five complementary techniques for tackling dark energy, and outlined high-priority tasks for the science collaboration during construction. They designated sixteen working groups (some of which already existed) to coordinate the work. This roadmap has been documented in a 133-page white paper (arxiv.org/abs/1211.0310). The white paper provides a guide for investigators looking for ways to contribute to the overall investigation. It may help in efforts to obtain funding, because it provides clear indication of the importance of the advance work and how the pieces fit together.

The investigation of Dark Energy is only one topic for LSST. It is important to develop similarly concrete roadmaps for work in other areas. After some discussion among the collaborations, it appears useful in some cases for different science collaborations to join forces on a single whitepaper. This is particularly true for topics that involve observations of distant galaxies. With the advent of the DESC, some of the science goals of the large-scale-structure, weak-lensing, and strong-lensing collaborations have found a new home. The remaining science goals of those collaborations tend to be focused on galaxy evolution and dark matter. Two other collaborations: AGN and Galaxies, also have those topics as major themes. This roadmap identifies the major high-level science themes of these investigations, outlines how complementary techniques will contribute, and identifies areas where advance work is essential. For this advance work, the emphasis is on areas that are not adequately covered in the DESC roadmap. As convenient shorthand, we use the acronym GALLA (Galaxies, AGN, Lensing Large-scale Structure and Astro-informatics) joint roadmap of the overlapping science collaborations.

Chapter 2 gives a brief summary of the science background. Many of the themes and projects are already set out in the Science Book, where more detail is provided for many of the science investigations. Chapter 3 sets out the highest priority preparatory work to enable these investigations. These tasks are laid out on the assumption that the work plan of the DESC will be carried out and that software and data products resulting from that work will be available to other science collaborations. The Appendix 4 organizes the tasks by science topic and desribes them in more detail.

2. Science Background

TBD

2.1 Overview

TBD

2.2 Science Background: AGN

2.3 Science Background: Galaxies

Galaxies represent fundamental astronomical objects outside our own Milky Way. The large luminosities of galaxies enable their detection to extreme distances, providing abundant and far-reaching probes into the depths of the universe. At each epoch in cosmological history, the color and brightness distributions of the galaxy population reveal how stellar populations form with time and as a function of galaxy mass. The progressive mix of disk and spheroidal morphological components of galaxies communicate the relative importance of energy dissipation and collisionless processes for their formation. Correlations between internal galaxy properties and cosmic environments indicate the ways the universe nurtures galaxies as they form. The evolution of the detailed characteristics of galaxies over cosmic time reflects how fundamental astrophysics operates to generate the rich variety of astronomical structures observed today.

Study of the astrophysics of galaxy formation represents a vital science of its own, but the ready observability of galaxies critically enables a host of astronomical experiments in other fields. Galaxies act as the semaphores of the universe, encoding information about the development of large scale structures and the mass-energy budget of the universe in their spatial distribution. The mass distribution and clustering of galaxies reflect essential properties of dark matter, including potential constraints on the velocity and mass of particle candidates. Galaxies famously host supermassive black holes, and observations of active galactic nuclei provide a window into the high-energy astrophysics of black hole accretion processes. The porous interface between the astrophysics of black holes, galaxies, and dark matter structures allows for astronomers to achieve gains in each field using the same datasets.

The Large Synoptic Survey Telescope (LSST) will provide a digital image of the southern sky in six bands (*ugrizy*). The area ($\sim 18,000$ deg^2) and depth ($r \sim 24.5$ for a single visit, $r \sim 27.5$ coadded) of the survey will enable research of such breadth that LSST may influence essentially all extragalactic science programs that rely primarily on photometric data. For studies of galaxies, LSST provide both an unequaled catalogue of billions of extragalactic sources and high-quality multiband imaging of individual objects. This section of the *Extragalactic Roadmap* presents scientific background for studies of these galaxies with LSST to provide a context for considering how the astronomical community can best leverage the catalogue and imaging datasets and for identifying any required preparatory science tasks.

LSST will begin science operations during the next decade, more than twenty years after the start of the Sloan Digital Sky Survey (**york2000a**) and subsequent precursor surveys including PanSTARRS (**kaiser2010a**), the Subaru survey with Hyper Suprime-Cam (**miyazaki2012a**), and the Dark Energy Survey (**flaugher2005a**). Relative to these prior efforts, extragalactic science breakthroughs generated by LSST will likely benefit from its increased area, source counts, and statistical samples, the constraining power of the six-band imaging, and the survey depth and image quality. The following discussion of LSST efforts focusing on the astrophysics of galaxies will highlight how these features of the survey enable new science programs.

2.3.1 Star Formation and Stellar Populations in Galaxies

Light emitted by stellar populations will provide all the direct measurements made by LSST. This information will be filtered through the six passbands utilized by the survey (*ugrizy*), providing constraints on the rest-frame ultraviolet SEDs of galaxies to redshift $z \sim 6$ and a probe of rest-frame optical spectral breaks to $z \sim 1.5$. By using stellar population synthesis modeling, these measures of galaxy SEDS will enable estimates of the redshifts, star formation rates, stellar masses, dust content, and population ages for potentially billions of galaxies. In the context of previous extragalactic surveys, LSST will enable new advances in our understanding of stellar populations in galaxies by contributing previously unachievable statistical power and an areal coverage that samples the rarest cosmic environments.

A variety of ground- and space-based observations have constrained the star formation history of the universe over the redshift range that LSST will likely probe (**madau2014a**). The statistical power of LSST will improve our knowledge of the evolving UV luminosity function, luminosity density, and cosmic star formation rate. The LSST observations can constrain how the astrophysics of gas cooling within dark matter halos, the efficiency of molecular cloud formation and the star formation within them, and regulatory mechanisms like supernova and radiative heating give rise to these statistical features of the galaxy population. While measurement of the evolving UV luminosity function can help quantify the role of these astrophysical processes, the ability of LSST to probe vastly different cosmic environments will also allow for the robust quantification of any changes in the UV luminosity function with environmental density, and an examination of connections between environment and the fueling of star formation.

Optical observations teach us about the established stellar content of galaxies. For stellar populations older than ~ 100 million years, optical observations provide sensitivity to the spectral breaks near a wavelength of $\lambda \approx 4000$ in the rest-frame related to absorption in the atmospheres of mature stars.

Such observations help constrain the amount of stellar mass in galaxies. For passive galaxies that lack vigorous star formation, these optical observations reveal the well-defined "red sequence" of galaxies in the color-magnitude plane that traces the succession of galaxies from recently-merged spheroids to the most massive systems at the centers of galaxy clusters. For blue, star-forming galaxies, optical light can help quantify the relative contribution of evolved stars to total galaxy luminosity, and indeed has led to the identification of a well-defined locus of galaxies in the parameter space of star formation rate and stellar mass (**noeske2007a**). This relation, often called the "star-forming main sequence" of galaxies, indicates that galaxies of the same stellar mass typically sustain a similar star-formation rate. Determining the physical or possibly statistical origin of the relation remains an active line of inquiry, guided by recently improved data from Hubble Space Telescope over the ~ 0.2 deg^{-2} Cosmic Assembly Near-Infrared Deep Extragalactic Survey (**grogin2011a**; **koekemoer2011a**). While LSST will be comparably limited in redshift selection, its $30,000$ times larger area will enable a much fuller sampling of the star formation–stellar mass plane, allowing for a characterization of the distribution of galaxies that lie off the main sequence that can help discriminate between phenomenological explanations of the sequence.

2.3.2 Galaxies as Cosmic Structures

The structural properties of galaxies arise from an intricate combination of important astrophysical processes. The gaseous disks of galaxies require substantial energy dissipation while depositing angular momentum into a rotating structure. These gaseous disks form stars with a surface density that declines exponentially with galactic radius, populating stellar orbits that differentially rotate about the galactic center and somehow organize into spiral features. Many disk galaxies contain (pseduo-)bulges that form through a combination of violent relaxation and orbital dynamics. These disk galaxy features contrast with systems where spheroidal stellar distributions dominate the galactic structure. Massive ellipticals form through galaxy mergers and accretions, and manage to forge a regular sequence of surface density, size, and stellar velocity dispersion from the chaos of strong gravitational encounters. Since these astrophysical processes may operate with great variety as a function of galaxy mass and cosmic environment, LSST will revolutionize studies of evolving galaxy morphologies by providing enormous samples with deep imaging of exquisite quality.

The huge sample of galaxies provided by LSST will provide a definitive view of how the sizes and structural parameters of disk and spheroidal systems vary with color, stellar mass, and luminosity. Morphological studies will employ at least two complementary techniques for quantifying the structural properties of galaxies. Bayesian methods can yield multi-component parameterized models for all the galaxies in the LSST sample, including the quantified contribution of bulge, disk, and spheroid structures to the observed galaxy surface brightness profiles. The parameterized models will supplement non-parametric measures of the light distribution including the Gini and M20 metrics that quantify the surface brightness uniformity and spatial moment of dominant pixels in a galaxy image (**abraham2003a**; **lotz2004a**). Collectively, these morphological measures provided by analyzing the LSST imaging data will enable a consummate determination of the relation between structural properties and other features of galaxies over a range of galaxy mass and luminosity previously unattainable.

While the size of the LSST sample supplies the statistical power for definitive morphological studies, the sample size also enables the identification of rare objects. This capability will benefit our efforts for connecting the distribution of galaxy morphologies to their evolutionary origin during the structure formation process, including the formation of disk galaxies. The emergence of ordered disk galaxies remains a hallmark event in cosmic history, with so-called "grand design" spirals like the Milky Way forming dynamically cold, thin disks in the last ~ 10 Gyr. Before thin disks emerged, rotating systems featured "clumpy" mass distributions with density enhancements that may originate from large scale gravitational instability. Whether the ground-based LSST can effectively probe the exact timing and duration of the transition from clumpy to well-ordered disks remains unknown, but LSST can undoubtedly contribute studying the variation in forming disk structures at the present day. Unusual objects, such as the UV luminous local galaxies identified by **heckman2005a** that display physical features analogous to Lyman break galaxies at higher redshifts, may provide a means to study the formation of disks in the present day under rare conditions only well-probed by the sheer size of the LSST survey.

Similarly, the characterizing the extremes of the massive spheroid population can critically inform theoretical models for their formation. For instance, the most massive galaxies at the centers of galaxy clusters contain vast numbers of stars within enormous stellar envelopes. The definitive LSST sample can capture enough of the most massive, rare clusters to quantify the spatial extent of these galaxies at low surface brightnesses, where the bound stellar structures blend with the intracluster light of their hosts. Another research area the LSST data can help address regards the central densities of local ellipticals that have seemingly decreased compared with field ellipticals at higher redshifts. The transformation of these dense, early ellipticals to the spheroids in the present day may involve galaxy mergers and environmental effects, two astrophysical processes that LSST can characterize through unparalleled statistics and environmental probes. By measuring the surface brightness profiles of billions of ellipticals LSST can determine whether any such dense early ellipticals survive to the present day, whatever their rarity.

Beyond the statistical advances enabled by LSST and the wide variation in environments probed by a survey of half the sky, the image quality of LSST will permit studies of galaxy structures in the very low surface brightness regime. Observational measures of the outer most regions of thin disks can constrain how such disks "end", how dynamical effects might truncate disks, and whether some disks smoothly transition into stellar halos. LSST will provide such measures and help quantify the relative importance the physical effects that influence the low surface brightness regions in disks. Other galaxies have low surface brightnesses throughout their stellar structures, and the image quality and sensitivity of LSST will enable the most complete census of low surface brightness galaxies to date. LSST will provide the best available constraints on the extremes of disk surface brightness, which relates to the extremes of star formation in low surface density environments.

The ability of LSST to probe low surface brightnesses also allows for characterization of stellar halos that surround nearby galaxies. Structures in stellar halos, from streams to density inhomogeneities, originate from the hierarchical formation process and their morphology provides clues to the formation history on a galaxy-by-galaxy basis (**bullock2005a**; **johnston2008a**). Observations with small telescopes (**martinez-delgado2008a**; **abraham2014a**) have already demonstrated that stellar halo structures display interesting variety (**van_dokkum2014a**). LSST, with its unrivaled entendue, can help build a statistical sample of stellar halos and cross-reference their morphologies with the observed properties of their central galaxies. Such studies may determine whether the formation histories reflected in the structures of halos also influence galaxy colors or morphological type. The examination of stellar halos around external galaxies may also result in the identification of small mass satellites whose sizes, luminosities, and abundances can constrain models of the galaxy formation process on the extreme low-mass end of the mass function.

2.3.3 Probing the Extremes of Galaxy Formation

The deep, multiband imaging LSST provides over an enormous area will enable the search for galaxies that form in the rarest environments, under the most unusual conditions, and at very early times. By probing the extremes of galaxy formation, the LSST data will push our theoretical understanding of the structure formation process.

The rarest, most massive early galaxies may form in conjunction with the supermassive black holes that power distant quasars. LSST can use the same types of color-color selections to identify extremely luminosity galaxies out to redshift $z \sim 6$, and monitor whether the stellar mass build-up in these galaxies tracks the accretion history of the most massive supermassive black holes. If stellar mass builds proportionally to black hole mass in quasars, then very rare luminous star forming galaxies at early times may immediately proceed the formation of bright quasars. LSST has all the requisite survey properties (area, mutliband imaging, and depth) to investigate this long-standing problem.

The creation of LSST Deep Drilling fields will enable a measurement of the very bright end of the high-redshift galaxy luminosity function. Independent determinations of the distribution of galaxy luminosities at $z \sim 6$ show substantial variations at the bright end. The origin of the discrepancies between various groups remains unclear, but the substantial cosmic variance expected for the limited volumes probed and the intrinsic rarity of the bright objects may conspire to introduce large potential differences between the abundance of massive galaxies in different areas of the sky. Reducing this uncertainty requires deep imaging over a wide area, and the LSST Deep Drilling fields satisfy this need by achieving sensitivities beyond the rest of the survey.

Lastly, the spatial rarity of extreme objects discovered in the wide LSST area may reflect an intrinsically small volumetric density of objects or the short duration of an event that gives rise to the observed properties of the rare objects. Mergers represent a critical class of short-lived epochs in the formation histories of individual galaxies. Current determinations of the evolving numbers of close galaxy pairs or morphological indicators of mergers provide varying estimates for the redshift dependence of the galaxy merger rate (**conselice2003a**; **kartaltepe2007a**; **lotz2008a**; **lin2008a**; **robotham2014a**). The identification of merging galaxy pairs as a function of separation, merger mass ratio, and environment in the LSST data will enable a full accounting of how galaxy mergers influence the observed properties of galaxies as a function of cosmic time.

2.3.4 Science Book

The contents of the Galaxies Chapter 9 of the Science Book (**LSSTSciBook**).

1. Measurements, Detection, Photometry, Morphology
2. Demographics of Galaxy Populations
   - Passively evolving galaxies
   - High-redshift star forming galaxies
   - Dwarf galaxies
   - Mergers and interactions
3. Distribution Functions and Scaling Relations
   - Luminosity and size evolution
   - Relations between observables
   - Quantifying the Biases and Uncertainties
4. Galaxies in their Dark-Matter Context
   - Measuring Galaxy Environments with LSST
   - The Galaxy-Halo Connection

    - Clusters and Cluster Galaxy Evolution
    - Probing Galaxy Evolution with Clustering Measurements
    - Measuring Angular Correlations with LSST, Cross-correlations
5. Galaxies at Extremely Low Surface Brightness
    - Spiral Galaxies with LSB Disks
    - Dwarf Galaxies
    - Tidal Tails and Streams
    - Intracluster Light
6. Wide Area, Multiband Searches for High-Redshift Galaxies
7. Deep Drilling Fields
8. Galaxy Mergers and Merger Rates
9. Special Populations of Galaxies
10. Public Involvement

2.4 Science Background: Astroinformatics

2.5 Science Background: Large Scale Structure

2.6 Science Background: Strong Lensing

2.7 Science Background: Weak Lensing

3. The Roadmap

4. Task Lists by Science Area

4.1 Active Galactic Nuclei Task Lists

AGN are phenomena that enable us to understand the growth of BHs, understand aspects of galaxy evolution, probe the high redshift universe and study other physical activity, including accretion physics, jets, magnetic fields, etc. There are distinct aspects of the study of AGN that can best be explored by considering AGN as an evolutionary stage of galaxies rather than a distinct type of source. The tasks listed here explore aspects of AGN study that are particularly important AGN as a stage in galaxy evolution.

T-1. AGN feedback in clusters

Motivation: Brightest Cluster/Group Galaxies (hereafter BCGs) are the most massive galaxies in the local Universe residing at/near the centres of galaxy clusters/groups. They will therefore contain the largest supermassive black holes. These black holes can influence their host BCG, the cluster gas and other cluster members via the mechanical energy produced by their 100s kpc scale jets (AGN feedback).

For low redshift galaxy clusters it is possible to perform detailed studies of the star, gas and AGN jets to analyse the details of AGN feedback. LSST will provide a large sample of moderate to high redshift clusters in which we can measure AGN feedback statistically. By combining X-ray, radio and optical observations we can assess the average influence of the BCG's AGN on the hot Intra-cluster medium (ICM) for different sub-populations [e.g. Stott et al. 2012].

Activities: By assembling a multi-wavelength dataset (optical, X-ray, Radio) we can obtain the BCG mass, cluster mass and ICM temperature, and the mechanical power injected into the ICM. We can use this to study the interplay between the BCG, its black hole and the cluster gas, to assess the balance of energies involved and for direct comparison with theoretical models of AGN feedback. This has been done with a few hundred clusters at $z<0.3$ using

SDSS but we may well be able to reach z=1 and therefore look for an evolution in their interplay and therefore AGN feedback. There are also implications for cosmology too as this will help with the selection of clusters for which the X-ray properties better represent the mass of the cluster rather than the complex interplay of baryonic physics.

Deliverables: Deliverables over the next several years from the activities described above include the following:

(a) Investigate the number of BCGs and the mass range of their clusters with redshift that LSST is likely to be able to observe.
(b) Assess radio and X-ray data available for AGN Feedback studies (XCS, eROSITA, SKA-pathfinders, SUMSS etc).
(c) Assess the theoretical predictions expected for the above (e.g. cosmological simulations such as EAGLE or more detailed single cluster studies).

T-1. AGN Selection from LSST Data

Motivation: Active Galactic Nuclei are selected using a variety of different methods. At optical and infrared wavelengths, photometric selection of AGN candidates is driven by their distinctive colors at particular redshifts. X-ray and radio observations can also be efficient selectors of candidates for additional follow-up. With spectral data, AGN can be selected using the ratios of their emission lines. LSST will also open up, in a more practical way, the identification of AGN based on their variability. Each of these samples probes aspects of the AGN phenomena and a better understanding of the AGN role in galaxy evolution requires that we understand how and why each of these selection methods includes or excludes particular sources. Furthermore, currently each of these methods for identifying AGN candidates requires spectral follow-up to cull these samples to positively identify the most reliably clean AGN sample.

Activities: For us to use LSST as a single way to identify the diversity of AGN, we must develop selection criteria that take advantage of the source parameters available with just LSST imaging, that is, color, morphology and variability. Already there are a number of AGN surveys with input from multiple wavelength observation and spectra. Precursor work needs to be done using these surveys to determine if AGN not easily identified using optical color selection can be selected using the additional parameters of morphology, variability and/or the additional filter that LSST provides.

Deliverables: Deliverables over the next several years from the activities described above include the following:

(a) Cross-matched catalog of known AGN selected and verified using different methods
(b) Development of morphology parameters beyond just star/galaxy separation and an understanding of the morphology parameters to be provided by LSST level 2 products.
(c) Development of color selection criteria that takes into account the morphology of the source
(d) Understanding of how AGN variability looks given the nominal LSST cadence
(e) Development of algorithms for color selection that take into account the variability of an AGN source

T-1. AGN Host Galaxy Properties from LSST Data

Motivation: We are requesting that basic morphological parameters (e.g., CAS, G-M20, etc.) be measured in the pipeline and made available as products to help in the identification of merging galaxies in LSST data. The issue here is how well this can be done when the host galaxies contain AGN that are likely identified via their variability. In other words, how well can we determine the host morphology of galaxies with variable AGN? This would be interesting for models of AGN fueling during mergers.

Activities: Simulations of the accuracy by which the pipeline (deblender) can measure the defined morphology parameters in host galaxies as a function of AGN brightness and wavelength. We could then "vary" the central source by expected levels in certain filters to see the effect on the morphological params. To constrain this it would be helpful to add in central sources with reasonable SEDs across the LSST bands, and a limited set of frequencies/amplitudes (based on real data - perhaps Pan-STARRS?).

Deliverables: Deliverables over the next several years from the activities described above include the following:

(a) Plots of the accuracy of the measured basic morphology parameters as a function of AGN brightness and wavelength.

(b) Effect of AGN brightness on classification diagrams.

T-1. AGN Variability Selection in LSST Data

Motivation: Most AGN exhibit broad-band aperiodic, stochastic variability across the entire EM spectrum on timescales ranging from minutes to years. Continuum variability arises in the accretion disk of the AGN, making it a powerful probe of accretion physics. The main LSST WFD survey will obtain $\sim 10^8$ AGN light curves (i.e. flux as a function of time) with ~ 1000 observations (~ 200 per filter band) over 10 years. The deep drilling fields will give us AGN lightcurves with much denser sampling for a small subset of the objects in the WFD survey. The science content of the lightcurves will critically depend on the exact sampling strategy used to obtain the light curves. For example, the observational uncertainty in determining the color variability of AGN will critically depend on the interval between observations in individual filter bands. It is of crucial importance to determine guidelines for an optimal survey strategy (from an AGN variability perspective) and determine what biases and uncertainties are introduced into AGN variability science as a result of the chosen survey strategy.

Activities: Study existing AGN variability datasets (SDSS Stripe 82, OGLE, PanSTARRS, CRTS, PTF + iPTF, Kepler, & K2) to constrain a comprehensive set of AGN variability models. Generate & study simulations using parameters selected from these models with the observationally determined constraints to determine goodness of simulations for carrying out various types of AGN variability science - PSD models, QPO searches, binary AGN models, etc.

Deliverables: Deliverables over the next several years from the activities described above include the following:

(a) Observational constraints on AGN variability models.

(b) MAF metrics quantifying the goodness of different survey strategies for AGN variability science.

T-1. AGN Photometric Redshifts from LSST Data

Motivation: Given the large number of AGN that will be observed with LSST, many of these will not be followed up with spectral observations. However, understanding the large scale structure of the universe, requires a 3-D understanding of the distribution of these galaxies in the universe. Photometric redshifts can provide relatively accurate redshifts for large numbers of galaxies. However, it is harder to obtain accurate photometric redshifts for galaxies that contain AGN compared to those that do not. We must understand how to get accurate photometric redshifts of galaxies with AGN.

Activities: An initial activity for this need to include comprehensive review of the state of the art in obtaining photo-z's for AGN host galaxy populations and how those compare to non-AGN galaxies. A comparison of model and/or observed AGN host SEDs with a matched set of non-host galaxies at a variety of redshifts will be used to determine color selection criteria for identifying AGN hosts. Explore whether variability can be used to break degeneracies.

Deliverables: Deliverables over the next several years from the activities described above include the following:

(a) Plots that show AGN host color selection criteria and where that color selection might become ambiguous (be degenerate) for non-host galaxies with different parameters.
(b) Plots that show if other parameters might break degeneracies.

T-1. AGN Merger Signature from LSST Data

Motivation: Understanding the role AGN play in galaxy evolution requires identifying the phenomenon at all stages and in all types of galaxies. AGN host galaxies are often found to be disturbed suggesting that the galaxy merger process is an important trigger of AGN activity. While the 'trainwrecks' may be easier to find, galaxies in other merger stages can be difficult to identify and those experiencing 'pre-merger' harassment may be particularly hard to recognize. Preliminary work needs to be done to understand how to identify mergers from the LSST data products and whether galaxy deblending and segmentation methods and procedures are adequate or mask galaxy mergers.

Activities: Create or Identify simulated and real images that contain known galaxy mergers, these images should contain mergers with and without AGN. Run LSST detection and identification software on these images. Identify metrics that describe/quantify the accurate detection of galaxy mergers (with and without AGN).

Deliverables: Deliverables over the next several years from the activities described above include the following:

(a) Give feedback to LSST software teams about metrics and detection of galaxy mergers
(b) Give feedback on structure or galaxy type that do and do not work well with current versions of LSST software

4.2 Photometric Redshifts

Summary

T-1. Impact of Filter Variations on Galaxy photo-z Precision

Motivation: For accurate photometric redshifts, well calibrated photometry is essential. Variations in the telescope system, particularly the broad-band ugrizy filters, will need to be very well understood if we are to meet the stringent LSST calibration goals. Photometry will be impacted by multiple factors that may vary as a function of position and/or time. The position of the galaxy in the focal plane will change the effective throughput both due to the angle of the light passing through the filter, and potential variations in the filter transmission itself due to coating irregularities across the physical filter. The spatially correlated nature of these effects can induce scale-dependent systematics that could be particularly insidious for measurements of local environment and clustering. The nominal plan from LSST

Data Management is to correct for variations across the focal plane. Such corrections will be SED dependent, and may leave residuals, particularly for specific populations with unusual SEDs. Tests of the amplitude of these residuals, and the impact on photo-z as a whole, and for particular object classes, is an important consideration. Beyond this, if the variations turn out to be very well calibrated, they could potentially be used to further improve, rather than degrade, photo-z performance. The variations in filter response can offer up additional a small amount of extra information on the object SED, given the slight variation in effective filter wavelength, particularly for objects with strong narrow features, i.e. emission lines. Tests of how much information is gained can inform whether or not the extra computational effort used in computing photo-z's from many slightly different filters as opposed to measurements corrected to the six fiducial filters of the survey.

Activities: XXX

Deliverables: Deliverables over the next several years from the activities described above include the following:

(a) AAA
(b) BBB

T-1. Photometric Reshifts in the LSST Deep Drilling Fields

Motivation: The LSST Deep Drilling Fields present different challenges than the main survey, including more confusion between sources, and the ability to use the best subsets of the images due to their being many repeat observations. These properties allow investigations of galaxies of brightness close to the noise in the main survey at higher signal to noise.

Activities: Assessing robustness of photo-zs with spectroscopic surveys will be difficult at the faintest fluxes, relationship to clustering redshifts important.

Deliverables: Deliverables over the next several years from the activities described above include the following:

(a) AAA
(b) BBB

T-1. Multivariate Physical Properties of Galaxies from Photometric Redshifts

Motivation: The knowledge of the derived physical properties underlies much of the work involving galaxies and their evolution. Derived physical properties include, among others: star formation rate (SFR), stellar mass ($M_{\star}$), specific SFR (sSFR), dust attenuation, and stellar metallicity. When it comes to scientific analysis, in recent years the derived physical properties have largely supplanted fluxes and luminosities in the UV, optical and near-IR bands. This is because derived properties require no redshift (K) corrections, are dust-corrected, and are therefore easier to relate across surveys and studies and to compare with the models. Stellar mass has emerged as a parameter of choice for selecting galaxy types and making apple-to-apple comparisons of galaxies at different redshifts. The sSFR (current SFR normalized by stellar mass) provides a rough estimate of galaxy's SF history. Dust attenuation and stellar metallicity are also indicative of various processes important for understanding galaxy evolution.

Activities: Deriving physical properties, usually accomplished by spectral energy distribution (SED) fitting, is an involved process and the results depend on the number of factors, including: underlying population models, assumed dust attenuation law, assumed star formation histories, choice of model priors, choice of IMF, emission line corrections, choice of input fluxes, type of flux measurements, treatment of flux errors, SED fitting methodology, interpretation of the resulting probability distribution functions (PDF) (e.g., Salim et al. 2016). In the case of LSST, the additional challenge is that the redshifts are for the most part photometric, and carry a PDF (a measure of uncertainty) of their own. In principle, the redshifts could be determined as part of the SED fitting (and vice versa, physical parameters can be obtained from some photo-z codes), but it is not clear whether this joint approach is the best. Alternatives are to use empirical training sets to obtain the photo-z (some "best" estimate or a PDF) and then feed it into the SED fitting code.

Activities will consist of testing whether the determination of physical parameters and photo-z should be performed jointly or not, based on training sets with spectroscopic redshifts, at a range of redshifts. Furthermore, testing should be performed on mock galaxies to understand which choices of methods and assumptions (specifically related to LSST data) produces the best results in the sense of retrieving the "known" properties.

Deliverables: Deliverables over the next several years from the activities described above include the following:

(a) Pre-LSST: A set of guidelines as to optimal practices regarding the derivation of both the photo-z and properties, together with the software to be used.

(b) With LSST data: the production of catalogs of properties to be used by the collaboration.

T-1. Identifying Spectroscopic Redshift Training Sets for LSST

Motivation: Require deep spectroscopic redshift data in order to help train algorithms, improve algorithms with clustering etc, and also provide a basis for determining accuracy of photo-z algorithms.

Activities: Collate existing spectroscopic redshift data over both DDF and wider fields, along with selection biases for each spectroscopic data set. Assess robustness of existing data, determine colour space where existing surveys lack statistics. Apply for additional spectroscopy to fill in parameter space not already covered by existing surveys.

Deliverables: Deliverables over the next several years from the activities described above include the following:

(a) AAA

(b) BBB

T-1. Develop Techniques to Identify Specific Sub-Populations of Galaxies

Motivation: Studying properties related with the star formation activity of galaxies, such as color and specific star formation rate (sSFR), as a function of mass, environment and redshift is relevant for understanding the different physical processes in galaxy formation and evolution. The aim is to develop techniques in order to identify specific sub-populations with the aforementioned properties (e.g. blue/star-forming and red/quenched galaxies) based on photometric data. Another interesting sub-population is galaxies which contain an active nucleus. The identification of AGN candidates will also be explored.

This task is potentially cross-cutting with the theory/mock catalogs, machine learning, clusters, lss, AGN, and DESC working groups and collaborations.

Activities: We can use simulations and mock catalogs to obtain prior estimates of the calibrations used to identify specific galaxy sub-populations. These calibrations will depend on mass and redshift (z). One technique to explore is fitting two Gaussians to the corresponding color and sSFR distributions in different mass and redshift bins to identify populations of red and blue galaxies. It is important that the mass definition assumed in the mocks be comparable to that estimated for observations. Note that the stellar mass would be used as the alpha parameter in the joint probability distribution functions, p(z,alpha).

Furthermore, we will make efforts to identify AGNs to obtain a sample of AGN candidates and, also, isolate them from "normal" galaxy samples without AGNs. The information of color and star formation described above can be used for this aim.

The techniques can be probed as a function of environment, which can be defined using different approaches at both small and large scales (e.g. number of neighbor galaxies, location in large-scale structures such as filaments, voids, knots, or Voronoi tessellation techniques). This would enable the characterization of galaxy sub-populations according to the environment. The resulting galaxy sub-populations can be used as training sets to be implemented on machine learning models.

Deliverables: Deliverables over the next several years from the activities described above include the following:

(a) Obtaining mass from mock catalogs compatible with the mass used in p(z,mass).
(b) Developing techniques that depend on this mass and redshift using mock catalogs for selecting samples with red/blue colors.
(c) Developing multiple techniques that depend on this mass and redshift using mock catalogs for selecting star-forming/quenched samples.
(d) Developing techniques that may depend on star formation, color and redshift for selecting AGN samples.
(e) Defining several environment estimators in simulated datasets.
(f) Probing techniques in b), c) and d) as a function of the environments defined in e).
(g) Obtaining training sets to be implemented on machine learning models.

T-1. Simulations with Realistic Galaxy Colors and Physical Properties

Motivation: As representative samples of spectroscopic redshifts will be very difficult to compile for LSST, simulations will play a key role in calibrating estimates of physical properties such as galaxy stellar mass, star formation rate, and other properties. This is particularly problematic for photometric surveys, where photometric redshift and physical property estimates must be calculated jointly. In addition, we must include prominent effects that will influence the expected photometric performance, for example the presence of an active galactic nucleus can significantly impact the color of a galaxy and the inferred values for the physical parameters, so models of AGN components of varying strength must be included in the simulations. Many current generation simulations cannot or do not simultaneously match observed color distributions and physical property characteristics for the galaxy population at high redshift. As photo-z algorithms are highly dependent on accurate photometry, realistic color distributions are required to test the bivariate redshift-physical property estimates. Working with the galaxy simulations and high redshift galaxy working groups to develop new simulations with more accurate high redshift colors is a priority. These photo-z needs are not unique, and the improved simulations will benefit the wider Collaboration as a whole.

Activities: The main activity for this task is to bring together the knowledge gained from observational studies of high redshift galaxies to act as input for improved simulation metrics. This will require expertise from the photo-z group, the high redshift galaxies group, the AGN group, and the simulations group. In order to test whether mock high-z populations agree with the real Universe, we must have some real data to compare against, even if it is a luminous subsample or only complete in certain redshift intervals. Once such comparison datasets are established, metrics can be developed to determine which simulations and simulation parameters most accurately reproduce the observed galaxy distributions. Assuming that the simulations are valid beyond the test intervals, we can then test bivariate photo-z/physical process determinations to develop improved algorithms.

Deliverables: Deliverables over the next several years from the activities described above include the following:

(a) Determination of a list of which physical parameters are important for galaxy science.
(b) Compiling observable datasets that can be used as comparators for simulated datasets.
(c) Developing a set of metrics to compare simulations to the observational data.
(d) Use the metrics in deliverable B to create updated simulations with more realistic parameter distributions.
(e) Development of improved joint estimators for redshift and physical properties (M*, SFR, etc. . .).

T-1. Using Galaxy Size and Surface Brightness distributions as Photo-z Priors

Motivation: Photometric redshift algorithms traditionally use galaxy fluxes and/or colors alone to estimate redshifts. However, morphological information in the form of the galaxy's size/shape/surface brightness (SB) profile adds additional information that can aid in constraining both the redshift and type of the galaxy, breaking potential degeneracies that using colors alone would miss. Adding type information beyond just the rest frame SED may help to constrain bivariate galaxy properties that correlate with morphological type as well. If sufficient training samples are available, a Bayesian prior on colors and SB profile, p(z|C,SB), can be constructed that should lead to improved photometric redshifts.

Activities: The primary activity in this task is to develop an algorithm to compute a parameterized SB profile fit (e.g. Sersic index, though other measures may be appropriate) for a large number of galaxies. The algorithm must be fast enough to compute SB profiles for large numbers of galaxies. Simulated datasets may be necessary to calibrate this code in the limits of galaxy sizes approaching the size of the PSF, and in the limit of low signal-to-noise ratios. With SB measurements in hand, the computation of a Bayesian prior on redshift given galaxy photometry and SB. This can be done with either simulated datasets, or real observations with spectroscopic redshifts. Tests will then show the performance of such a prior relative to using galaxy photometry alone.

Deliverables: Deliverables over the next several years from the activities described above include the following:

(a) A fast, scalable algorithm for measuring the surface brightness profile of galaxies.
(b) A cross matched catalog with objects at known redshifts and measured surface brightness profiles.
(c) A Bayesian prior $p(z|C,SB)$ that can be used to improve photo-z measurements.