Final Technical Report on Award DE-SC0011635

Calibration and Commissioning of the LSST

Andrew Connolly
University of Washington
August 1st 2022

Organization: Office of High Energy Physics
Project Period: 04/01/2017 - 03/31/2020

Abstract

Understanding the nature of dark energy and dark matter remains one of the fundamental questions in physics today; impacting our understanding of particle physics, cosmology, and possibly theories of gravity. Given the scale and complexity of the next generation of cosmology experiments (e.g., the Rubin Observatory, the Euclid satellite mission, and the Roman space telescope) we are entering an era where statistical noise no longer determines the accuracy to which we can measure cosmological parameters. Our ability to control and correct for systematics will ultimately determine the scientific impact of these experiments. This award addressed the challenge of how we determine what limits the accuracy of our cosmological measures, what techniques are appropriate for measuring and calibrating the properties of galaxies to best constrain cosmological models, how to develop statistical techniques that are insensitive to systematic errors, and how to optimize survey strategies in order to minimize systematics while maximizing the speed at which an experiment can achieve its science objectives. In this final technical report for award DE-SC0011635 we describe a set of open-source frameworks that simulate the characteristics and properties of current and planned cosmology surveys and the application of these frameworks to the development of new methodologies for estimating the properties and distances to galaxies that are robust to noisy and incomplete data.

Introduction

This final technical report for award DE-SC0011635 describes the development and application of tools and software to simulate and characterize the expected performance of DOE-funded dark energy surveys such as the Legacy Survey of Space and Time (LSST) that will be undertaken by the Rubin Observatory. This work includes the creation of a simulation framework that can generate catalogs of sources and images with the same properties as those that will be observed by the LSST, the evaluation of survey strategies to optimize the performance of the LSST including the creation of new optical filters to supplement those used by the LSST, the characterization of how well researchers can estimate the distances to galaxies using photometric redshifts, and novel techniques to combine images and to generate image templates from which transient and variable sources might be detected. This work contributed to 10 publications over the period of the award. In the following sections we describe the main results from this work.

A simulation framework for modelling the universe

The impact on cosmological measures of dark energy surveys such as the Rubin Observatory will be determined by the properties of the observations and our ability to measure, characterize, and remove systematics within the data. Development of a simulation framework capable of modelling these systematics was a primary deliverable from this award. This framework can generate astrophysical data sets with properties and uncertainties appropriate for a broad range of imaging surveys (e.g., the LSST, the Roman space telescope and the Euclid satellite), as well as model systematics within these surveys that impact measures of gravitational weak lensing, signatures of large-scale structure, and the detection of transient sources such as supernova (SN).

Under this award we enhanced the simulation framework to include observational effects such as Galactic extinction that modulates the density of sources as a function of position on the sky, extended the wavelength range of the spectral energy distribution of sources to infrared wavelengths to support ongoing work to understand the benefits of infrared observations from the Roman space telescope and the Euclid satellite, and integrated transient objects such as Type Ia supernovae used in estimating the expansion rate of the universe and variable sources (such as variable stars) that contaminate the classification of Type Ia supernova.

The simulation framework was used in the production of catalogs for the Dark Energy Science Collaboration's (DESC) Data Challenge 2 (DC2). This was used by DESC to evaluate the technical and scientific performance of the Rubin Observatory's science pipelines (software that will process and analyse Rubin data). DC2 was the first large-scale data challenge for DESC. It covered approximately 300 square degrees on the sky with five years of simulated LSST data. The data and early results are described in "The LSST DESC DC2 Simulated Sky Survey", LSST Dark Energy Science Collaboration (LSST DESC), Abolfathi, B., et al, 2021, ApJS, 253, 31.

A follow-on and related paper described the application of the simulation framework to determine how the cadence of Rubin observations (i.e., where on the sky the telescope observed, or how often, and in which filters) will impact science with LSST. This work was published in a series of papers including "Optimization of the Observing Cadence for the Rubin Observatory Legacy Survey of Space and Time: A Pioneering Process of Community-focused Experimental Design", Bianco, F. B., et al, 2022, ApJS, 258, 1

Models for transient and variable sources developed for the simulation frameworks were used in the PLAsTiCC time series classification challenge to develop new machine learning methodologies for identifying and classifying variable and transient sources. This classification challenge attracted 1380 applicants and led to several new approaches for classifying time series data to identify Type Ia supernova. The simulations used in PLAsTiCC were published in "Models and Simulations for the Photometric LSST Astronomical Time Series Classification Challenge (PLAsTiCC)", Kessler, R., et al, 2019, PASP, 131, 094501

Estimating the distances to galaxies using photometric redshifts

Broadband optical colors of galaxies can be used to estimate the distances to galaxies. These estimates are called photometric redshifts and many techniques for photometric redshift estimation rely on knowing the underlying spectral properties of the galaxies themselves. Under this award, two new techniques were developed to estimate the colors or spectra of distant galaxies.

A method based on Gaussian Processes was developed to estimate the spectral energy distribution (SED) of a galaxy from a set of photometric observations and a small training set of template SEDs. In this work we showed that Gaussian Process interpolation can create a continuous mapping of SEDs to photometric color space, allowing interpolation and extrapolation from a small training set of SEDs. This mapping provided SEDs that better matched galaxy colors than standard interpolation techniques such as nearest neighbor or linear interpolation. For the wavelength ranges where photometric filters and spectra overlap (i.e., the optical bands for the LSST), we have shown that we can reduce the mean square error between the estimated and true SED by over 65% when compared to the linear interpolation. This approach reduced the scatter (standard deviation) in the photometric redshift estimation of galaxies by over 24.8% and lowered the bias due to outliers by over 87.5% when compared to current SED-based photometric redshift techniques. This work was published in "Estimating Spectra from Photometry", Kalmbach, J. B. and Connolly, A. J., 2017, AJ, 154, 277

We extended the SED modelling to the question of learning the underlying SEDs of galaxies from an ensemble of ~100K galaxies with measured redshifts and colors (i.e., recovering the SED from broadband photometry). We have shown that we can reconstruct SEDs including emission and absorption lines at a significantly higher resolution than the broadband filters used to measure the photometry. When applied to photometric redshifts we found that our learned SEDs reduced the fraction of outliers in the derived photometric redshifts by up to 28%, decreased the bias by up to 91%, and reduced the scatter (standard deviation) in the redshift estimates by up to 25%, when compared to estimates using standard techniques. This work was published in "Learning Spectral Templates for Photometric Redshift Estimation from Broadband Photometry", Crenshaw, J. F. and Connolly, A. J., 2020, AJ, 160, 191

In a series of papers, we evaluated how well photometric redshifts can be derived using LSST data as a function of the length of the survey and whether utilizing external data sets such as those from the Euclid and Roman space telescopes would improve the redshift estimates. In these papers we showed that reducing the fraction of visits allotted by the LSST to u-band observations significantly deteriorates the photometric redshift accuracy, but that limiting the number of y-band visits had a smaller effect. We found that if the photometric errors for the LSST increased by 50% or suffered a systematic of +0.01 magnitudes, the accuracy of photometric redshifts derived by the LSST would not be sufficient to meet the requirements of the survey. Adding infrared data (e.g., from the Roman space telescope) could reduce the scatter (standard deviation) in the photometric redshifts by $\gtrsim$50% at high redshift and reduce the number of outliers (or catastrophic failures) at high redshift by a factor of almost 5. These results were published in "Photometric Redshifts with the LSST. II. The Impact of Near-infrared and Near-ultraviolet Photometry", Graham, M. L., et al, 2020, AJ, 159, 258 and "Photometric

Redshifts with the LSST: Evaluating Survey Observing Strategies", Graham, M. L., Connolly, A. J., et al, 2018, AJ, 155, 1

The technique developed to evaluate the photometric redshift performance was included in a paper to validate a range of different photometric redshift techniques, and was published in "Evaluation of probabilistic photometric redshift estimation approaches for The Rubin Observatory Legacy Survey of Space and Time (LSST)", Schmidt, S. J., et al, 2020, MNRAS, 499, 1587

Optimizing LSST operations

Building on this earlier work we developed a new approach for designing optimal filters for photometric redshift estimation. This technique (SIGGI) applied ideas from information theory to determine how well the SED and redshift of a galaxy could be constrained given a specific set of filters and observing time. The objective was to define the shape of a set of photometric passbands such that they can maximally discriminate between different galaxy spectral energy distributions. This was expressed in terms of the conditional entropy of the color distributions of the sources (i.e., we designed the filters to minimize the conditional entropy). For a realistic set of six filters covering optical wavelengths, we found that we can improve the standard deviation of the photometric redshift error by ~7% overall and up to 70% for some redshifts, when compared to the photometric redshift performance using standard LSST filters. This work was published in "Applying Information Theory to Design Optimal Filters for Photometric Redshifts", Kalmbach, J. B., VanderPlas, J. T., and Connolly, A. J., 2020, ApJ, 890, 74

In collaboration with Prof Budavari and Dr Lee at Johns Hopkins University, we developed a new approach for combining imaging data taken at a range of different airmasses or elevations. Refraction by the atmosphere causes the positions of sources to depend on the airmass through which they were observed. This shift is dependent on the underlying spectral energy distribution of the source as well as the filter or bandpass through which it was observed (the effect is referred to as differential chromatic refraction or DCR).

With surveys such as the LSST undertaking repeated observations of the same part of the sky over a range of different airmasses DCR should introduce a detectable and measurable astrometric signal. Through this award we developed a novel procedure that takes this astrometric signal and uses it to infer the underlying spectral energy distribution of a source (essentially creating a low resolution spectrum for each source pixel). We have demonstrated, using simulations, that we can generate partially deconvolved images at higher spectral resolution than the input images (i.e., increasing the effective u and g band spectral resolution by a factor of three), for surveys such as the LSST.

This work was published as "Sub-band Image Reconstruction Using Differential Chromatic Refraction", Lee, M. A., Budavári, T., Sullivan, I. S., and Connolly, A. J., 2019, AJ, 157, 182

Publications

"Optimization of the Observing Cadence for the Rubin Observatory Legacy Survey of Space and Time: A Pioneering Process of Community-focused Experimental Design", Bianco, F. B., Ivezić, Ž., Jones, R. L., Graham, M. L., Marshall, P., Saha, A., Strauss, M. A., Yoachim, P., Ribeiro, T., Anguita, T., and 37 colleagues, 2022, ApJS, 258, 1

"The LSST DESC DC2 Simulated Sky Survey", LSST Dark Energy Science Collaboration (LSST DESC), Abolfathi, B., Alonso, D., Armstrong, R., Aubourg, É., Awan, H., Babuji, Y. N., Bauer, F. E., Bean, R., Beckett, G., Biswas, and 70 colleagues, 2021, ApJS, 253, 31

"Evaluation of probabilistic photometric redshift estimation approaches for The Rubin Observatory Legacy Survey of Space and Time (LSST)", Schmidt, S. J., Malz, A. I., Soo, J. Y. H., Almosallam, I. A., Brescia, M., Cavuoti, S., Cohen-Tanugi, J., Connolly, A. J., DeRose, J., Freeman, P. E., and 17 colleagues, 2020, MNRAS, 499, 1587

"Learning Spectral Templates for Photometric Redshift Estimation from Broadband Photometry", Crenshaw, J. F. and Connolly, A. J., 2020, AJ, 160, 191

"Photometric Redshifts with the LSST. II. The Impact of Near-infrared and Near-ultraviolet Photometry", Graham, M. L., Connolly, A. J., Wang, W., Schmidt, S. J., Morrison, C. B., Ivezić, Ž., Fabbro, S., Côté, P., Daniel, S. F., Jones, R. L., and 3 colleagues, 2020, AJ, 159, 258

"Applying Information Theory to Design Optimal Filters for Photometric Redshifts", Kalmbach, J. B., VanderPlas, J. T., and Connolly, A. J., 2020, ApJ, 890, 74

"Models and Simulations for the Photometric LSST Astronomical Time Series Classification Challenge (PLAsTiCC)", Kessler, R., Narayan, G., Avelino, A., Bachelet, E., Biswas, R., Brown, P. J., Chernoff, D. F., Connolly, A. J., Dai, M., Daniel, S., and 21 colleagues, 2019, PASP, 131, 094501

"Sub-band Image Reconstruction Using Differential Chromatic Refraction", Lee, M. A., Budavári, T., Sullivan, I. S., and Connolly, A. J., 2019, AJ, 157, 182

"Photometric Redshifts with the LSST: Evaluating Survey Observing Strategies", Graham, M. L., Connolly, A. J., Ivezić, Ž., Schmidt, S. J., Jones, R. L., Jurić, M., Daniel, S. F., and Yoachim, P., 2018, AJ, 155, 1

"Estimating Spectra from Photometry", Kalmbach, J. B. and Connolly, A. J., 2017, AJ, 154, 277