UNIVERSITY OF CALIFORNIA,
IRVINE

Toward dark energy: DESI & LSST

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Physics & Astronomy

by

Bela Abolfathi

Dissertation Committee:
Professor David Kirkby, Chair
Professor Mu-Chun Chen
Professor Simona Murgia

2022

DEDICATION

To my grandfather, Mikail Abolfathi

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

After my first physics class in high school, I never wanted to take another physics class again. The fact that I ended up pursuing a PhD in the subject is, first and foremost, a testament to my physics professors at Reed College. Despite the many people who have helped and encouraged me, the genesis of my path to pursuing a PhD began with them. I am deeply indebted to Mary James and David Griffiths for giving me the confidence to pursue physics, and for Joel Franklin, Darrell Schroeder, John Essick, and Johnny Powell for their exceptional instruction, their inspiration and their immeasurable contribution to my academic success.

I am grateful to the people at the Kavli Institute for Particle Astrophysics and Cosmology, specifically Aaron Roodman, for his mentorship and guidance on a topic I had no prior familiarity with; Jim Chiang and Seth Digel, for always keeping their doors open and for their patience in answering my questions; Pat Burchat, for graciously offering me a place to stay in Palo Alto; and Eric Charles and Greg Madejski, for always being so kind to me and going out of their way to make me feel included at KIPAC.

DESC collaboration meetings were always something I looked forward to, but Eric Gawiser made them infinitely better, whether it was organizing a game of soccer, or getting a group together to wander the streets of Paris. I am so grateful for his empathy and support. I am thankful to Joe Zuntz for organizing the DESC Sprint Week in Edinburgh, which was where I was first introduced to work on photomeric redshifts; and Sam Schmidt, Will Hartley, and Markus Rau, for initiating me into the DESC Photometric Redshifts Working Group.

One of the most consequential experiences of my time in graduate school was participating in the LSSTC Data Science Fellowship Program. I am so thankful to Lucianne Walkowicz and Adam Miller for showing me the human side of science, and for giving me hope at a time when I didn't feel like there was a place for me in the astronomy community.

I was very lucky to have amazing mentors Simona Murgia and Mu-Chun Chen at UCI, who were a vital source of encouragement and support. There are always those people who, after seeing them, can completely turn your day around for the better. Gary Chanan is one of them, and I am incredibly grateful for our conversations about classical music, orchestra, and for keeping me in the loop with performances on campus. I'd also like to thank Aaron Barth for the opportunity to participate in the LAMP project, where I got to conduct both on-site and remote telescope observations for the first time, and for sharing his treasured collection of old American folk music with me.

I am extremely grateful to my advisor, David Kirkby, for his endless patience, caring and empathy. The high standards he set in his work, as well as the work-life balance he achieved are examples I continue to strive for.

I owe so much thanks to Melissa, Soffi, Emily L., Ben, Yasha, Zaina, Manu, Carolyn, Mandy, Minhzuy, Sahel, Arash, Tianxing, and Emily P.L., whose friendship and encouragement have

meant the world to me; to Annie and Toby – my fairy godmother – for their love and wisdom, and for giving me a sanctuary at a time when I desperately needed it; to Maryanne, my spiritual sister, for her unending love and empathy; and to Jamie, for giving me strength and perspective.

I'd like to thank the Department of Energy and the LSST Corporation for providing me with the financial support to conduct my research.

And lastly, I'd like to thank my family: Payman, Soosan, Dersu, Taras, Ameh Valo, Mama Parvin, Maman Forough, Baba Mahmoud, and Elsa.

VITA

Bela Abolfathi

EDUCATION

Doctor of Philosophy in Physics & Astronomy **2022**
University of California, Irvine *Irvine, CA*

Master of Science in Physics & Astronomy **2016**
University of California, Irvine *Irvine, CA*

Bachelor of Arts in Physics **2011**
Reed College *Portland, OR*

RESEARCH EXPERIENCE

Graduate Student Researcher **2014-2022**
University of California, Irvine *Irvine, California*

Graduate Student Researcher **2018**
SLAC National Accelerator Laboratory *Menlo Park, CA*

REU Student **Summer 2010**
Cornell University *Ithaca, New York*

TEACHING EXPERIENCE

Teaching Assistant **2014–2016**
University of California, Irvine *Irvine, California*

REFEREED JOURNAL PUBLICATIONS

Establishing the Foundations of BH Mass Estimators I: Searching for Observational Proxies to Determine Individual Scale Factors for Future Reverberation Mapping Campaigns **2022**
Astrophysical Journal

The LSST-DESC 3x2pt Tomography Optimization Challenge **2021**
Open Journal of Astrophysics

The Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: The Damped Lyman- systems Catalog **2022**
Astrophysical Journal, Supplement

The Lick AGN Monitoring Project 2016: Velocity-Resolved H Lags in Luminous Seyfert Galaxies **2021**
Astrophysical Journal

DESC DC2 Data Release Note **2021**

The LSST DESC DC2 Simulated Sky Survey **2021**
Astrophysical Journal, Supplement

Overview of the DESI Legacy Imaging Surveys **2019**
Astronomical Journal

The Fourteenth Data Release of the Sloan Digital Sky Survey: First Spectroscopic Data from the extended Baryon Oscillation Spectroscopic Survey and from the second phase of the Apache Point Observatory Galactic Evolution Experiment **2018**
Astrophysical Journal, Supplement

Sloan Digital Sky Survey IV: Mapping the Milky Way, Nearby Galaxies, and the Distant Universe **2017**
Astronomical Journal

Discovery and Follow-up Observations of the Young Type Ia Supernova 2016coj **2017**
Astrophysical Journal

SOFTWARE

pz_bayes `https://github.com/LSSTDESC/pz_bayes`
Photometric redshift estimation using a hierarchical Bayesian framework.

SOMVIZ `https://github.com/belaa/SOMVIZ`
Tools for generating self-organizing maps.

ABSTRACT OF THE DISSERTATION

Toward dark energy:
DESI & LSST

By

Bela Abolfathi

Doctor of Philosophy in Physics & Astronomy

University of California, Irvine, 2022

Professor David Kirkby, Chair

One of the most compelling problems in physics today is understanding the nature of dark energy, a mysterious component driving the current accelerated cosmic expansion. The Dark Energy Spectroscopic Instrument (DESI) and the Rubin Observatory Legacy Survey of Space and Time (LSST) are Stage-IV Department of Energy (DOE) projects aimed at better understanding the nature of dark energy and its influence on the evolution of the universe. While DESI is a spectroscopic survey, and LSST provides multi-band photometry, their observations are complementary and can be combined to improve measurements of cosmological parameters.

One area of synergy lies in estimating the redshifts of extragalactic sources. The overlap between the DESI and LSST footprints is approximately 4,000 square degrees. While DESI will have a lower density of galaxies per square degree, having spectra for these targets will help to improve constraints on measured redshift distributions. The first part of the thesis will focus on developing and testing a state-of-the-art photometric redshift estimation algorithm on simulated LSST data. The algorithm employs a hierarchical Bayesian framework to simultaneously incorporate photometric, spectroscopic, and clustering information to constrain redshift probability distributions of populations of galaxies, as well as provide redshift

estimates of their individual members. Once data from LSST arrives, this method can be tested and refined through training on real LSST targets whose counterparts lie within the DESI footprint. This will ultimately improve redshift estimates for other targets in LSST by providing spectroscopic prior information, and will be especially useful in the context of tomographic weak gravitational lensing, which derives a significant amount of uncertainty from imprecise redshift estimates.

Another crucial step in limiting weak lensing systematics involves understanding and mitigating image artifacts in the camera. This is important for identifying blended objects, as well as pinpointing biases in shear measurements. The third chapter of the thesis focuses on studying the systematics of the LSST instrument response by investigating anomalies in calibration sequences and developing testing software to analyze irregularities in bias frames. Making reliable, quantitative measurements that can be compared to requirements at the 1% level is necessary to avoid systematic biases in weak lensing shape measurements, which are often of the same order as the sensor distortions.

The second half of the thesis is devoted to developing software pipelines in preparation for the DESI survey. The fourth chapter discusses using a Gaussian mixture model (GMM) to characterize galaxy magnitudes and colors from DESI targeting data for the purpose of generating mock spectra. Results from the GMM are compared to density estimates for these features using extreme deconvolution, which simultaneously models the data and the noise to provide error-deconvolved distribution functions.

One of the final stages in the processing of mock spectra involves accounting for the noise contributions due to the atmosphere and the spectrograph response. The last chapter is devoted to reconfiguring a DESI software package that simulates this response to produce synthetic spectra for Lyα studies in the Extended Baryon Oscillation Survey (eBOSS). The original configuration is then used to validate the DESI sky model by comparing real sky brightnesses with simulated brightnesses generated under similar observing conditions.

Chapter 1

Introduction

1.1 A history of dark energy

Since its publication in 1915, Albert Einstein's theory of general relativity has been consistently proven to be the correct description of gravity on large scales (e.g., Dyson et al. 1920; Abbott et al. 2016; Kramer et al. 2021). The nature of a key feature in his theory, however, is still under debate. At the time he formulated his field equations, Einstein assumed that space was neither expanding nor contracting. To accommodate a static universe, Einstein introduced a cosmological constant, Λ, to his field equations to balance the attractive force of gravity induced by matter:

$$G^{\mu\nu} + \Lambda g^{\mu\nu} = -\frac{8\pi G}{c^4} T^{\mu\nu} \,. \tag{1.1}$$

Not only was this static formulation unstable, but motivation for a cosmological constant waned after Edwin Hubble's 1929 discovery that the universe was in fact expanding (see Fig. 1.1). The cosmological constant was generally abandoned until the 1990's, when Riess et al.

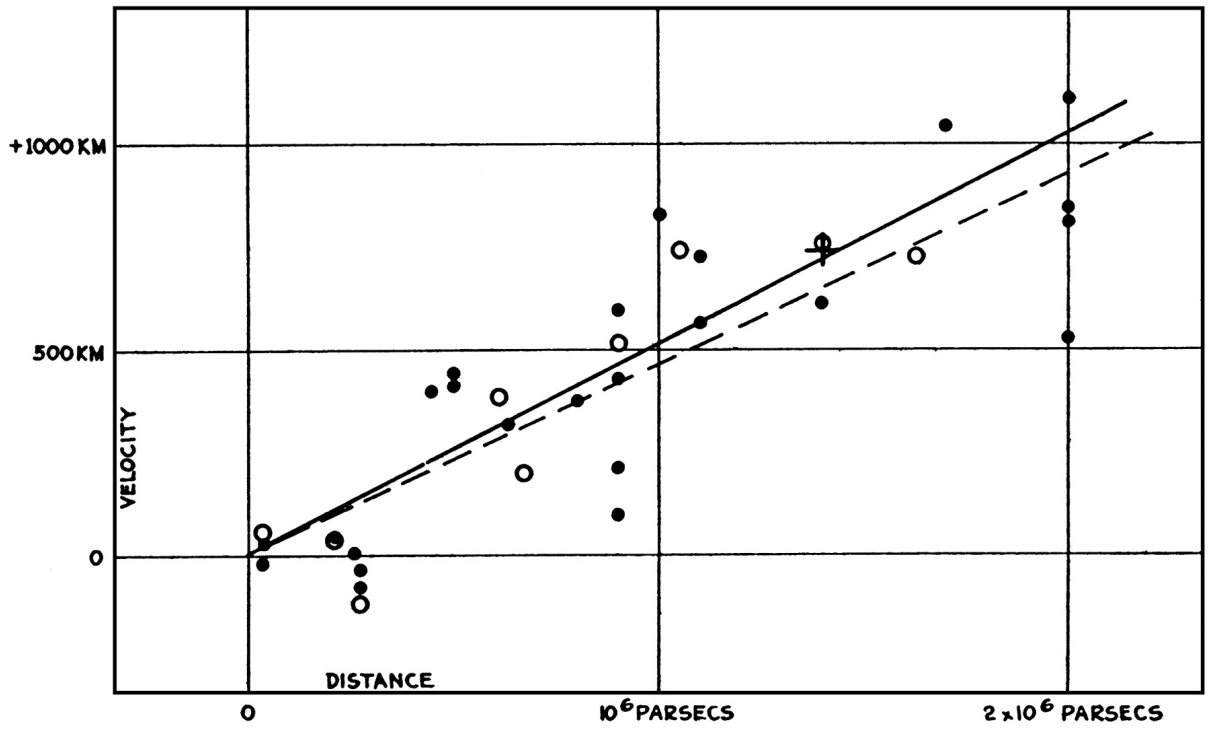


Figure 1.1: A figure from Hubble's original paper showing the radial velocities of extragalactic nebulae as a function of distance. The dark circles and solid line represent the solution to the velocity-distance relation using individual nebulae, while the circles and dashed line combine the nebulae into groups. The units of the Y-axis should be in km/s. Figure is from Hubble (1929).

(1998) and Perlmutter et al. (1999) independently observed that the measured luminosity distances of Type Ia supernovae were larger compared to models assuming a universe solely composed of matter and radiation (see Fig. 1.2). This discovery not only proved that cosmic expansion was accelerating, but it provided the first direct evidence of dark energy as a new component responsible for driving this acceleration. It also established a new benchmark model for cosmology, called ΛCDM, describing a universe born in a hot big bang, and composed of dark energy in the form of a cosmological constant, cold dark matter (CDM), and baryonic matter – governed by the laws of general relativity.

Although we still lack a compelling theoretical explanation for dark energy, there is a growing amount of observational evidence supporting the benchmark model. Measurements from a multitude of probes, including temperature anisotropies in the cosmic microwave background

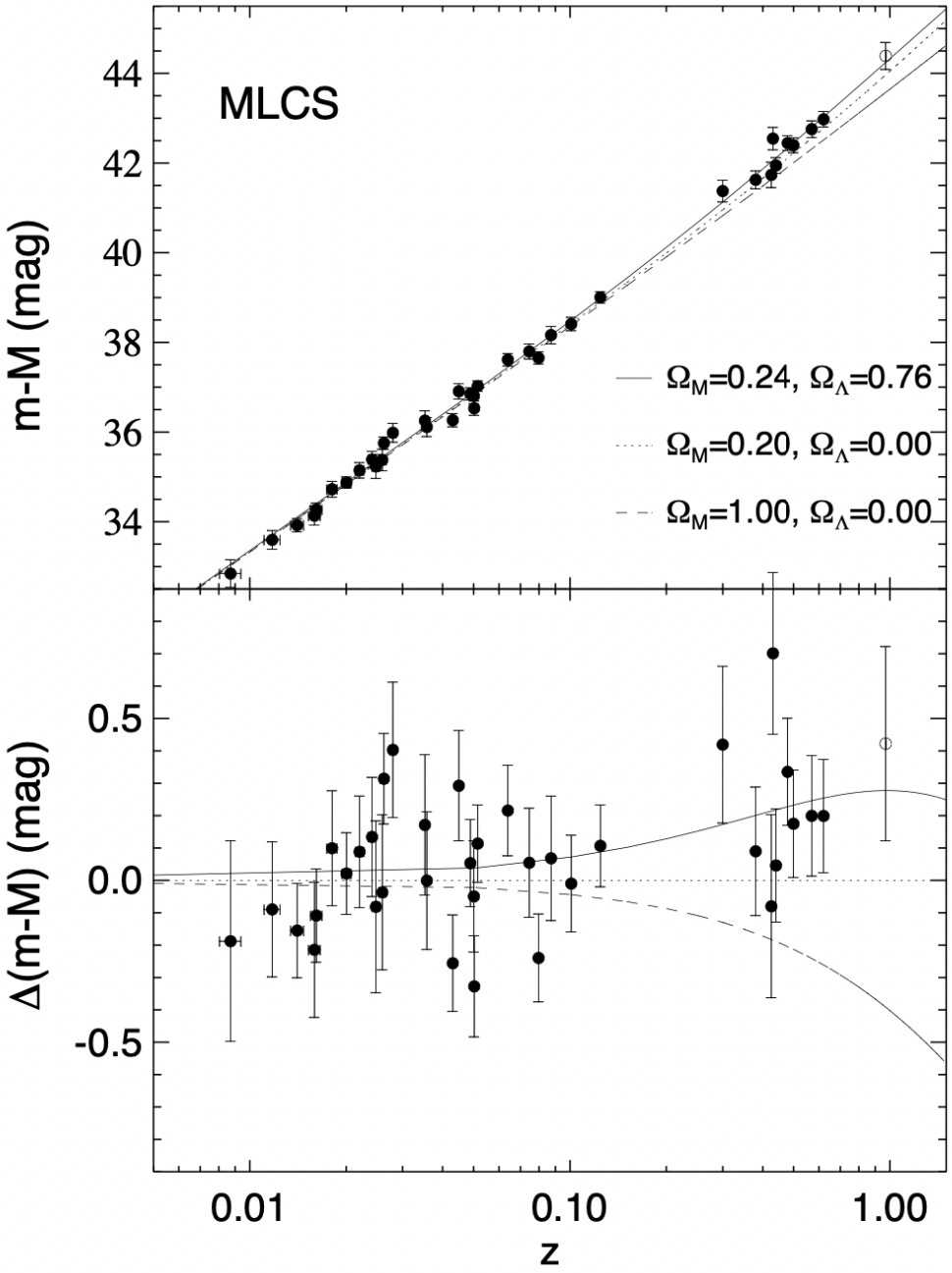


Figure 1.2: Hubble diagram showing the velocity-distance relation for Type Ia supernovae (SNe Ia). The top panel shows the observed magnitudes of SNe Ia samples as a function of redshift, along with three different cosmologies. Distance measurements were obtained by fitting SNe light curves using the Multi-Color Light Curve Shape method (MLCS). The bottom panel shows the residuals between the data and each model cosmology. The best fit to the data is a model with 24% matter and 76% dark energy today. Figure is from Riess et al. (1998).

(Hinshaw et al. 2013; Aghanim et al. 2020), the baryon acoustic scale in the clustering of galaxies and opacity fluctuations in the spectra of high-redshift quasars (Zhao et al. 2021; du Mas des Bourboux et al. 2020), and the deflection of light from background galaxies due to the gravitational pull of foreground mass concentrations (Asgari et al. 2021; Amon et al. 2022), are so far in excellent agreement with ΛCDM. The latest results from the Planck Collaboration[1] reveal the current rate of cosmic expansion to be 68km $s^{-1}Mpc^{-1}$, and a universe made up of $\Omega_\Lambda = 68.9\%$ dark energy, $\Omega_{DM} = 26.1\%$ dark matter, and $\Omega_b = 4.9\%$ baryonic matter (Aghanim et al. 2020). Next-generation dark energy surveys, such as the Dark Energy Spectroscopic Instrument (DESI[2]; DESI Collaboration (2016)), the Nancy Grace Roman Space Telescope[3] (Spergel et al. 2015), Euclid[4] (Laureijs et al. 2011), and the Vera C. Rubin Observatory Legacy Survey of Space and Time (LSST[5]; Ivezić et al. (2019)), will constrain cosmological parameters to sub-percent precision in order to better understand the dynamical behavior of dark energy and its influence on the formation of large-scale structure. These experiments will not only provide stringent tests of ΛCDM, but they will allow cosmologists to discriminate between alternative theories of dark energy such as quintessence, which posit dark energy as a scalar field ϕ with a slowly varying energy density, as well as theories that abandon dark energy altogether, instead suggesting that general relativity requires modification.

1.2 Measuring cosmic expansion

Insight into the nature and cause of accelerated expansion rests on accurate measurements of a number of cosmological parameters. In this section we will discuss some of the fundamental

[1] https://www.cosmos.esa.int/web/planck
[2] https://www.desi.lbl.gov
[3] https://roman.gsfc.nasa.gov
[4] sci.esa.int/euclid/
[5] http://www.lsst.org/lsst

parameters used to study cosmic dynamics and expansion. For a more in-depth treatment, the reader is encouraged to refer to Dodelson (2003), Ryden (2003), and Peacock (1999).

To discuss gravity's effect on distances in an expanding universe, we first need a way of quantifying expansion. We introduce a function called the scale factor $a(t)$ to relate physical separations to their separation due to expansion at time t since the Big Bang. The scale factor is a dimensionless, non-negative number with a present-day value defined to be $a(t_0) = 1$, where t_0 is today. It can be expressed in terms of the cosmological redshift, z, of electromagnetic radiation emitted at time t_{emit}:

$$a(t_{emit}) = (1 + z)^{-1} \,. \tag{1.2}$$

To quantify the rate of change of the scale factor, we introduce the Hubble rate, $H(t)$:

$$H(t) \equiv \frac{\dot{a}(t)}{a(t)} \,. \tag{1.3}$$

The value of the Hubble rate today is the Hubble constant, H_0, and is commonly parameterized in terms of the constant h:

$$H_0 \equiv h \, (100\text{km s}^{-1}\text{Mpc}^{-1}) \,. \tag{1.4}$$

While physical distances are modified by $a(t)$, the comoving distance, D_M, is referenced to today, $t = t_0$, and factors out this expansion. It is related to the Hubble rate, $H(t)$, via:

$$D_M(z) = \frac{c}{H_0} \int_0^z dz' \frac{H_0}{H(z')} \,. \tag{1.5}$$

Other distance measures (for a flat universe) include the luminosity $D_L(z)$ and angular diameter $D_A(z)$ distances:

$$D_L(z) = D_M(z)(1+z) \tag{1.6}$$

$$D_A(z) = D_M(z)/(1+z)\,. \tag{1.7}$$

The rate of cosmic expansion is inextricably tied to its composition. The relative densities of matter, radiation, and dark energy factor into this rate, and these densities scale differently as expansion speeds up or slows down. The Friedmann equation describes the evolution of the scale factor in terms of the density parameter for each component:

$$\frac{H^2(z)}{H_0^2} = \Omega_m(1+z)^3 + \Omega_r(1+z)^4 + \Omega_k(1+z)^2 + \Omega_\phi \frac{u_\phi(z)}{u_\phi(z=0)}\,, \tag{1.8}$$

where the density parameters for matter Ω_m (baryonic Ω_b and dark Ω_{DM}), radiation Ω_r, curvature Ω_k, and a generic form of dark energy Ω_ϕ, are defined as the ratio of the present-day energy density u_x to the critical density $\rho_{crit} \equiv \frac{3c^2H_0^2}{8\pi G}$:

$$\Omega_x \equiv \frac{u_x}{\rho_{crit}}\,. \tag{1.9}$$

If dark energy were associated with a cosmological constant Λ, the last term in Eq. 1.8 would not scale with expansion, such that $\Omega_\phi \frac{u_\phi(z)}{u_\phi(z=0)} \to \Omega_\Lambda$. In a flat universe, the sum of the densities of each component is equal to the critical density. A departure from flatness is described in terms of the curvature term Ω_k:

$$\Omega_k \equiv 1 - \Omega_m - \Omega_r - \Omega_\phi \,. \tag{1.10}$$

A universe with $\Omega_k = 0$ indicates a spatially flat universe, whereas $\Omega_k < 0$ and $\Omega_k > 0$ indicate closed and open universes, respectively.

The evolution of the dark energy density with the scale factor is expressed in terms of the equation of state parameter w:

$$\frac{u_\phi(z)}{u_\phi(z=0)} = (1+z)^{3(1+w)} \,, \tag{1.11}$$

where w is defined as the ratio of a component's pressure to its energy density $w \equiv pc^2/u$. For a cosmological constant, $w = -1$, while $w = 0$ and $w = +\frac{1}{3}$ for non-relativistic matter and radiation, respectively. The scaling of a dynamical dark energy field with expansion is slightly more complicated:

$$\frac{u_\phi(z)}{u_\phi(z=0)} = \exp\Big[\int_0^z [1+w(z')]\frac{dz'}{1+z'}\Big] \,. \tag{1.12}$$

The equation-of-state for dark energy is commonly parameterized as:

$$w(a) = w_0 + w_a(1+a) \,, \tag{1.13}$$

where w_a captures the linear evolution of $w(a)$ with scale factor a. In models that favor dark energy as a cosmological constant, $w_a = 0$. Constraints on the dark energy equation of state from a diverse range of probes are shown in Fig. 1.3.

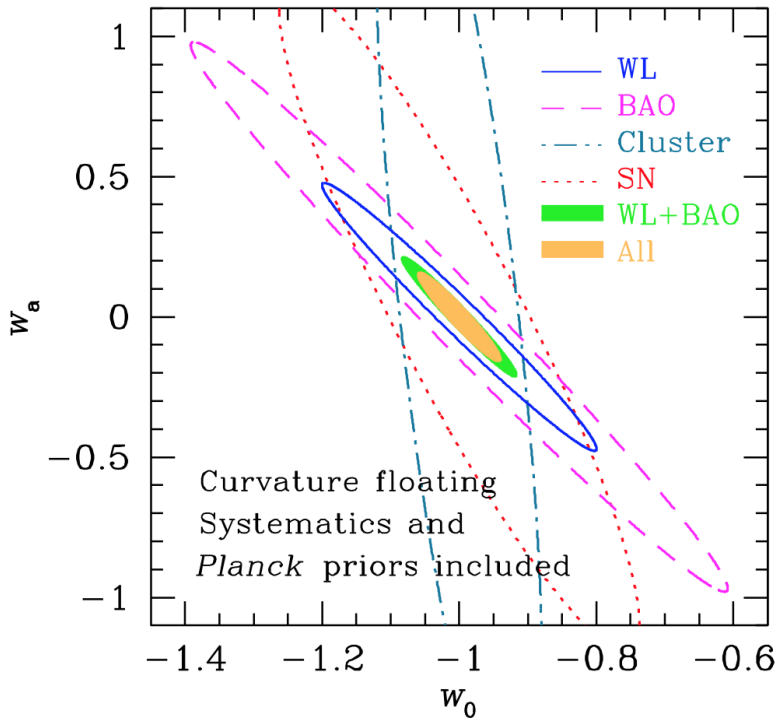


Figure 1.3: Constraints on the dark energy equation of state $w(a) = w_0 + w_a(1+a)$ from both individual and combined probes. Individual probes constrain degenerate combinations of cosmological parameters and are sensitive to different systematics. A combination of probes, therefore, allows for calibrating systematics so that degeneracies can be broken, resulting in tighter constraints on parameters. The ellipses show contours for weak lensing (WL), baryon acoustic oscillations (BAO), clusters, supernovae (SN), WL+BAO, and a combination of all probes. Figure is from Ivezić et al. (2019).

1.3 Using large-scale structure to study dark energy

1.3.1 Seeds of cosmic structure

A key assumption underpinning any cosmological study of the universe is the cosmological principle, which states that the universe is homogeneous and isotropic on scales larger than 100Mpc. At its core, it means that there is no special observer or location in the universe, so that any sample of the sky can be assumed to be obey the same physical laws.

There is good evidence that the universe is isotropic on large scales from measurements of the cosmic microwave background (CMB), where variations in the 2.75K background are $\delta T/T \lesssim \mathcal{O}(10^{-4})$ on angular scales between $1'$ to 180° (Kolb & Turner 1990). These temperature variations trace the distribution of matter at the time of decoupling, when matter was free of the radiation pressure induced by photons, and density perturbations began to grow via the Jeans instability. The density perturbations in the CMB vary as:

$$\frac{\delta\rho}{\rho} = \alpha \times (\frac{\delta T}{T}) \lesssim \mathcal{O}(10^{-3} - 10^{-2})\,, \tag{1.14}$$

where α is $\mathcal{O}(10 - 10^2)$. The history of the universe is the playing out between the forces of expansion and gravity, the latter of which is necessary for the formation of structure. While the relic photons constituting the CMB did not participate in the gravitational collapse of matter, they provide a fossil record of the primordial imhomogeneities at decoupling. These inhomogeneities are associated with small, Gaussian fluctuations in the matter density field that resulted in the galaxies, clusters, voids and filaments we see today.

1.3.2 The matter power spectrum

To discuss the matter density field of the universe, we introduce the density contrast:

$$\delta(\mathbf{x}, t) \equiv \frac{\delta\rho_m(\mathbf{x}, t) - \bar{\rho}_m(t)}{\bar{\rho}_m(t)}\,, \tag{1.15}$$

where $\bar{\rho}_m(t)$ is the mean matter density at time t, and $\mathbf{x}$ is the comoving spatial coordinate. Density fluctuations grow with time according to a given cosmological model, obeying the following differential equation:

[6] `https://virgo.dur.ac.uk`

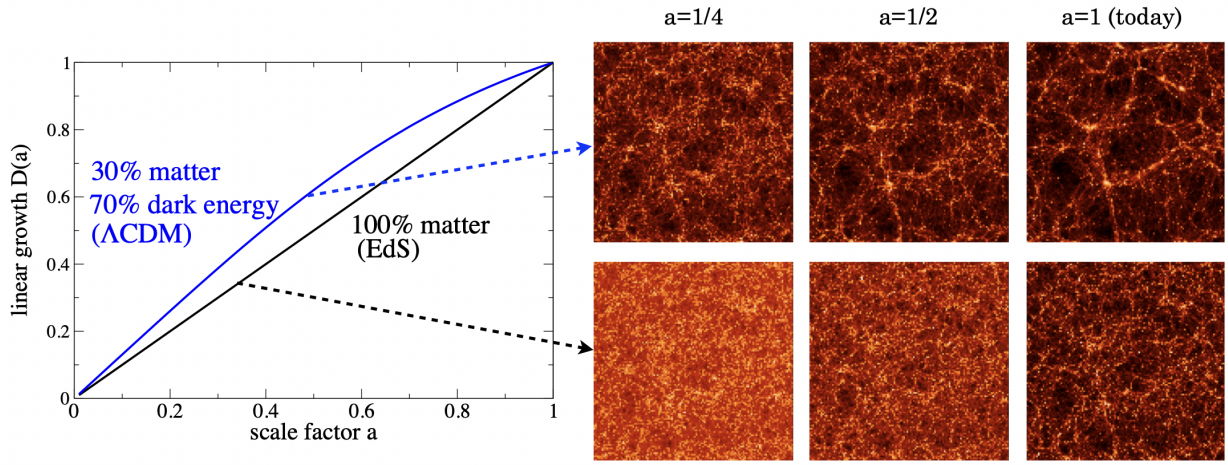


Figure 1.4: The growth of structure as a function of scale factor. The left panel shows the linear growth function $D(a)$ for ΛCDM (blue) and Einstein-de Sitter (EdS; black) universes, where the latter is a model for a flat, matter-only universe. In the ΛCDM model, there is less dark energy in the past to suppress structure growth, so the growth function is larger for values of the scale factor where the universe was matter-dominated. The right panel shows numerical N-body simulations of matter density fluctuations by the Virgo consortium[6] (Jenkins et al. 1998). Again, we see that dark energy-dominance at late times suppresses structure growth, resulting in less "lumpiness" on large scales. Figure is from Huterer et al. (2015).

$$\ddot{\delta}(\mathbf{x},t) + 2H\dot{\delta}(\mathbf{x},t) - 4\pi G\rho_m\delta(\mathbf{x},t) = 0\,, \tag{1.16}$$

where G is the gravitational constant, and $\dot{\delta} \to \frac{d\delta}{dt}$. In the linear regime (where $|\delta| \ll 1$ and at large spatial scales k), the amplitude of fluctuations is normalized with respect to the linear growth function $D(t)$:

$$\delta(\mathbf{x},t) = \delta(\mathbf{x},t_i) \times \frac{D(t)}{D(t_i)}\,, \tag{1.17}$$

where t_i is an arbitrary time, usually taken to be the present t_0. The evolution of the growth function as a function of scale factor $D(a)$ is shown in Fig. 1.4. $D(t)$ is indicative of the growth of primeval matter perturbations, whereby initially over-dense regions continue to

attract more matter over time, becoming increasingly more over-dense. The matter power spectrum $P(k)$ is used to characterize the clustering of matter on different scales $\mathbf{k}$:

$$\langle \tilde{\delta}(\mathbf{k}) \tilde{\delta}(\mathbf{k}') \rangle = (2\pi)^3 P(k) \delta_D^3(\mathbf{k} - \mathbf{k}') \,, \tag{1.18}$$

where δ_D^3 is the three-dimensional Dirac delta function, and the average is taken over $|\mathbf{k}| = k$. The Fourier transform of the density contrast $\tilde{\delta}(\mathbf{k})$, and its inverse $\delta(\mathbf{r})$, are given by:

$$\tilde{\delta}(\mathbf{k}) = \int d^3r \, e^{-\mathbf{k}\cdot\mathbf{r}} \delta(\mathbf{r}) \,, \qquad\qquad \delta(\mathbf{r}) = (2\pi)^{-3} \int d^3k \, e^{\mathbf{k}\cdot\mathbf{r}} \tilde{\delta}(\mathbf{r}) \,. \tag{1.19}$$

The power spectrum is expressed in terms of wavenumber $\mathbf{k}$, which corresponds to the different Fourier modes of the density perturbations. Isotropy guarantees that $P(\mathbf{k}) = P(k)$, since $k = |\mathbf{k}|$. Because the power spectrum has units of volume, it is often expressed in terms of the dimensionless quantity:

$$\Delta^2(k) \equiv \frac{k^3}{2\pi^2} P(k) = \frac{d\sigma^2}{d\ln k} \,. \tag{1.20}$$

The last equality describes the contribution of the over-density to the variance $\sigma^2 \equiv \langle \delta^2 \rangle$ per logarithmic interval in k. The density field is typically convolved with a window function $W_R(r)$. The variance of the density field over the window $W_R(r)$ is:

$$\sigma_R^2(a) = \int_0^\infty \frac{dk}{k} \Delta^2(k) \widetilde{W}_R^2(k) \,. \tag{1.21}$$

Eq. 1.21 describes the *rms* mass fluctuation with respect to the average density field averaged over a spherical volume of radius R, assuming the fluctuations are linear. An important

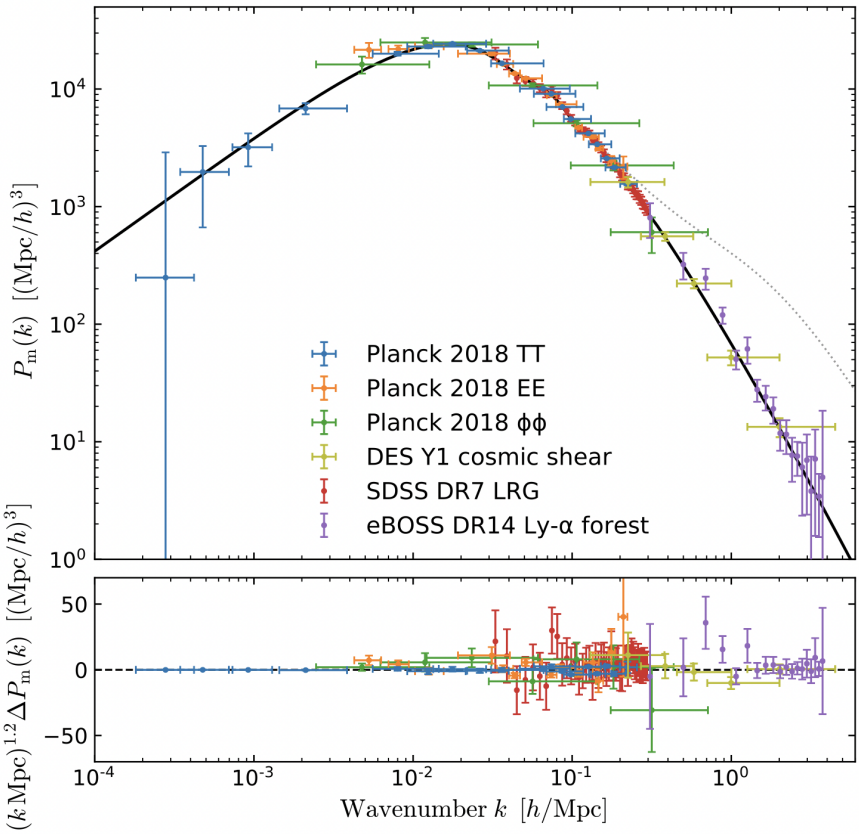


Figure 1.5: The top panel shows the 3D linear matter power spectrum as a function of spatial scale. Data points from Planck CMB data populate the largest scales (small k), SDSS galaxy clustering data are shown on intermediate scales, and SDSS Lyα clustering and DES cosmic shear data are on the smallest scales (large k). The solid black line represents the Planck 2018 best-fit ΛCDM model, and the dotted line shows theoretical predictions of non-linear effects on small scales. The bottom panel shows deviations of the data from the Planck best-fit ΛCDM 3D matter power spectrum. Figure is from Chabanier et al. (2019).

parameter used to measure clustering is S_8, which is given in terms of the amplitude of fluctuations within a radius of 8Mpc, σ_8, and the matter density parameter:

$$S_8 \equiv \sigma_8(\Omega_m/0.3)^{0.5}\,. \tag{1.22}$$

The two-point correlation function is the 3D Fourier transform of the power spectrum, and is a useful method of characterizing fluctuations in configuration space:

$$\xi(r) \equiv \langle \delta(\mathbf{x})\delta(\mathbf{x}+\mathbf{r})\rangle = \int_0^\infty \frac{dk}{k}\Delta^2(k)\frac{\sin(kr)}{kr}\,. \tag{1.23}$$

A plot of linear matter power spectrum measurements using clustering, shear, and CMB data, along with a ΛCDM best-fit model, is shown in Figure 1.5.

1.4 Observational probes of large-scale structure

Dark energy manifests itself through its impact on the distance-redshift relation (see Fig. 1.1) (or equivalently, the expansion rate), and via the rate at which matter clusters with time, which involves a balance between the gravitational attraction of matter over-densities and cosmic expansion. The study of structure formation examines the clustering of galaxies to quantify the rate of growth of structures from the early universe to the present. Accurate measurements of the clustering at different spatial scales and as a function of redshift can therefore provide stringent constraints on different cosmological models.

Maps such as the one shown in Fig. 1.6 reveal that the distribution of matter in the universe is not random. Rather, the cosmic web consists of entangled filaments and sheets interrupted by large voids. The seeds of this large-scale structure are density perturbations in the early universe that have evolved over time through a tug-of-war between radiation, matter, and dark energy. Probing this structure by measuring the amount of clustering at different scales and at different points in time can therefore constrain theories of gravity, the expansion history of the universe, and the growth of cosmic structure over time.

Accelerated expansion suppresses the clustering of matter due to gravity, and hence leaves an imprint on the expansion rate $H(z)$, distances $D_A(z)$, the matter power spectrum $P(k)$, and the "lumpiness" of the universe, S_8. We will first discuss a standard method for constraining

[7] `https://www.sdss.org`

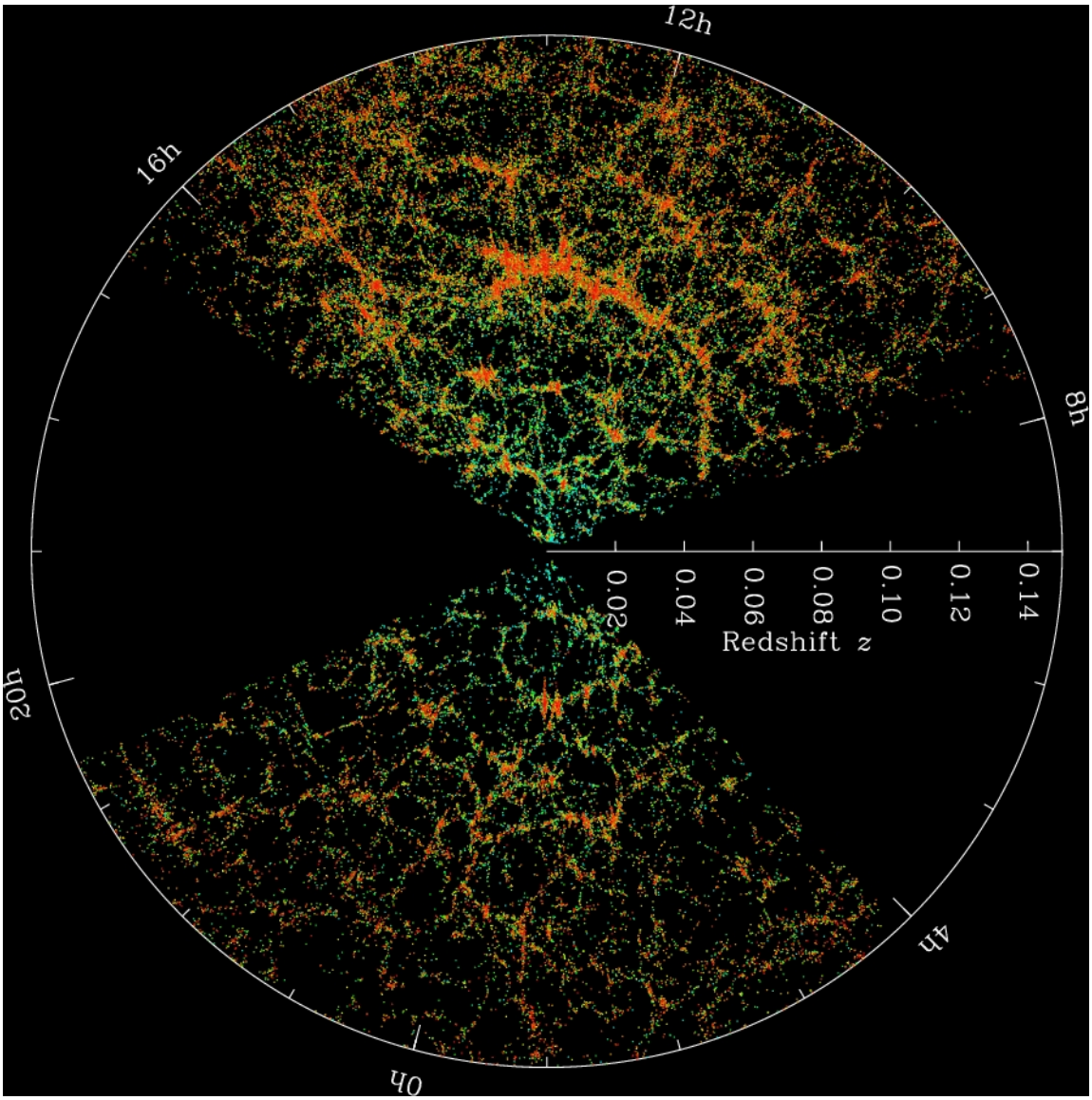


Figure 1.6: A 3D map showing galaxy clustering up to $z = 0.16$ using SDSS[7] data, with Earth at the center. Galaxies are colored based on the ages of their stars, with redder, more strongly clustered galaxies containing older stellar populations. The dark wedges were not mapped due to obscuration caused by dust in the Milky Way in these directions. Credit: M. Blanton and the Sloan Digital Sky Survey.

the latter two parameters via a combined analysis of clustering and shape distortions of galaxies called 3x2-point tomography. Next, we will cover a spectroscopic probe involving baryon acoustic oscillations that provides constraints on $H(z)$ and $D_A(z)$.

1.4.1 3x2-point tomography

The 3x2-point probe combines measurements from two different observational probes (clustering and cosmic shear) via three different 2-point correlation functions (galaxy-galaxy, galaxy-shear, and shear-shear). These measurements can be analyzed tomographically (as a function of cosmological redshift) to constrain the growth of structure over cosmic time. We covered clustering measurements from correlations between the density contrast at different spatial coordinates in Section 1.3.2. Before introducing the three observables for the 3x2-point probe, we will briefly discuss the theory of weak gravitational lensing – the mechanism behind cosmic shear. For more detailed coverage of this topic, the reader is encouraged to refer to Bartelmann & Schneider (2001), Weinberg et al. (2013), Mandelbaum (2018), Kilbinger (2015), and LSST Science Collaboration and LSST Project (2009).

Weak gravitational lensing

With the rise in the statistical power of photometric surveys, tomographic weak lensing has become one of the most promising cosmological probes of accelerated expansion. Gravitational lensing refers to distortions in the intrinsic shapes of galaxies due to the gravitational pull of intervening matter. When these distortions cause small perturbations in the position, size, brightness, and shape of galaxies (rather than multiple images or arcs), they fall into the category of weak lensing. When the shapes of neighboring galaxies are correlated due to the fact that their light passes through the same nearby structures, it is referred to as

cosmic shear. The amplitude of this correlation and its dependence on angular scales can be used to infer a map of the underlying distribution of matter.

For a single lensed source galaxy, the mapping from lensed (x_l, y_l) to unlensed (x_u, y_u) coordinates is given via the following linear transformation:

$$\begin{pmatrix} x_u \\ y_u \end{pmatrix} = \begin{pmatrix} 1-\gamma_1-\kappa & -\gamma_2 \\ -\gamma_2 & 1+\gamma_1-\kappa \end{pmatrix} \begin{pmatrix} x_l \\ y_l \end{pmatrix} . \tag{1.24}$$

The complex shear $\gamma = \gamma_1 + i\gamma_2$ describes the stretching of the galaxy image, and is represented as a two-dimensional spinor. The convergence κ describes the change in size and surface brightness of the galaxy due to lensing. Because we lack precise information about the unlensed galaxies' sizes and shapes, the transformation in Eq. 1.24 can be recast in terms of the reduced shear $g_i = \gamma_i/(1-\kappa)$, which combines the shear and convergence components:

$$\begin{pmatrix} x_u \\ y_u \end{pmatrix} = (1-\kappa) \begin{pmatrix} 1-g_1 & -g_2 \\ -g_2 & 1+g_1 \end{pmatrix} \begin{pmatrix} x_l \\ y_l \end{pmatrix} . \tag{1.25}$$

The lensing shear induces a change in the intrinsic galaxy ellipticities, ε, resulting in the observed ellipticity ε':

$$\varepsilon' = \frac{\varepsilon + g}{1 + g^*\varepsilon} . \tag{1.26}$$

The ellipticity is also expressed as a complex spinor $\varepsilon = \varepsilon_1 + i\varepsilon_2$, with components ε_1 and ε_2 corresponding to changes in the intrinsic galaxy shape along and 45° to the $\hat{x}$-direction, respectively.

A key assumption in weak lensing studies is that intrinsic galaxy ellipticities are randomly oriented on the sky, and therefore uniformly distributed in angle. The reduced shear can then be expressed as a first-order approximation of the average of the observed ellipticities:

$$\langle \varepsilon' \rangle = g \,. \tag{1.27}$$

Recent cosmic shear studies from Stage-III surveys include results from the Dark Energy Survey[8] (DES; Amon et al. (2022)), the Hyper Suprime-Cam Subaru Strategic Program[9] (HSC; Hamana et al. (2020)), and the Kilo-Degree Survey[10] (KiDS; Asgari et al. (2021)). The increase in available data from Stage-IV experiments will enable averaging the lensing signal over large ensembles of galaxies, resulting in a significant reduction in statistical errors. Systematic errors, therefore, will become the dominant source of error for weak lensing studies. One of the main sources of systematics will be shape noise due to the shear signal being small compared to intrinsic galaxy ellipticities. Extracting these ellipticities requires an accurate characterization of both the atmospheric and telescope point spread function (PSF)[11]. Moreover, because shape noise dominates over pixel noise in LSST sources with a detection significance above a 5σ point-source depth, galaxies used for weak lensing studies are typically small and faint. Atmospheric turbulence, as well as anomalies and distortions in the instrument-camera system can also contribute to weak lensing systematics (see Chapter 3). Survey requirements and details of the systematic error budget for LSST will be further discussed in Section 1.5.1.

[8] `https://www.darkenergysurvey.org`

[9] `https://hsc.mtk.nao.ac.jp/ssp/`

[10] `https://kids.strw.leidenuniv.nl`

[11] The PSF is the 2D flux representation of a point source on the focal plane. Contributions to the PSF come from atmospheric turbulence (or seeing) as well as the instrument, which includes the telescope optics, deviations from a flat focal plane, and sensor anomalies. Together these effects manipulate point sources such that they are rendered as anisotropic blobs on the detector.

3x2-point observables

The three different correlation functions for the 3x2-point probe measure the degree of clustering as a function of the angular separation between galaxies η on the sky, $\langle\delta_\eta\delta_\eta\rangle$, the correlated distortions in the shapes of lensed galaxies, $\langle\gamma\gamma\rangle$, and the cross-correlation between galaxy positions and shapes, $\langle\delta_\eta\gamma\rangle$. The latter function quantifies the degree to which foreground galaxies in one part of the sky give rise to distortions in the ellipticities of background galaxies in another region.

Although galaxy clustering is a tracer of the density field, it is a biased tracer, as galaxies tend to form only in peaks with a high enough density contrast. Therefore, the galaxy density contrast is equal to the matter contrast, modified by a linear bias term, b, $\delta_\eta = b\,\delta_m$. Galaxy bias changes the amplitudes of $\xi(r)$ or $P(k)$, but not their shapes. The target observable for clustering is the correlation function $C^{ij}_{\delta_\eta\delta_\eta}(l)$, given in terms of the angular multipole l[12]:

$$C^{ij}_{\delta_\eta\delta_\eta}(l) = \int d\chi \frac{q^i_{\delta_\eta}(\frac{l+1/2}{\chi}\chi,\chi)\, q^j_{\delta_\eta}(\frac{l+1/2}{\chi}\chi,\chi)}{\chi^2} P_{NL}\Big(\frac{l+1/2}{\chi}, z(\chi)\Big)\,. \tag{1.28}$$

The superscripts i, j indicate distinct galaxies, the non-linear (NL) matter power spectrum P_{NL} is given in Eq. 1.18, and χ is the comoving distance. The weights $q^i_{\delta_\eta}(k,\chi)$ are given in terms of the bias b^i of the clustering sample:

$$q^i_{\delta_\eta}(k,\chi) = b^i \frac{n^i_\eta(z(\chi))}{\bar{n}^i_\eta}\frac{dz}{d\chi}\,, \tag{1.29}$$

where the average surface density of galaxies per bin is:

[12]The angular multipole l corresponds to an angular scale $\delta\theta = l/D_A(z)$, where $D_A(z)$ is the angular-diameter distance to the last scattering surface.

$$\bar{n}^i_{\eta/\gamma} = \int dz\, n^i_{\eta/\gamma}(z)\,. \tag{1.30}$$

A lensing sample containing galaxy redshift and ellipticity measurements is used to measure the shear-shear correlation function:

$$C^{ij}_{\gamma\gamma}(l) = \int d\chi \frac{q^i_\gamma(\chi)\, q^j_\gamma(\chi)}{\chi^2} P_{NL}\Big(\frac{l+1/2}{\chi}, z(\chi)\Big)\,. \tag{1.31}$$

Unlike galaxy clustering, there is no galaxy bias in cosmic shear, and the weights $q^i_\gamma(\chi)$ are:

$$q^i_\gamma(\chi) = \frac{3H_0^2\Omega_m}{2c^2}\frac{\chi}{a(\chi)}\int_\chi^{\chi_h} d\chi' \frac{n^i_\gamma(z(\chi'))dz/d\chi'}{\bar{n}^i_\gamma}\frac{\chi'-\chi}{\chi'}\,, \tag{1.32}$$

where the average surface density of galaxies per bin is given in Eq. 1.30. Finally, the observable for galaxy-galaxy lensing is given by:

$$C^{ij}_{\delta_\eta\gamma}(l) = \int d\chi \frac{q^i_{\delta_\eta}(\frac{l+1/2}{\chi}\chi, \chi)\, q^j_\gamma(\chi)}{\chi^2} P_{NL}\Big(\frac{l+1/2}{\chi}, z(\chi)\Big)\,. \tag{1.33}$$

Here, the weights are a combination of the clustering and lensing weights, $q^i_{\delta_\eta}(k,\chi)$ and $q^i_\gamma(\chi)$, and contain a bias factor which depends on the redshift distribution of the lensing and clustering samples.

1.4.2 BAO in the Lyα forest of high-redshift quasars

The early universe was a hot, dense plasma comprised of photons, baryons, electrons, neutrinos, and dark matter. The baryons and photons were initially coupled together, as seen in the first panel of Fig. 1.7. If we consider a single perturbation in the plasma, whose

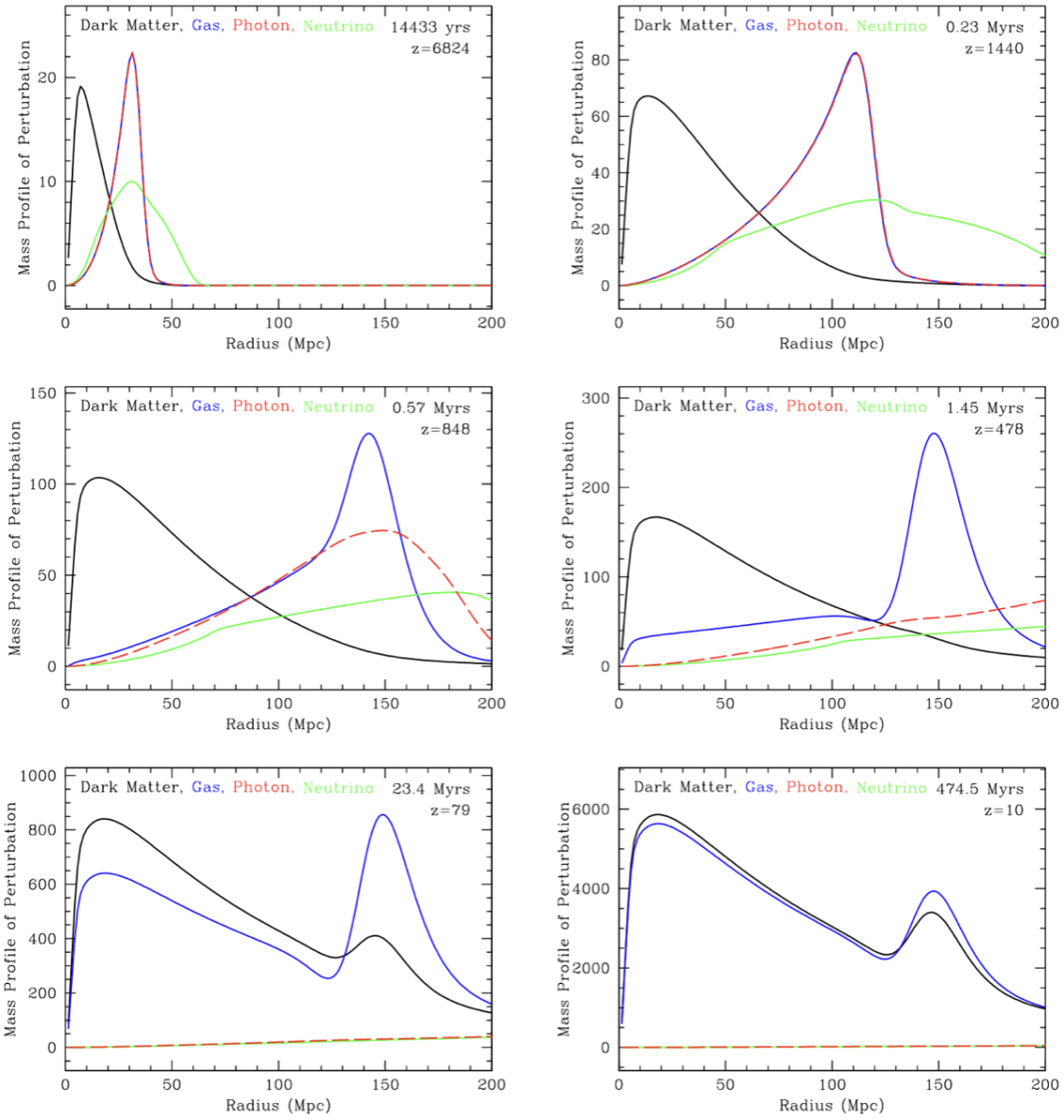


Figure 1.7: Snapshots showing the evolution of perturbations in the primordial plasma for different species of dark matter, baryonic gas, photons, and neutrinos. The redshift and time after the Big Bang are given in each pane. An initially point-like overdensity at the origin is gradually pushed out due to radiation pressure from the photons until the universe has expanded enough for photons to decouple from the plasma. Without the outward radiation pressure, protons and electrons recombine and are gradually drawn back in by the gravitational pull of the central dark matter over-density. The photon density profile is flat at late times, representing the homogeneity of the CMB. Figure is from Eisenstein et al. (2007).

density is uniform except for an excess of dark matter at the center, the baryon-photon fluid is pushed out due to the radiation pressure of the photons. This expansion, exceeding half the speed of light (around 170,000km/s), continues for about 100,000 years, until the universe has expanded just enough for photons to travel without scattering off of other particles. Once the photons decouple and begin to free-stream through the universe, the baryons, no longer under the influence of outward pressure from the photons, remain fixed in place (see the third pane of Fig. 1.7). A shell of baryonic matter is left behind at a radius of ~ 100Mpc, while the photons become more and more smoothly distributed. Eventually, the gravitational pull of the dark matter starts to draw the baryons back toward the center of the perturbation until they reach equilibrium densities (see the last pane of Fig. 1.7), leaving a peak at the center with an echo in a shell at $r_d = 150$Mpc. This comoving scale is a feature of these primordial density perturbations, called baryon acoustic oscillations (BAO), and is imprinted in the distribution of large-scale structure at late times. The BAO method uses r_d to directly constrain the comoving distance $D_M(z)/r_d$ and the Hubble distance $D_c(z)/r_d$ at different redshifts. In doing so, it provides a path to studying expansion via the Hubble parameter, as well as the geometry of the universe via D_M.

1.4.3 The Lyα forest

The first BAO measurements were made by Eisenstein et al. (2005) and Cole et al. (2005) using the auto-correlation of galaxy positions and the galaxy power spectrum, respectively, at $z \leq 1$. Because the number density of galaxies necessary for high precision measurements is lacking at $z \geq 2$, discrete galaxies are insufficent tracers of the BAO scale in this redshift regime. Measurements of the BAO scale at these redshifts have instead been performed by studying the absorption of neutral hydrogen in the Lyα region of the spectra of high-redshift quasars.

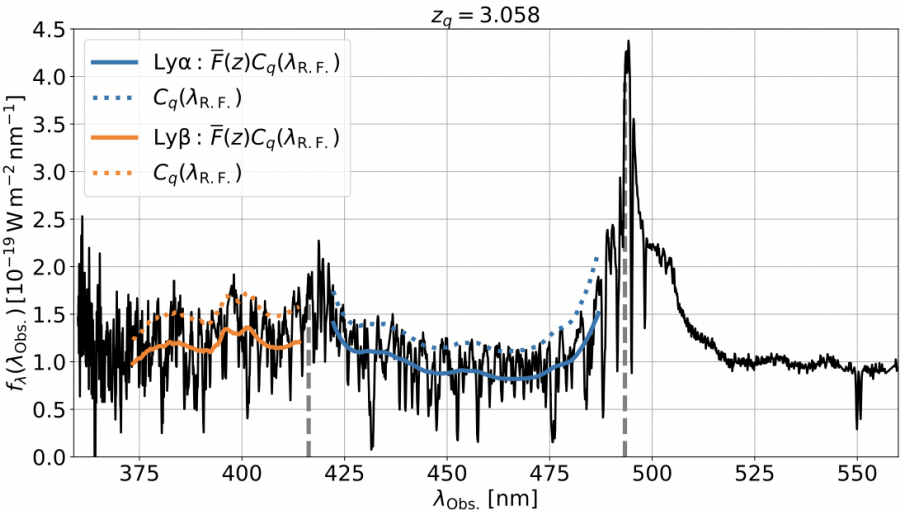


Figure 1.8: A high signal-to-noise quasar at $z = 3.058$, with the best-fit models for $\bar{F}(\lambda)C_q(\lambda_{RF})$ in the Lyα and Lyβ regions shown as solid blue and orange lines, respectively. The unabsorbed continuum model for each is shown as a dotted line. Figure is from du Mas des Bourboux et al. (2020).

The Lyα line is a spectral line of hydrogen corresponding to the transition from the $n = 2$ orbital to the ground state $n = 1$. This transition occurs in the spectrum of a background source each time its light interacts with clouds of hydrogen gas along the line of sight, producing a succession of transition lines called the Lyα forest (see Fig. 1.8). Because the rest wavelength of this transition is 1215.67Å (in the UV part of the electromagnetic spectrum), it can only be observed if it has been sufficiently redshifted into the optical wavelength range. High-redshift quasars fulfill both the number density and redshift requirements to be continuous tracers of this absorption feature, and opacity fluctuations in their spectra are therefore a highly sensitive probe of density fluctuations in the intergalactic medium.

1.4.4 BAO formalism

Rather than extracting the BAO scale from the correlated positions of galaxies, the Lyα probe measures correlations between the transmitted flux along the line-of-sight of high-

redshift quasars. The flux-transmission field of quasar q is defined in terms of the ratio of the observed flux, $f_q(\lambda)$, to the mean expected flux, $\bar{F}(\lambda)C_q(\lambda)$:

$$\delta_q(\lambda) = \frac{f_q(\lambda)}{\bar{F}(\lambda)C_q(\lambda)} - 1\,. \tag{1.34}$$

The mean expected flux is a product of the unabsorbed quasar continuum, $C_q(\lambda)$, and the mean transmission, $\bar{F}(\lambda)$, at a given wavelength. The separation between two tracers is given by the vector $\mathbf{r} = (r_{||}, r_{\perp})$, which indicates separations along and perpendicular to the light of sight. Once we measure $\Delta\theta$ and Δz, we can assume a cosmology $H(z)$ to determine this separation:

$$r_{||} = \big[D_c(z_i) - D_c(z_j)\big] \cos\left(\frac{\Delta\theta}{2}\right) \tag{1.35}$$

$$r_{\perp} = \big[D_M(z_i) + D_M(z_j)\big] \sin\left(\frac{\Delta\theta}{2}\right), \tag{1.36}$$

where $D_c = c/H(z)$ and $D_M = (1+z)D_A$ are the comoving distance and the comoving angular diameter distance, respectively. The Ly$\alpha\times$Lyα correlation function is computed for pairs of points at fixed separation $(\Delta\theta, \Delta z)$:

$$\xi(\Delta\theta, \Delta z) = \langle \delta_q(\theta_1, z_1)\delta_q(\theta_2, z_2)\rangle \tag{1.37}$$

$$\propto \frac{1}{N} \sum_{\text{all pairs}} \delta_q(\theta_1, z_1) \times \delta_q(\theta_2, z_2)\,, \tag{1.38}$$

where all pairs satisfy $|\theta_1 - \theta_2| < \Delta\theta$ and $|z_1 - z_2| < \Delta z$. The expected comoving scale of the acoustic peak is $r_{BAO} \sim 100h^{-1}$Mpc. Therefore, the different correlations are typically

computed for separations up to twice the BAO scale ($\pm 200 h^{-1}$Mpc along and across the line-of-sight), within bins of a given width (typically $4h^{-1}$Mpc wide in both directions). For two spectral pixels in the flux-transmission field, i and j, and their associated weights w_i and w_j, which account for the statistical power of each pixel, the correlation function is:

$$\xi_A = \frac{\sum\limits_{(i,j)\in A} w_i w_j \delta_i \delta_j}{\sum\limits_{(i,j)\in A} w_i w_j} \,. \tag{1.39}$$

Results of the Ly$\alpha\times$Lyα auto-correlation function for the complete extended Baryon Oscillation Spectroscopic Survey (eBOSS[13]; du Mas des Bourboux et al. (2020)) are shown in Fig. 1.9. Each of the four panels show the correlation function for different angular separations along the line of sight. This separation is parameterized by $\mu = r_{||}/r$, or the cosine of the angle formed by the median line-of-sight of both tracers and the vector $\mathbf{r}$.

1.5 Stage-IV dark energy experiments

The distinguishing feature of Stage-IV surveys is the unprecedented quantity and quality of data they will produce compared to Stage-III. Specifically, Stage-IV surveys are expected to improve upon a quantitative measure of the sensitivity of a cosmological probe, defined by the Dark Energy Task Force (DETF), called the figure-of-merit (FoM), by at least a factor of 10 over Stage-II surveys. This section covers the main features of LSST and DESI, along with a discussion of their respective survey requirements and figures-of-merit.

[13] `https://www.sdss.org/surveys/eboss/`

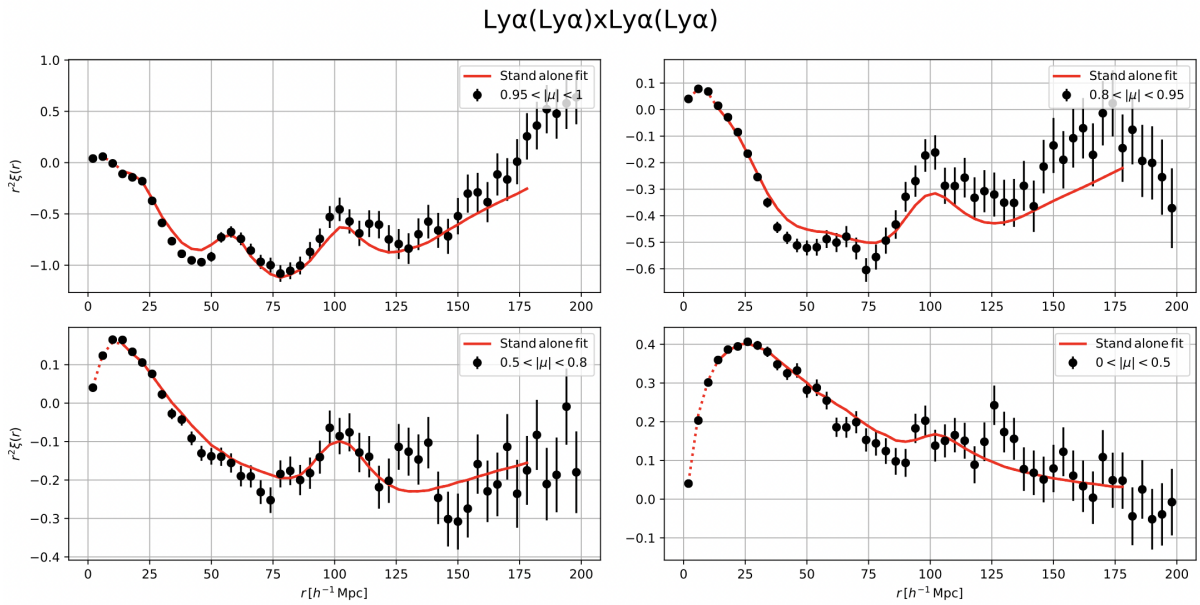


Figure 1.9: The Lyα auto-correlation function for different ranges of $|\mu| = |r_{||}|/r$, multiplied by r^2 to render the y-axis dimensionless and accentuate the BAO peak. In each of the four panels there is a noticeable bump in the correlation function at the expected scale of $r = 100h^{-1}$ Mpc. The best-fit model is given in red, with the dashed red lines showing the best-fit model interpolated beyond the fitted range. Figure is from du Mas des Bourboux et al. (2020)

1.5.1 The Legacy Survey of Space and Time

LSST is a 10-year optical survey that will map the southern sky from the Cerro Pachón ridge in Chile. It will focus on addressing four key areas of science: probing dark matter and dark energy, taking an inventory of the solar system, studying the transient sky, and mapping the Milky Way (Željko Ivezić, and the LSST Science Collaboration 2018). The Dark Energy Science Collaboration[14] (DESC; LSST Dark Energy Science Collaboration (2012)) is a community of scientists dedicated to designing and implementing the infrastructure needed to conduct dark energy research with LSST data products. This includes developing software for data reduction, simulations to both verify science pipelines and test cosmological models, and computational resources for data storage and processing. The DESC will probe the growth of structure via tomographic galaxy clustering, cosmic shear, and clusters, and pro-

[14] `https://lsstdesc.org/index.html`

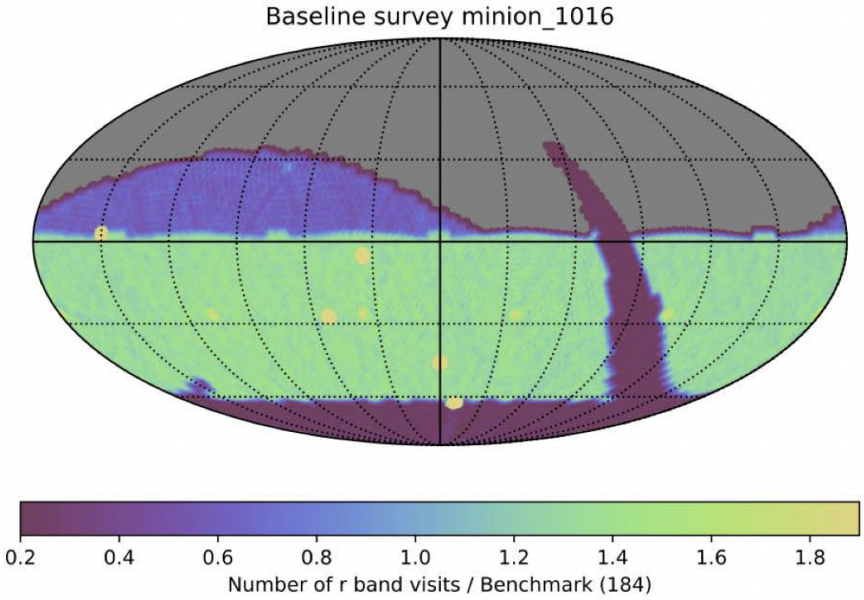


Figure 1.10: The number of visits in the r-band, normalized by 184, for the baseline survey over the LSST footprint. Note that regions corresponding to the Galactic plane (arc on the right) and the northern Ecliptic region (upper left), and the South Celestial Pole (bottom) show fewer visits. Deep drilling fields have a higher number of visits and are shown as small circles. Figure is from Ivezić et al. (2019).

vide constraints on the expansion rate using Type Ia supernovae, strong lensing, and baryon acoustic oscillations (The LSST Dark Energy Science Collaboration 2021). The diversity of these probes can be leveraged to break degeneracies, as well as to test and compare different cosmological models, specifically ΛCDM, against theories of modified gravity.

LSST system design

LSST has adopted a straightforward survey strategy to maximize the return on its science goals. The baseline survey, which accounts for 90% of observing time, involves wide, deep, and fast scans of the sky, known as "universal cadence." The remaining 10% of survey time will be dedicated to other programs. One such program, the Deep Drilling Field, will devote an hour of observing time to image a single field to substantially greater depth for

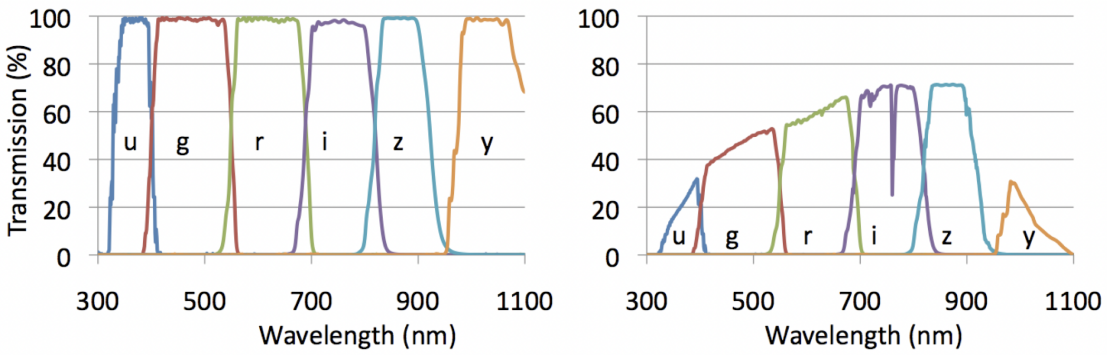


Figure 1.11: The figure on the left shows the theoretical transmission efficiency of the six LSST filters. In practice, the throughput is diminished due to absorption by the atmosphere, loss due to telescope optics, and the quantum efficiency of the sensors, as seen on the figure on the right. Figure is from LSST Science Collaboration and LSST Project (2009).

Solar System, Galactic, and extragalactic studies. The exceptionally wide field-of-view of LSST's dedicated Simony Survey Telescope means that each field covers 9.6 deg^2 on the sky, approximately 40 times the area of the full moon, where a visit consists of a pair of back-to-back 15-second exposures. The distribution of visits in the r-band for the baseline main survey is shown in Fig. 1.10. Because uniform coverage and depth of the sky is necessary for cosmological parameter estimation, LSST will visit each field up to 800 times to reach sufficient depth in six different passbands (*ugrizy*) (see Fig. 1.11). This translates to imaging every field in its 18,000 deg^2 footprint every few nights. Imaging over such a broad range of filters, and to such high depth is critical for accurate photometric redshifts, as well as for achieving the high signal-to-noise and control of systematic uncertainties necessary for weak lensing measurements. Upon completion, LSST will have imaged roughly 20 billion galaxies, 17 billion Milky Way stars, and 6 million Solar System objects.

LSST's survey strategy is made possible by its optical and telescope design. The telescope consists of an 8.4-meter primary mirror (M1) with a 5.1-meter inner clear aperture, a 3.4-meter convex secondary mirror with an 1.8-meter inner opening (M2), and a 5-meter tertiary mirror (M3), along with three refractive lenses (L1, L2, L3). A filter wheel rotates between

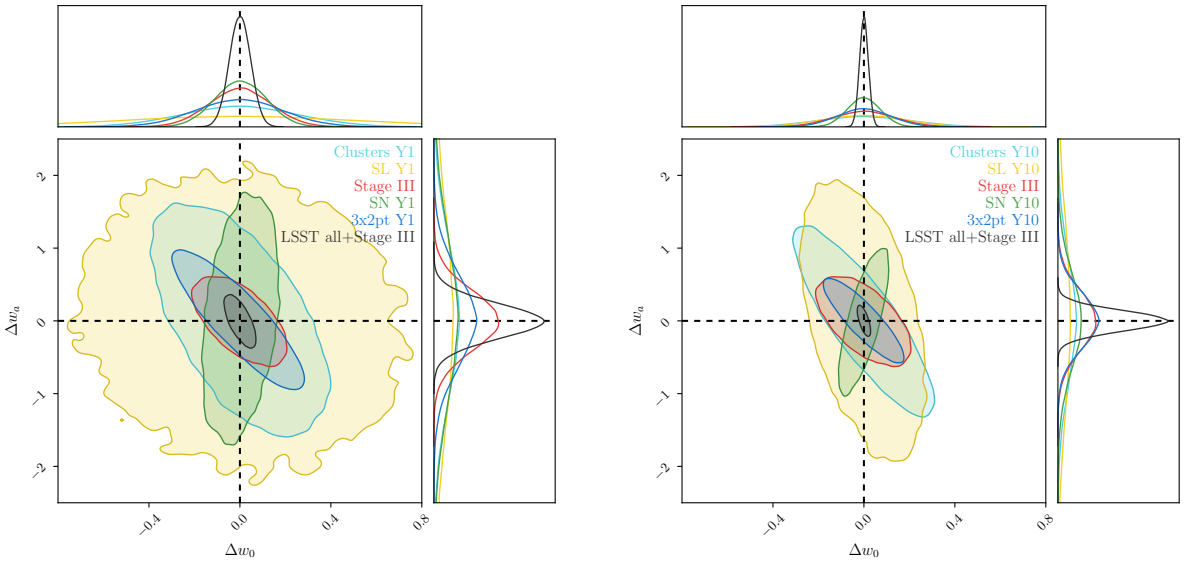


Figure 1.12: Y1 (left) and Y10 (right) forecasts for dark energy constraints from individual probes, as well as the joint forecast including Stage-III priors. The contours represent 68% confidence intervals for all cases, and the quanitities Δw_0 and Δw_a represent the difference between w_0 and w_a from their fiducial values of -1 and 0, respectively. Figure is from The LSST Dark Energy Science Collaboration (2021).

each of the six different filters. A 3.2 Gigapixel camera enables a two-second readout of the entire focal plane. Details of the LSST camera are deferred to Chapter 3.

LSST science requirements

The LSST Science Requirements Document (Željko Ivezić, and the LSST Science Collaboration 2018) outlines specifications for achieving the DETF FoM for Stage-IV dark energy experiments. Given a model for the dark energy equation of state (Eq. 1.13), the FoM is defined as $\sqrt{|F|}$, where F is the Fisher matrix for (w_0, w_a), marginalized over all other parameters (both cosmological and nuisance parameters). The Fisher matrix forecasts the best possible performance of a given model experiment in constraining the parameters of the model. In the case of constraining the dark energy equation of state, it provides (Gaussian) uncertainties on $\sigma(w_0)$ and $\sigma(w_a)$ in terms of the contours enclosing the 1-2σ confidence level

Analysis	Priors	Y1 FoM (ceiling)	Y10 FoM (ceiling)	Target
LSS	Stage-III (not w_0, w_a)	10 (13)	10 (14)	1.5
LSS	None	6.7 (8.4)	6.6 (9.1)	1.5
WL+LSS	Stage-III (not w_0, w_a)	31 (37)	66 (87)	40
WL+LSS	None	22 (27)	49 (68)	40
All	Stage-III	142 (156)	505 (711)	500
All	None	108 (135)	461 (666)	-

Table 1.1: Figures-of-merit for Y1 and Y10 analyses of clustering and weak lensing, as well as for a combination of all probes (Željko Ivezić, and the LSST Science Collaboration 2018).

in the w_0-w_a plane. Fisher forecasts for the dark energy equation of state for LSST Year 1 (Y1) and Year 10 (Y10) analyses are shown in Figure 1.12.

LSST system requirements include a single visit depth of $r \sim 24.5$[15] for a 5σ point source, which is necessary to image faint objects, as well as variable transient sources. The co-added survey depth should reach $r \sim 27.5$ in order to average over systematics in PSF estimation for weak lensing science. The depth requirement is also tied to stringent constraints on image quality, which should be atmosphere-limited, with a median seeing of 0.7 arcsec in the r-band. In addition, the combined number of visits of any field in all filters should be on the order of 1,000. The distribution of visits per filter should enable accurate photometric redshifts, and the distribution of visits on the sky should cover LSST's footprint to obtain the required number density of galaxies for weak lensing studies. Photometry must be measured to 5 millimag precision to maintain zeropoint stability across the sky of 10 millimag, and to keep band-to-band calibration errors smaller than 5 millimag. Finally, transmission through each of the six filters should only be limited by atmospheric absorption and silicon detection efficiency, with no large gaps in coverage. An overview of LSST camera requirements is detailed in Chapter 3.

In addition to the baseline survey requirements, the DESC Science Requirements Document (The LSST Dark Energy Science Collaboration 2021) details its own set of high-level objectives specific to each observational probe, as well as for combined probes. The requirements

[15]The r passband is shown in Figure 1.11.

for the joint-probe combined 1σ statistical and systematic uncertainties on the equation-of-state parameters (w_0, w_a), consistent with the goals outlined by the DETF for Stage-IV experiments, are $\sigma(w_0) = 0.02$ and $\sigma(w_a) = 0.14$ (after marginalization over all other parameters). For the full LSST Year 10 (Y10) analysis, DESC dark energy probes will achieve a combined FoM exceeding 500 when including both statistical and systematic uncertainties, using Stage-III priors. For each individual probe, the improvements in the FoM should be at least twice the corresponding Stage-III probe (see Table 1.1).

Due to the unprecedented amount of high-quality data from LSST, statistical uncertainties will be sub-dominant compared to quantifiable measurement offsets due to observational or astrophysical issues. In other words, cosmological parameter measurements will be systematics-limited. Some of the self-calibrated systematic uncertainties for galaxy clustering include instrinsic alignments, mass function uncertainty, baryonic effects, and cluster large-scale bias. A dominant systematic for all probes concerns photometric redshift estimates (see Chapter 2). Clustering requirements therefore include limits on the systematic uncertainty in the mean redshift of each tomographic bin so that it does not exceed $0.003(1 + z)$, and keeping the systematic uncertainty in the photometric redshift scatter, σ_z, from exceeding $0.03(1 + z)$.

These requirements are more strict for the combined clustering and cosmic shear measurements, where systematics arise due to (non-linear) galaxy bias, magnification, intrinsic alignments (Krause & Eifler 2017), baryonic effects, and blending (Sanchez et al. 2021). For the 3x2-point probe, the uncertainties in the mean redshift and photometric redshift scatter should not exceed $0.001(1 + z)$ and $0.003(1 + z)$, respectively. Moreover, systematic uncertainty in the redshift-dependent shear calibration should not exceed 0.003. Since shape measurements are significantly influenced by improper characterization of the PSF, the systematic uncertainty in the PSF model size should not exceed 0.1%. All probes of structure growth will need to optimize tomographic binning schemes to reduce redshift-dependent

effects, and to eliminate redshift-related systematic biases. Attempts to find an optimal binning strategy within the DESC are ongoing, with preliminary results detailed in Zuntz et al. (2021).

1.5.2 Dark Energy Spectroscopic Instrument

DESI is a 5-year dark energy project that will measure the expansion rate with unprecedented precision using the BAO technique up to redshifts of $z \sim 3.7$, before the transition to a dark-energy dominated universe. Sited at the 4-meter Mayall Telescope at Kitt Peak National Observatory, the DESI instrument will use its 5,000 robotically-controlled fibers and 10 triple-arm spectrographs to obtain spectra of 35 million galaxies and quasars over a 14,000 $\deg^2$ area on the sky. The fibers are equally divided onto 10 petals on the focal plane. Light from each of the 500 fibers on a single petal is transmitted to a slit head in its designated spectrograph via $\sim$50 m cables, where it is split into three separate channels by a pair of dichroics to produce a continuuous spectrum for each object from 360-980 nm (Jelinsky et al. 2018).

In addition to measuring the cosmic distance scale using BAO, DESI will construct a 3D map of the large-scale structure in the universe to study the growth of structure through redshift-space distortions (RSD). It will primarily target three different classes of objects in distinct redshift regimes, namely luminous red galaxies (LRGs) between $0.4 < z < 1$, emission-line galaxies (ELGs) covering $0.6 < z < 1.6$, and Lyα quasars (QSOs) at $z > 2$ (Levi et al. 2019). DESI will also conduct a flux-limited survey of bright galaxies to $r \sim 19.5$, with a median redshift of $z \sim 0.2$ (see Table 1.2). Further discussion of these classes and how they are targeted is deferred to Section 4.1.2.

Object Type	Redshift Range	N/deg^2	# of successful objects	
			9k	14k
LRGs	0.4 – 1.0	300	2.7×10^6	4.2×10^6
ELGs	0.6 – 1.6	1280	11.5×10^6	17.9×10^6
Tracer QSOs	< 2.1	120	1.1×10^6	1.7×10^6
Lyα QSOs	> 2.1	50	0.45×10^6	0.7×10^6

Table 1.2: Survey parameters for the minimum-requirement 9,000 deg^2 and the baseline 14,000 deg^2 surveys (DESI-doc-318).

DESI science requirements

Survey requirements are carefully detailed in the DESI Science Requirements Document (SRD; DESI-doc-318). Over the course of the survey, DESI will cover at least 9000 deg^2, with a baseline survey area of 14,000 deg^2. For LRGs, ELGs, and tracer QSO targets, the BAO cosmic distance scale will be measured to 0.28% precision aggregated over the redshift range $0.0 < z < 1.14$ and to 0.39% precision in the redshift range $1.1 < z < 1.9$. For the baseline survey, the precision to which this scale is measured for these redshift ranges will be 0.22% and 0.31%, respectively. At higher redshifts, Lyα quasars will measure the Hubble parameter at $1.9 < z < 3.7$ from the BAO method to 1.05% precision for the minimum survey, and 0.84% for the baseline survey. Finally, systematic errors from the instrument and observational methods should not exceed 0.16% for measurements of the angular diameter distance $D_A(z)$ and 0.26% for $H(z)$ using galaxies at $z < 1.5$. FoM requirements are given in Table 1.3

The DESI SRD also details requirements for each target class. Those pertaining specifically to Lyα QSOs will be listed here. The average density of quasars at $z > 2.1$ and $r < 23.5$ used for Lyα or QSO clustering shall be at least 50 per deg^2. Lyα QSOs used as tracers in cross-correlations should have a redshift accuracy of $\sigma_z = 0.0025(1 + z)$, with redshift catastrophic failures below 2%. Moreover, the signal-to-noise per Ångstrom in the observer frame (360-570nm) of the Lyα forest must be greater than 1.

Technique	BOSS	DESI	
		9k	14k
Galaxy + Lyα BAO	37	114	143
+ DESI galaxy broad-band power to $k = 0.1\,h\mathrm{Mpc}^{-1}$		229	303
+ DESI galaxy broad-band power to $k = 0.2\,h\mathrm{Mpc}^{-1}$		525	687

Table 1.3: Figures-of-merit for the mininum-requirements 9,000 deg^2 and baseline 14,000 deg^2 surveys (DESI-doc-318).

1.6 Motivation for this thesis

Recent advances in survey design and instrumentation, along with a concomitant increase in the quality and amount of astrophysical data, will enable sub-percent precision on measurements of cosmological parameters, particularly those constrained by weak lensing and BAO. With DESI having launched in 2021, and LSST set to begin formal operations in 2022, there is significant potential to explore synergy between these two surveys. Failure to properly understand and mitigate systematics will ultimately limit the sensitivity for all dark energy probes. Fully exploiting the investment in both DESI and LSST therefore requires a detailed characterization of the systematics arising from redshift-estimation algorithms, as well as from the instrument and camera system. As more data becomes available, we can leverage DESI spectra to test and refine LSST's photometric redshift capabilities.

The following chapters aim to develop the infrastructure necessary to maximize the amount of information extracted from these two experiments by investigating the performance of a state-of-the art photometric redshift estimation algorithm on simulated LSST data, improving the systematic understanding of the LSST camera and sensors, and developing realistic simulations in DESI. This will enable more accurate measurements of cosmological parameters related to dark energy and ultimately allow us to better characterize the role of dark energy on cosmic expansion and structure growth.

Chapter 2

Redshift inference using a hierarchical Bayesian model

2.1 Introduction

The advent of Stage-IV dark energy surveys will bring forth an unprecedented amount of data, totalling tens of billions galaxies over vast cosmic volumes. Nearly all of the dark energy probes for DESI and LSST will rely on accurate redshift estimates of these galaxies to measure observables – such as the growth rate of matter density perturbations, the expansion rate, and baryon acoustic oscillations – as a function of cosmic time. The increase in available data will substantially reduce statistical uncertainties for all probes, most notably for tomographic weak lensing, which has the potential to be the most powerful individual technique among all Stage-IV surveys. One of the major limitations to studying cosmic expansion via tomographic cosmic shear measurements, however, other than instrumental effects on shape measurements, are systematic uncertainties due to redshift errors. Failure to accurately characterize redshift distributions can ultimately lead to biases in the estimation

of comsological parameters (Huterer et al. 2006). This is an issue of particular importance for photometric surveys like LSST, which will reach higher depths and rely on broad spectral features to estimate distances to galaxies.

2.2 Photometric redshift estimation

Precision redshifts can be measured via spectroscopy by fitting observed spectral features in high-resolution spectra to their rest frame values. Despite their robustness, spectroscopic redshifts are expensive in both time and resources. Not only does spectroscopy require longer exposure times, but there are also a limited number of targets that can be observed during each exposure. Moreover, the dearth of observed targets in deep spectroscopic surveys results in high incompleteness at faint magnitudes. Altogether, the pace at which galaxy images are obtained greatly surpasses the rate at which we can process galaxy spectra, making it infeasible to spectroscopically observe the vast majority of extragalactic objects in next-generation dark energy surveys. Stage-IV experiments must therefore rely on imaging to access this third spatial dimension by measuring photometric fluxes in multiple bands.

Photometric redshifts (photo-z's) use broadband flux information from galaxies that have been observed through a set of optical filters. This essentially provides a low-resolution spectrum of each object, albeit with a higher signal-to-noise ratio due to both deeper observations, and the wide wavelength coverage of each passband when combining fluxes (Newman et al. 2015). Broadly speaking, there are three classes of photo-z algorithms. Template-fitting methods (`Hyperz` (Bolzonella et al. 2000); `BPZ` (Benitez 2000); `LePhare` (Arnouts & Ilbert 2011); `EAZY` (Brammer et al. 2008)) utilize a library of empirical or synthetic galaxy spectral templates to minimize the χ^2 between their derived fluxes and observed photometric fluxes as a function of redshift. Machine learning methods (`ANNz` (Collister & Lahav 2004); `ArborZ` (Gerdes et al. 2010); `TPZ` (Carrasco Kind & Brunner 2013); `SkyNet` (Bonnett 2015)) involve

training a flexible model to directly learn the flux-redshift relation using measured and "true" fluxes. Clustering redshifts constrain redshift distributions using projected two-point galaxy clustering statistics to compare the spatial separation of galaxies in the survey to a tracer population with secure redshifts (Newman (2008); Matthews & Newman (2010); Ménard et al. (2014); `THE-WIZZ` (Morrison et al. 2017)).

Although these approaches have been successful in different contexts, each one comes with its own set of potential drawbacks. Spectroscopic surveys are usually incomplete at high redshifts and faint magnitudes, which happen to be the regions we are most interested in probing. The "true" fluxes used in photometric inferences can be biased, and a lack of tracer populations in certain redshift regimes can hinder clustering estimates (Schmidt et al. 2020). A hybrid approach, one that incorporates attributes of all three methods, can potentially bypass some of these issues by compensating for the limitations of using each technique used on its own.

This chapter investigates the application of a composite redshift inference method developed by Sanchez & Bernstein [SB18] (Sánchez & Bernstein 2018) on data simulated for the purpose of the DESC Data Challenge 2 (DC2) (Abolfathi et al. 2021). Generation of the feature likelihood in SB18 is then compared to a method using normalizing flows developed by David Kirkby (David Kirkby 2020).

2.3 Hierarchical Bayesian redshifts

In their paper, SB18 use a hierarchical Bayesian framework to simultaneously infer redshift distributions for populations of galaxies, as well as provide redshift estimates for their individual constituents. They perform a proof-of-principle by testing their method on a toy model, and then apply it to a simulated data set in a subsequent paper (Alarcon et al. 2020).

We will first focus on outlining their process by reproducing the results of their toy model in Section 2.5. This will provide a foundation for applying the method to simulated data from the cosmoDC2 catalog (Korytov et al. 2019) in Section 2.6. Finally, we will compare the discrimination of galaxy types in generating the feature likelihood with a method involving normalizing flows in Section 2.9.

2.3.1 Hierarchical models

Bayesian statistics is an interpretation of probability based on Bayes' theorem, which is a natural consequence of the law of conditional probabilities. Bayes' theorem is expressed in terms of events A and B as:

$$p(A|B) = \frac{p(B|A)\,p(A)}{p(B)}\,. \tag{2.1}$$

The terms $p(A|B)$, $p(B|A)$, $p(A)$, and $p(B)$ are commonly referred to as the posterior, the likelihood, the prior, and the evidence, respectively. In the context of cosmological inference, A is typically a parameter to be estimated, and B is observable data from one or more surveys.

The utility of this framework is that it allows us to update our degree of belief in an outcome by accommodating new data or information in a straightforward manner. Thus, a Bayesian approach, unlike frequentist statistics, can assign a degree of belief to a single event because the model parameters and hypotheses, rather than being fixed, are expressed as probabilities.

Eq. 2.1, as it stands, is not amenable to more multiplexed problems involving interdependent variables. For this purpose, we can extend this equation in a hierarchical fashion to express more complex parameter dependencies. For example, in the case where A follows a dis-

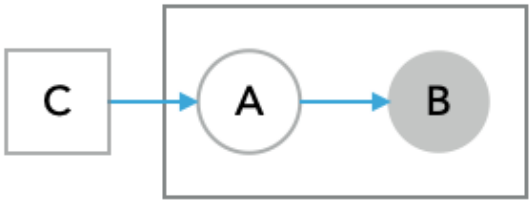


Figure 2.1: An example of a directed acyclic graph for the posterior in Eq. 2.2, where observed quantities are shaded, variable quantities are circles, and fixed quantities are squares. The larger box represents repetition over an index.

tribution parameterized by a hyperparameter C, the joint posterior $p(A, C|B)$ is expressed as:

$$p(A, C|B) \propto p(B|A, C)\, p(A|C)\, p(C)\,. \tag{2.2}$$

This factorization of probabilities is often represented in the form of a directed acyclic graph, as in Fig. 2.1. Additional information in the form of hyperparameters are expressed in terms of cascading dependencies of variable and fixed quantities. This hierarchical framework will be used in the next section to combine information from multiple galaxy populations in order to simultaneously infer redshift distributions of a sample of photometric galaxies, as well as provide estimates for individual redshifts in that sample.

2.4 Redshifts with colors and clustering

2.4.1 Galaxy populations

SB18 incorporate clustering, photometric and spectroscopic information in their hierarchical Bayesian model (HBM). The photometric sample (the observed galaxies whose redshift distributions we are trying to constrain) comes from survey photometry of single galaxies in the

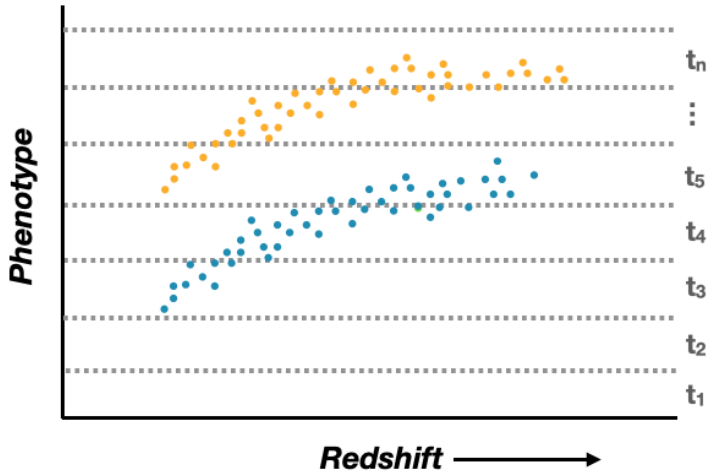


Figure 2.2: The variation of galaxies (circles) with similar genotypes, or rest-frame properties, with redshift. The two different genotypes are shown in yellow and blue, with phenotypes labeled as $t \in \{t_1 ... t_n\}$.

form of fluxes F_i and positions θ_i, with a known selection function s. The underlying density field of a well-characterized tracer population allows for the estimate of fluctuations in the density of sources for positions and redshifts spanning the survey to further constrain redshift distributions. This is similar to the method of clustering redshifts, which measures the angular cross-correlation between galaxies in two samples that occupy the same sky coverage, but that do not necessarily have similar galaxy properties (Newman 2008). The main requirement for galaxies from the tracer population is that they have well-established redshifts, either from spectroscopy, or from deep, multi-band photometry. Finally, a spectroscopic sample, which can also come from either spectroscopic observations or from deep photometry, provide galaxies with known redshifts to be used as an informative prior. Critically, the spectroscopic population should span the range of types of galaxies in the photometric sample. The characterization of these types is described in the following section.

2.4.2 Characterizing the color-redshift relation using phenotypes

As described in Section 2.2, photometric redshift methods exploit the relationship between observed features of galaxies with secure redshifts to make redshift estimates for galaxies obtained via imaging. This is done explicitly in template-fitting methods by fitting each galaxy to templates with different redshift-reddening combinations, and, implicitly, via machine learning methods by discovering the mapping between individual types, or combinations of observed features, and redshift (Leistedt et al. 2016).

In biology, an animal's phenotype is defined as the set of external characteristics that result from the interaction of its genotype, or genetic code, with its environment. Similarly, if a galaxy's inherent rest-frame features constitute its genotype, then its phenotype would be its observed appearance. Fig. 2.2 gives an example of how the evolution of galaxies of the same genotype can vary with respect to redshift. Although the galaxies shown in the figure come from two distinct genotypes, their observed appearances clearly span a range of phenotypes. Moreover, galaxies of the same phenotype can live at different redshifts. Most notably, there are galaxies from each genotype which, after being redshifted, fall into the same phenotype category, as in cell t_5. These galaxies could easily be mistaken as being redshifted by the same amount, despite their distinct rest-frame properties. Additional information, such as angular correlations between galaxy positions in the form of clustering, as well as spectroscopic information, are helpful in breaking these degeneracies.

Defining the color-redshift relation in terms of a set of discrete phenotypes can help to identify subtle differences between sets of features in order to better constrain redshift distributions. Specifically, because galaxies of the same type are assumed to have the same selection function $p(s|t,\theta)$, and the same probability $p(F,s|t,\theta)$ of being selected and measured to have observed features F, this essentially decouples the issue of measurement-noise from the color-redshift relation. When applying the SB18 method to data from cosmoDC2, galaxies will

be assigned phenotypes using a self-organizing map (Kohonen 1982). This technique will be discussed in further detail in Section 2.6.2.

Table 2.1: A list of the notation used in this method, following the notation in SB18.

F	set of galaxy features
t	galaxy phenotype
z	galaxy redshift
θ	galaxy angular position in the sky
$\boldsymbol{F, t, z, \theta}$	set of properties for all galaxies in the sample
s	galaxy selection outcome
$\delta_z(\theta)$	density fluctuation at redshift z and angular position θ
$\boldsymbol{\delta}$	$\{\delta_z(\theta)\}$ for all redshifts and positions
π_δ	hyperparameters of the density fluctuation field
b_z^t	linear galaxy bias for redshift z and type t
$\boldsymbol{b}$	$\{b_z^t\}$ for all redshifts and types
f_{zt}	joint redshift-type probability $p(z,t)$
$\boldsymbol{f}$	$\{f_{zt}\}$ for all types and redshifts
$\mathcal{L}_{it}$	probability of measuring galaxy i with F_i given t_i
n	mean density per unit solid angle
$n(z)$	mean density per unit solid angle per z
N_{zt}	number of sources assigned to redshift z and type t
$\boldsymbol{N}$	$\{N_{zt}\}$ for all redshifts and types
N_t	number of types
N_z	number of redshifts
M_{zt}	number of sources in the prior at redshift z and type t
$\boldsymbol{M}$	$\{M_{zt}\}$ for all redshifts and types
A	effective survey area
Δz	difference between the sample means and the true $n(z)$ mean

2.4.3 Redshift inference formalism

Observations of galaxies consist of a noisy set of individual galaxy features F_i identified by their phenotype t_i, redshift z_i, and angular position on the sky θ_i, where positions are viewed as being drawn from some latent density field δ_z that varies as a function of redshift z. The spatial distribution of galaxies, or galaxy bias b_z^t, varies linearly with respect to the fluctuations δ_z. Furthermore, the probability of a galaxy being selected is expressed in terms of a selection function $p(s|t,\theta)$. The vector quantities $\boldsymbol{F}$, $\boldsymbol{t}$, $\boldsymbol{z}$, and $\boldsymbol{\theta}$ denote the full set of properties of all selected galaxies. Both redshifts and phenotypes are represented as

indices identifying bins of the corresponding continuous quantities. A (z, t) pair for a galaxy therefore determines its position in a discrete 2D redshift-phenotype space. The number of sources assigned to redshift z and phenotype t is given by N_{zt}. The notation used in the remainder of the chapter is summarized in Table 2.1.

The HBM rests on the fundamental assumption that galaxies are drawn from a Cox process (Cox 1955), also known as a doubly stochastic Poisson process. Unlike a Poisson process, which assumes a fixed parameter λ, a Cox process generalizes this so that the number of events in disjoint intervals are not necessarily stochastically independent (Grandell 1976). This is relevant to the case of how we will sample galaxies since the rate or concentration of galaxies depends on fluctuations in the density field.

Let us assume for now that the mean density per unit solid angle of galaxies in the sky is n, and that all galaxies have identical feature vectors such that F is constant. Moreover, our galaxy sample has an intrinsic redshift distribution where $f_z = p(z)$ with $\sum_z f_z = 1$, and linear bias b_z with respect to the underlying density fluctuation δ_z. The density field from which these galaxies are sampled from is:

$$\rho(z, \theta | n, \boldsymbol{f}, \boldsymbol{b}, \boldsymbol{\delta}) = n f_z \left[1 + b_z \delta_z(\theta)\right] p(s|\theta) \,. \tag{2.3}$$

The first term on the right gives the number of galaxies that are expected to be found at redshift z, while the last two terms describe how the expected detection rate varies spatially due to both density fluctuations and observing conditions. The expected number of detections, expressed as the integral of Eq. 2.3 over all angular positions, is the variable parameter in our Cox process. In accordance with standard Poisson sampling, the likelihood of drawing a set of galaxies at redshifts $\boldsymbol{z}$ and positions $\boldsymbol{\theta}$ becomes:

$$
\begin{aligned}
p(\boldsymbol{z}, \boldsymbol{\theta}|n, \boldsymbol{f}, \boldsymbol{b}, \boldsymbol{\delta}) &= \exp\Big\{ -\int \mathrm{d}^2\theta \sum_z \rho(z,\theta)\Big\} \prod_{t=1}^{N} \rho(z_i, \theta_i) \\
&= \exp\Big\{ -n \int \mathrm{d}^2\theta \, p(s|\theta) \sum_z f_z \left[1 + b_z \delta_z(\theta)\right]\Big\} \\
&\times \prod_{i=1}^{N} p(s|\theta_i) \;\times\; n\, f_{z_i} \left[1 + b_{z_i} \delta_{z_i}(\theta_i)\right] .
\end{aligned}
\tag{2.4}
$$

The discrete sum over redshifts is due to the fact that each galaxy's redshift is transformed to an integer corresponding to a specific finite-width redshift bin. There are also a finite set of phenotypes indexed by integer t, which identify the type bin a galaxy belongs to. Eq. 2.3 can be generalized to include discrete phenotypes:

$$
p(z, \theta, t|n, \boldsymbol{f}, \boldsymbol{b}, \boldsymbol{\delta}) = n f_{tz} \left[1 + b_z^t \delta_z(\theta)\right] p(s|t, \theta) \, . \tag{2.5}
$$

The joint type and redshift probability $f_{zt} = p(z, t)$ is given by:

$$
f_{zt} \equiv p(z,t) = p(z|t)p(t) = f_z^t f_t \, , \tag{2.6}
$$

where $f_t = p(t)$ is subject to the constraint $\sum_t f_t = 1$. We can also define a feature-selection likelihood of a galaxy i with phenotype t and angular position θ:

$$
\mathcal{L}_{it} \equiv p(F_i, s|t_i, \theta_i) \, . \tag{2.7}
$$

Note that this likelihood is independent of redshift, since galaxies with identical phenotypes could live at distinct redshifts (see Fig. 2.2). The quality of observations and selection

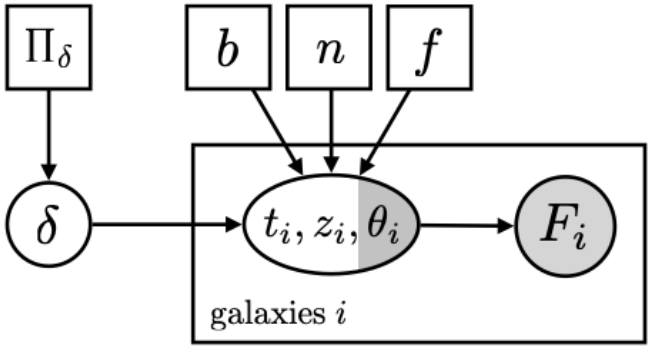


Figure 2.3: A directed acyclic graph for the hierarchical Bayesian model. The observed quantities (individual positions θ_i and fluxes F_i) are shaded. Squares indicate model parameters, while circles correspond to stochastic quantities.

algorithms are assumed to be known *a priori* for both the toy model and the application to the cosmoDC2 catalog. Eq. 2.4 can now be recast to include types and features:

$$\begin{aligned} p(\boldsymbol{F}, \boldsymbol{\theta}, \boldsymbol{t}, \boldsymbol{z} | n, \boldsymbol{f}, \boldsymbol{b}, \boldsymbol{\delta}) = & \exp\left[- n \sum_t f_t A^t(\boldsymbol{f^t}, \boldsymbol{b^t}, \boldsymbol{\delta}) \right] \\ & \times \prod_i \mathcal{L}_{it} n f_{t_i z_i} \left[1 + b^t_{z_i i} \delta_{z_i}(\theta_i)\right], \end{aligned} \tag{2.8}$$

where A^t is defined as the effective survey area for each phenotype:

$$\begin{aligned} A^t(\boldsymbol{f}, \boldsymbol{b}, \boldsymbol{\delta}) &\equiv \sum_z \int \mathrm{d}^2\theta \, p(s|t,\theta) f^t_z \left[1 + b^t_z \delta_z(\theta)\right] \\ &= \int \mathrm{d}^2\theta \, p(s|t,\theta) \left[1 + \sum_z f^t_z b^t_z \delta_z(\theta)\right] \\ &\approx \int \mathrm{d}^2\theta \, p(s|t,\theta) \, . \end{aligned} \tag{2.9}$$

The approximation in the last line assumes that the density of fluctuations over the survey area is small, making $\{A^t\}$ known constants of the survey. For the remainder of the studies in this chapter, A^t is assumed to be independent of type and therefore constant for *all* galaxies. This will simplify the joint posterior distribution for redshifts and types.

2.4.4 Sampling from the posterior

The goal of the HBM is to estimate individual redshifts and types, as well as the joint redshift-type probabilites f_{zt}, which are key to determining the shape of the underlying redshfit distribution:

$$n(z) = n \sum_t f_{zt} \,. \tag{2.10}$$

We can invoke Bayes' theorem to sample these individual and joint probabilities from the following posterior distribution, expressed in terms of the likelihood function in Eq. 2.8:

$$\begin{aligned} p(\boldsymbol{f}, \boldsymbol{z}, \boldsymbol{b}, \boldsymbol{t} | \boldsymbol{F}, \boldsymbol{\theta}, \pi_\delta) \propto \int \mathrm{d}n \, \mathrm{d}\delta \\ p(\boldsymbol{F}, \boldsymbol{\theta}, \boldsymbol{t}, \boldsymbol{z} | n, \boldsymbol{f}, \boldsymbol{b}, \boldsymbol{\delta}) \, p(\boldsymbol{\delta} | \pi_\delta) \, p(n) \, p(\boldsymbol{f}) \, p(\boldsymbol{b}) \,. \end{aligned} \tag{2.11}$$

Fig. 2.3 shows the directed acyclic graph for this distribution. The hyperparameters π_δ describe the stochastic density fluctuations. Rather than sampling from $p(\theta|\pi_\delta)$, we will replace $\delta_z(\theta)$ with some deterministic estimator $\hat{\delta}_z(\theta)$. This gives a single realization of the density field of the tracer population, transforming $p(\boldsymbol{\delta}|\pi_\delta)$ into a Dirac delta function where $\delta_z = \hat{\delta}_z$. In reality, the density fields remain stochastic even after specification of the tracer population but, since we are using simulated data, we already have prior knowledge of the

true density fields. After marginalizing over the overall density n, and assuming a logarithmic Jeffreys prior on $p(n)$, the full, joint posterior distribution for redshifts and types becomes:

$$p(\boldsymbol{f}, \boldsymbol{z}, \boldsymbol{t}, \boldsymbol{b}|\boldsymbol{F}, \boldsymbol{\theta}) \propto p(\boldsymbol{f}) \prod_i \mathcal{L}_{it_i} f_{t_i z_i} \left[1 + b_{z_i}^{t_i} \hat{\delta}_{z_i}(\theta_i)\right]. \tag{2.12}$$

Sampling every variable in Eq. 2.12 at once is computationally intractable. However, the fact that individual components of the posterior can be sampled separately means that Eq. 2.12 is amenable to Gibbs sampling. Because the posterior has been simplified by assuming that the effective survey area for source detection is independent of type ($A(\boldsymbol{f})$ is constant), and that the tracer and target populations have the same density fluctuations (the biases $b_z^t = 1$), the Gibbs sampling can be performed in two steps. The first step involves sampling from the joint redshift-type probability distribution, conditioned on the observables $\boldsymbol{F}$, $\boldsymbol{\theta}$, prior information $p(\boldsymbol{f})$, and counts N_{zt} of the number of sources belonging to a distinct redshift and type bin from the previous iteration of the sampler:

$$p(\boldsymbol{f}|\boldsymbol{z}, \boldsymbol{t}, \boldsymbol{F}, \boldsymbol{\theta}) \propto p(\boldsymbol{f}) \prod_{z,t} f_{zt}^{N_{zt}}. \tag{2.13}$$

The individual types and redshifts for each galaxy are sampled in the second step, conditioned on the joint type-redshift probabilities $\boldsymbol{f}$ from Eq. 2.13:

$$p(z_i, t_i|\boldsymbol{f}, F_i, \theta_i) \propto \mathcal{L}_{it_i} f_{t_i z_i} \left(1 + \hat{\delta}_i z_i\right). \tag{2.14}$$

The z, t pairs for each galaxy comprise the next realization of counts N_{zt} in the next iteration of the sampler. This process is performed until the type and redshift distributions have converged (the condition for convergence will be described in detail in Section 2.5.2).

2.4.5 Incorporating the prior

The prior $p(\boldsymbol{f})$ is a fixed quantity – the probabilities of each redshift-type combination in the spectroscopic sample are only computed once and then used in subsequent iterations of the sampler. Therefore, after the first iteration, the conditional posterior in Eq. 2.13 only depends on the number of counts N_{zt} from the second step of the Gibbs sampler. This dependency can be expressed in a straightforward way using Bayes' theorem:

$$p(\boldsymbol{f}|\boldsymbol{N}) \propto p(\boldsymbol{N}|\boldsymbol{f})\, p(\boldsymbol{f})\,. \tag{2.15}$$

The likelihood $p(\boldsymbol{N}|\boldsymbol{f})$ is a single realization of probabilities for a discrete set of classifications (in this case, unique combinations of discrete type and redshift bins), and therefore follows a categorical distribution (Wikipedia contributors 2021). Since the Dirichlet distribution is the conjugate prior to the categorical distribution, we can choose the prior to be Dirichlet distributed (which appropriately satisfies $0 \leq f_{zt} \leq 1$ and $\sum_{zt} f_{zt} = 1$), so that the posterior will also follow a Dirichlet distribution. This is commonly done in Bayesian statistics, and has the advantage of easily allowing the incorporation of new counts in each cycle. The Dirichlet distribution has the following probability density function:

$$f(x_1, ..., x_K; \alpha_1, ...\alpha_K) = \frac{1}{B(\alpha)} \prod_{i=1}^{K} x_i^{\alpha_i - 1}\,, \tag{2.16}$$

where K is the number of parameters, $B(\alpha)$ is the beta function, α_i are concentration parameters whose domain spans the positive reals, and x_i fulfill the standard normalization requirement for probabilities $0 \leq x_i \leq 1$ and $\sum_i x_i = 1$. In our model, the $\{x_i\}$ are the joint redshift-type probabilities f_{zt}, the number of parameters are the number of unique type-redshift bins, and the concentration parameters are the number of counts in each (z, t) bin.

Since we are using an informative prior where we include information from a representative spectroscopic sample, the posterior in Eq. 2.15 is Dirichlet-distributed on the number of counts in the prior sample $\boldsymbol{M} = \{M_{zt}\}$ and the counts from the last iteration $\boldsymbol{N} = \{N_{zt}\}$, with the concentration offset parameter $\alpha_{\text{offset}} = 1$:

$$\begin{aligned} p(\boldsymbol{f}|\boldsymbol{N}) &\propto \prod_{z,t}^{N_{zt}} f_{zt}^{(N_{zt}+M_{zt}+\alpha_{\text{offset}}-1)} \\ &\sim \text{Dir}(\boldsymbol{N}+\boldsymbol{M})\,. \end{aligned} \tag{2.17}$$

2.5 Demonstration on a toy model

The formalism outlined in Section 2.4.3 is applied to a toy model as a proof of concept. The following two sections detail how these simplified data were generated and show the redshift and type estimates of the HBM. A metric is introduced to study redshift estimation errors, and the effect of clustering on the performance of the method is also investigated.

2.5.1 Generating the toy data

Redshifts and types in the toy model are constrained to lie in the range $z, t \in [0, 1]$ in 50 equally spaced bins, such that $N_z = N_t = 50$. The density fluctuation field δ_z in each of these redshift bins is generated from a deterministic density estimator using a Gaussian Random Field (GRF) with a resolution of 1024 x 1024 pixels. The posterior will therefore not include variance due to shot noise in the tracer population. Each redshift bin has a density fluctuation field $\delta_z(\theta)$ that is independent of any other bin such that $\langle \delta_{z_i}(\theta)\delta_{z_j}(\theta)\rangle = 0$ for $z_i \neq z_j$. Moreover, it's assumed that the biases $b_z^t = 1$ so that targets and tracers have the same fluctuations. Perfect selection is also assumed ($p(s) = 1$). The field in each bin is

drawn from a power spectrum $P(k) \sim k^{-3}$, where k is the scale in Fourier space, with an $rms = 2.5$, and clipped to only include non-negative values.

The overdensity field is defined in terms of the value of the field at each position and redshift $\rho(\theta, z)$:

$$\delta(\theta, z) = \frac{\rho(\theta, z)}{\bar{\rho}(z)} - 1\,. \tag{2.18}$$

Individual galaxy types are drawn from the following distribution:

$$p(t) \propto t^a \exp -(t/t_0)^2\,, \tag{2.19}$$

with $a = 1.5$ and $t_0 = 0.3$. Redshifts are assigned piecewise as a function of type:

$$p(z|t) = \begin{cases} \text{if } t = 0 & \begin{cases} 0.8 & z = t \\ 0.2 & z = t + 0.02 \end{cases} \\ \text{if } t = 1 & \begin{cases} 0.8 & z = t \\ 0.2 & z = t - 0.02 \end{cases} \\ \text{else} & \begin{cases} 0.6 & z = t \\ 0.2 & z = t \pm 0.02 \end{cases} \end{cases} \tag{2.20}$$

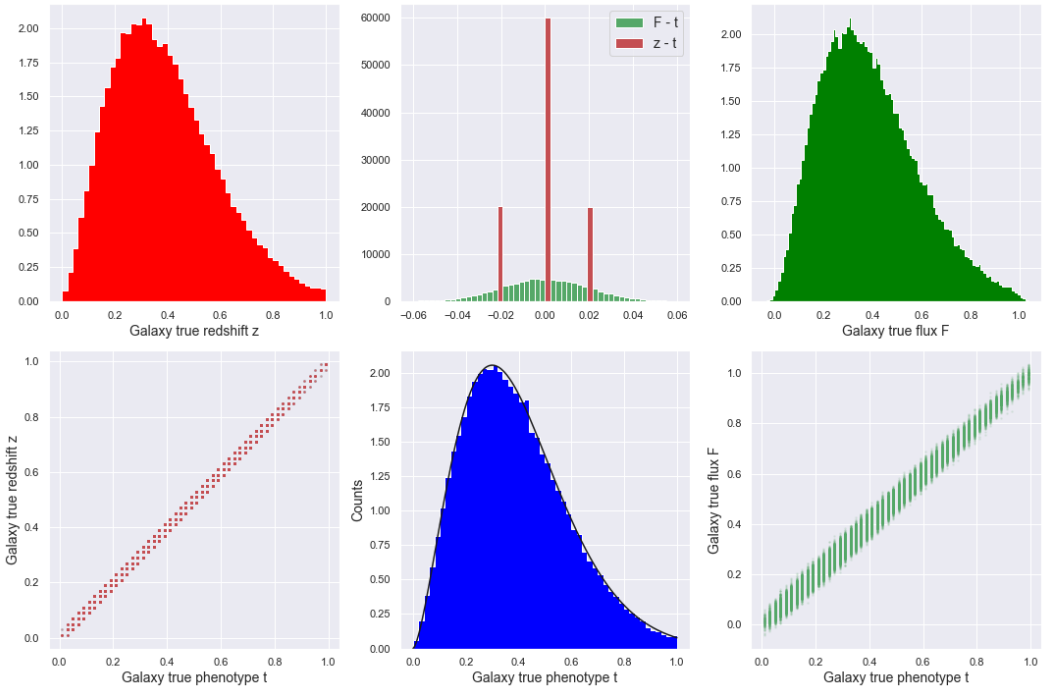


Figure 2.4: Distributions of the true redshifts, phenotypes, and fluxes of the 1×10^5 galaxies used in the toy model.

Each galaxy is also assigned a single feature, or flux, based on its type. This flux value is randomly selected from a 1D Gaussian with variance $\sigma_F^2 = 0.02$ to simulate flux measurement errors:

$$p(F|t) = \mathcal{L}_{Ft} = \mathcal{N}(t - F, \sigma_F^2) \tag{2.21}$$

Lastly, line-of-sight positions are assigned by randomly sampling from the overdensity field in Eq. 2.18:

$$p(\theta|z) \propto (1 + \delta(\theta, z)) \,. \tag{2.22}$$

Distributions of the redshifts, types, and features of 1×10^5 simulated galaxies are shown in Fig. 2.4, and the resulting density field in all 50 redshift bins is shown in Fig. 2.5, where the

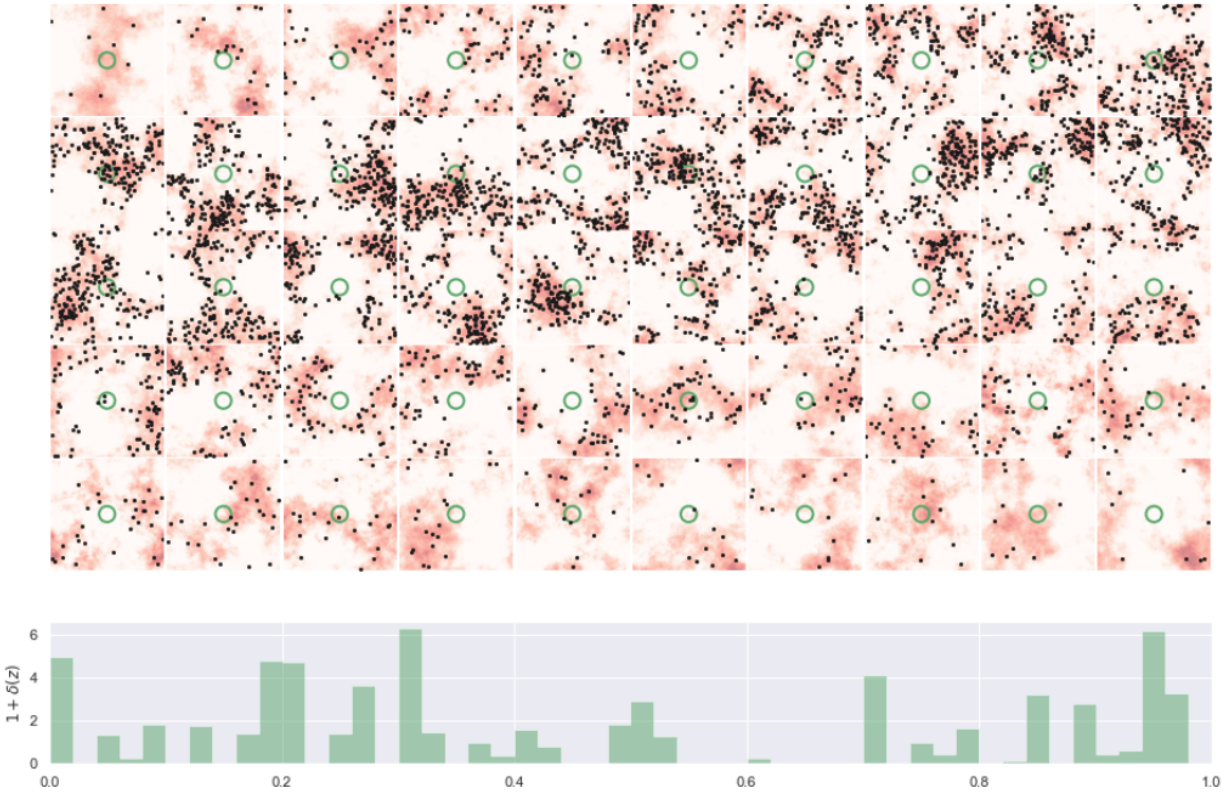


Figure 2.5: Angular density fluctuation field in each of the 50 redshift slices in the toy model (top), and a histogram of the value of the density field in the green circle in each slice (bottom).

clustering is in red, and a random subset of 5% of the data in each slice is shown in black. The histogram below shows the value of the field $\rho(\theta, z)$ in the center pixel of each 1024 x 1024 pixel redshift slice, indicated by the green circle.

The toy data set consists of two distinct subsets of the 1×10^5 galaxies simulated above: an $N_{phot} = 8 \times 10^4$ photometric sample and an $N_{spec} = 1 \times 10^3$ spectroscopic sample.

2.5.2 Convergence of the Gibbs chains

The Gibbs sampler was run for four different posterior conditions, including running on just the prior, on the prior and the feature likelihood, the prior, feature likelihood, and clustering

Table 2.2: Gelman & Rubin convergence results for each of the four different posteriors in the toy model.

Posterior condition	Description of posterior condition	R - 1
`prior`	Prior only	0.0003
`noz`	Prior + Feature likelihood	0.0117
`halfz`	Prior + Feature likelihood + $\delta(z \leq 0.5)$	0.0146
`fullz`	Prior + Feature likelihood + δ(All z)	0.0128

up to $z \leq 0.5$, and on the full posterior including clustering up to $z \leq 1.2$ (see Table 2.2). Four different walkers were run for each condition, initialized by different seeds, where each walker ran 2500 cycles of the Gibbs sampler. This resulted in a total of 1×10^4 samples per chain.

A simple metric was derived to assess the performance of the model for each sampling scheme. This involved computing the difference between the estimated redshift in each sample j and the mean of the true redshift distribution of the photometric population:

$$\Delta z_j = \langle z_{\mathrm{est},j} \rangle - \langle z_{\mathrm{true}} \rangle \,. \tag{2.23}$$

Convergence of the chains was determined using the Gelman & Rubin convergence metric (Gelman & Rubin 1992):

$$R = \sqrt{\hat{V}/W} \,, \tag{2.24}$$

where the variance estimate is:

$$\hat{V} = \frac{n-1}{n} W + \frac{1}{n} B \,. \tag{2.25}$$

W and B/n are the average of the within-sample variances and the variance between the sample means, respectively, and n is the number of samples in each chain. SB18 requires that the Gelman & Rubin metric $R - 1 < 0.03$ for the chain to have converged. Table 2.2 shows that this was satisfied for all four posterior conditions.

2.5.3 Results of the toy model

The following results are for the full posterior (`fullz`). A histogram of inferred redshifts for four different galaxies over all 1×10^4 samples is shown in Fig. 2.6. The true redshift in each histogram is shown as the dotted black line. In each case, the true redshift is closely aligned with the peak of the redshift distribution. The mean redshifts of all galaxies in each iteration of the sampler were also computed and are shown in Fig. 2.7, along with the true population mean. These results also show good agreement overall between true and sampled means.

Next, redshift and type distributions were examined separately for each posterior condition. Fig. 2.8 shows estimates of redshift and type in each bin, along with a histogram of the true distributions from the photometric sample. It is clear how the incorporation of additional information such as spectroscopic redshifts and clustering help constrain the posterior distribution, particularly for redshift estimates, resulting in tighter distributions. In the third panel of Fig. 2.8, we can see the impact of only including clustering information at redshifts below $z < 0.5$. Both redshift and type distributions for lower redshifts are more tightly constrained compared to galaxies above $z > 0.5$, which reduce to the photometry-only case in the second panel.

Next, the Δz metric from Eq. 2.23 was computed for each posterior condition. A histogram of Δz values for all four chains is shown in the left panel of Fig. 2.9. The median and standard deviation of all samples for each posterior type are given in Table 2.3.

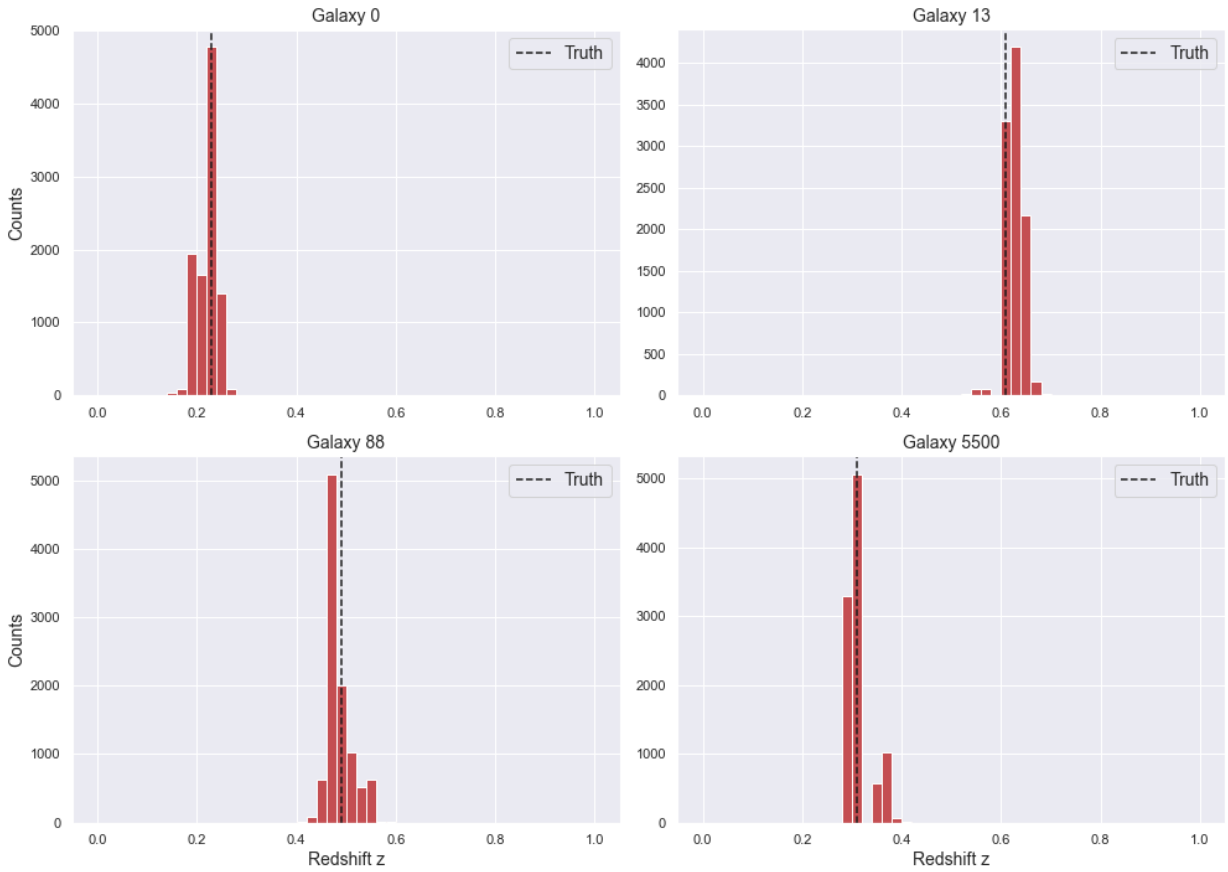


Figure 2.6: Histograms of estimated redshifts for four galaxies in the photometric sample.

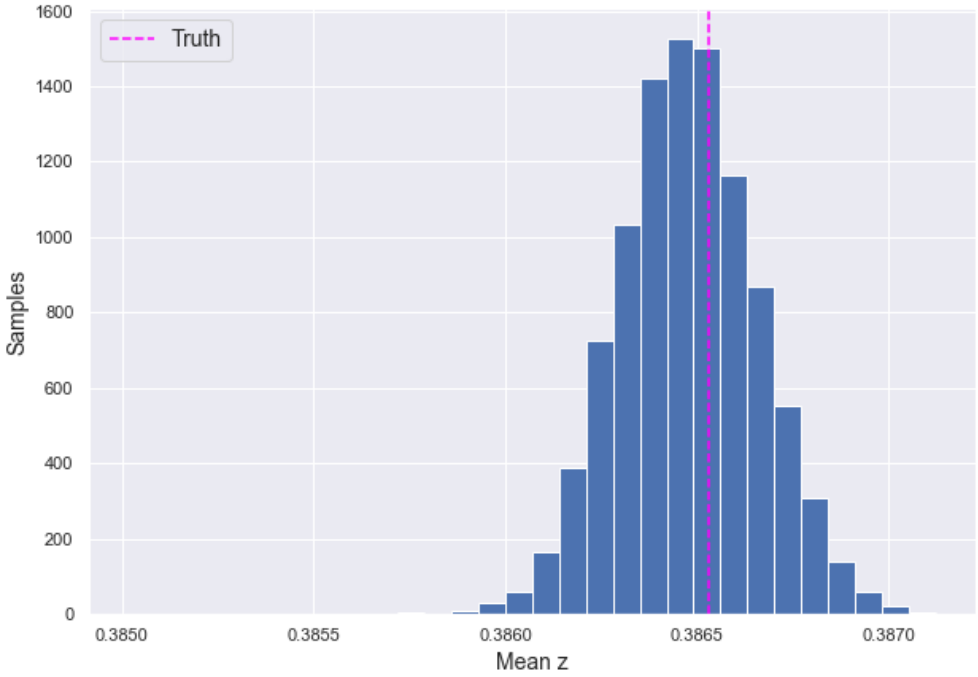


Figure 2.7: Estimated population mean redshifts for samples with the full posterior.

Table 2.3: Median and standard deviation of Δz for each fiducial posterior.

Posterior condition	Median (Δz)	Standard Deviation (Δz)
`prior`	0.00653	0.00639
`noz`	-0.00091	0.00045
`halfz`	-0.00139	0.00024
`fullz`	-0.00006	0.00018

The offsets in `noz` and `halfz` were the results of systematic biases in each chain. These biases were confirmed to be due to the prior being slightly non-representative of the full sample due to statistical fluctuations. The right panel of Fig. 2.9 shows the same metric, but instead of using a prior of 1×10^3 galaxies, the chains are run with a prior sample of 2×10^4 galaxies.

Restricting the clustering portion of the likelihood to $z < 0.5$ shows a slight improvement in the metric for low-redshift galaxies, where it effectively traces the full posterior histogram, as expected. Fig. 2.11 demonstrates this for a single galaxy. However, using the `halfz`

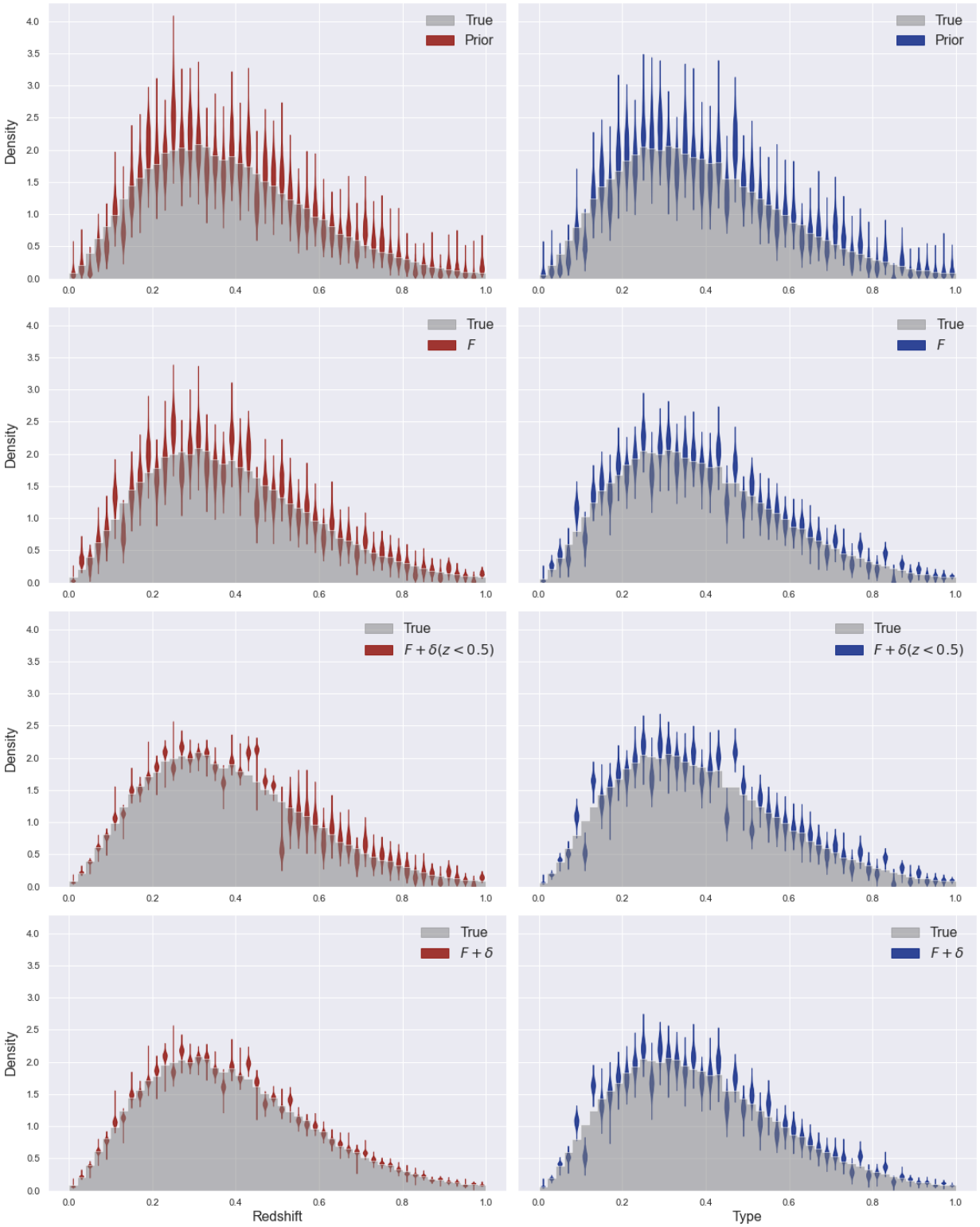


Figure 2.8: Results of Gibbs sampling from four different posterior conditions (see Table 2.2) shown as violin plots. Estimates for redshifts and types are shown in red and blue, respectively. Including clustering information significantly reduces the scatter in each bin, and reduces to only using photometric and prior information in the case of no clustering.

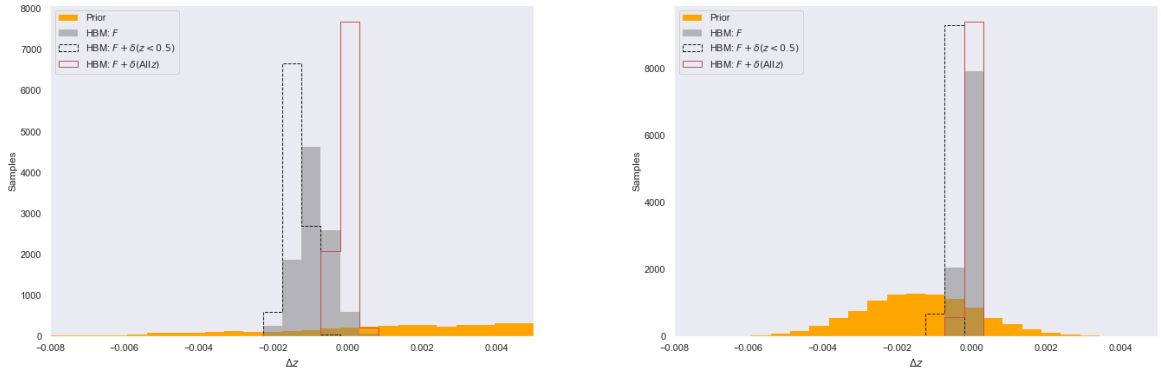


Figure 2.9: The Δz metric for all four chains for the case of using (a) 1×10^3 spectroscopic galaxies and (b) 2×10^4 spectroscopic galaxies for the prior. Using a larger spectroscopic sample reduces the systematic biases in `noz` and `halfz`.

posterior to make redshift estimates for high-redshift galaxies degrades the accuracy to be comparable to using the feature-only posterior, as seen in Fig. 2.12.

2.6 Redshifts using simulated data from DC2

SB18 discuss the application of their method to simulated data in their subsequent paper [A19] (Alarcon et al. 2020). This section follows the approach outlined in A19 to estimate redshifts on realistic simulations provided by the cosmoDC2 catalog[1], which is a synthetic version of a small subset of the LSST survey generated as part of DC2. A self-organizing map (SOM) is used to build the feature likelihood by treating cells in the map as distinct phenotypes. Finally, we will compare the constraining power of using of a SOM to classify galaxy phenotypes to a method based on using normalizing flows (Papamakarios et al. 2021).

[1]The full catalog can be accessed here: `https://github.com/LSSTDESC/cosmodc2`

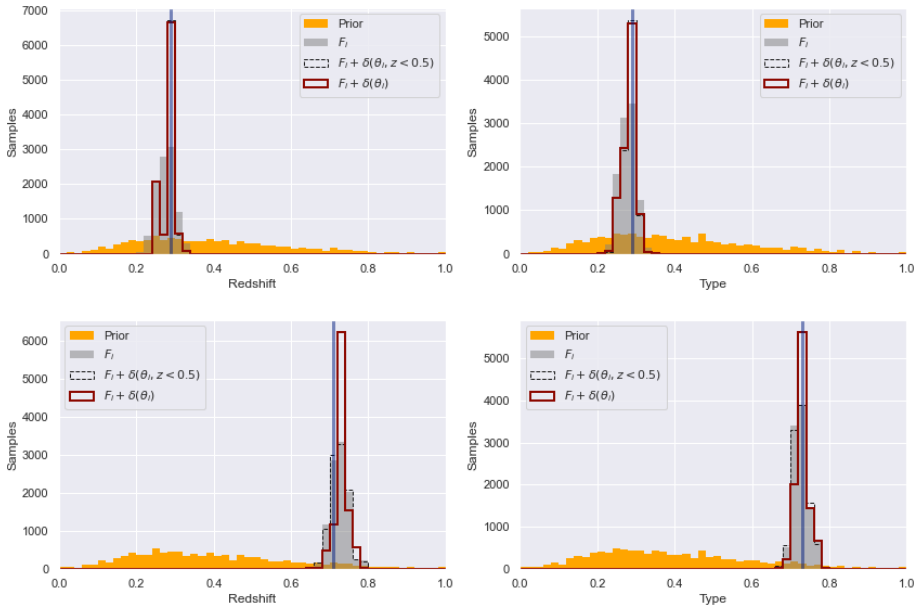


Figure 2.10: Redshift and type distributions for the four posterior conditions for a (a) low-z and a (b) high-z galaxy. The `halfz` and `fullz` in (a) are identical, consistent with the fact that high-z clustering information should have no effect on galaxies at redshifts below $z < 0.5$. The `halfz` estimates reduce to the `noz` case for the high-z galaxy in (b), as clustering information below $z < 0.5$ is not expected to have an impact for this redshift regime.

$50 < \text{RA} < 60.5$
$-44.5 < \text{DEC} < -36.5$
i-band < 25
$0.2 \leq z \leq 1.2$

Table 2.4: `cosmoDC2` catalog cuts.

2.6.1 The cosmoDC2 catalog

The cosmoDC2 catalog is based on a trillion-particle cosmological N-body simulation spanning a 440 deg^2 area on the sky up to redshift $z \leq 3$, with a limiting r-band magnitude depth of 28 (LSST Dark Energy Science Collaboration 2021). The simulations are based on semi-analytic models parameterized by underlying baryonic processes that model the evolution of galaxies within individual dark matter halos.

We use a version of this catalog created by Sam Schmidt for our photometric redshift studies, which is based on the extragalactic mock catalog `cosmoDC2_v1.1.4_image`. This catalog

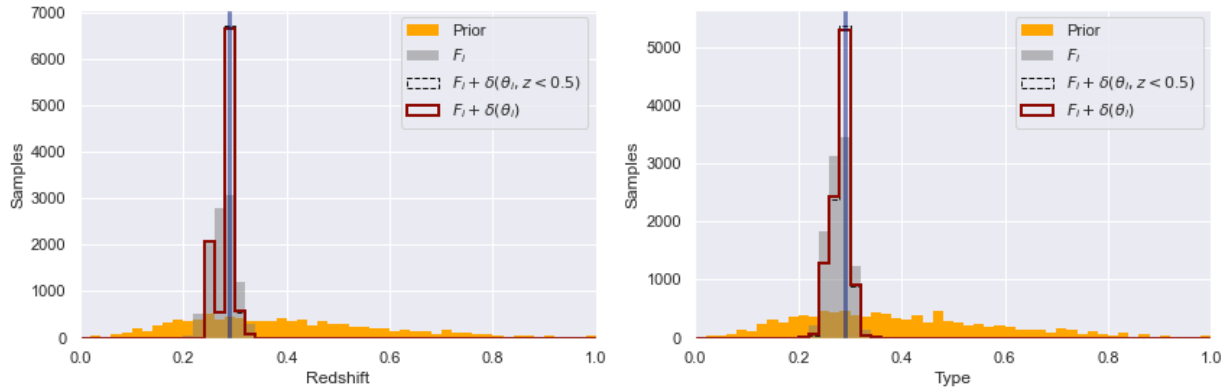


Figure 2.11: Redshift and type estimates for a low redshift galaxy, where there is only a small difference between using the `halfz` and `fullz` posteriors.

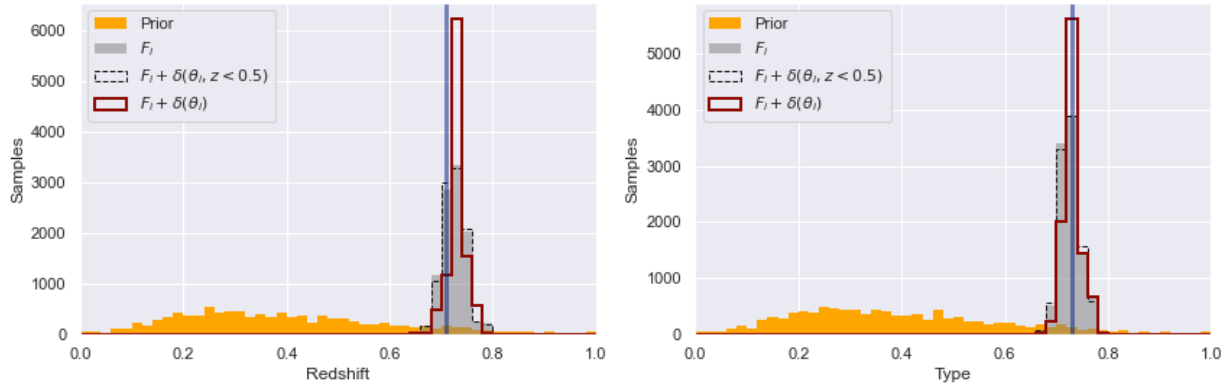


Figure 2.12: Redshift and type estimates for a high redshift galaxy, where estimates from the `fullz` posterior come closest to the true redshift.

contains observed magnitudes with predicted 10-year errors added to the fluxes according to the model presented in Ivezić et al. (2019). We apply magnitude and RA/DEC cuts to extract an 8×8 deg^2 patch in the South Galactic cap, with a limiting magnitude matching that of the LSST gold sample ($i < 25$) (LSST Science Collaboration and LSST Project 2009). The RA range is slightly wider than the DEC range to account for the cos(DEC) correction when measuring the distance between two points on a sphere.[2] Redshifts have been restricted to lie between $0.2 \leq z \leq 1.2$. A summary of the cuts applied to this catalog

[2]The RA range is multiplied by the cosine of the mean declination for the cut considered: $10.5° \times cos(40.5°) = 8°$.

is given in Table 2.4. Details on the specific galaxy samples used in this study, and how they were prepared, will be discussed in Section 2.7.

2.6.2 Defining phenotypes using self-organizing maps

One of the primary differences between the application of the hierarchical model to toy data in SB18 and to realistic simulations in A19 is the use of a self-organizing map (SOM) to define galaxy phenotypes. We will first describe the formalism behind this machine learning technique and then discuss its application to data from cosmoDC2 in the following section.

A SOM, also called a Kohonen map (Kohonen 1982), is an unsupervised machine learning method used to map high-dimensional data onto a low-dimensional (typically 2D) grid. It does this by mapping the input features to its cells through a competitive learning process, whereby cells rival each other to most closely resemble each input. One useful property of SOMs is that they preserve the topological properties of the data, so that neighboring cells have similar features. This makes it easier to visualize correlations that exist in the data, as well as identify gaps or anomalies in the feature space. For example, SOMs have been used to characterize the mapping between color and redshift in order to assess spectroscopic coverage in large surveys by locating parts of the grid that are under-populated (Masters et al. 2015). The SOM algorithm outlined below follows that of Buchs et al. (2019), where it was used for a similar purpose of estimating redshifts via phenotypes.

The SOM algorithm

Suppose that the training sample consists of n galaxies, where each galaxy is characterized by m features consisting of magnitudes or colors, but excluding redshift. The input feature vector associated with each galaxy is given by $\boldsymbol{x} \in \mathbb{R}^m$. The C cells in the SOM are arranged

on a p x q grid, and periodic boundaries are enforced so that cells along the edges of the map are considered neighbors of cells on the opposite edge. Each cell is associated with a weight vector $\boldsymbol{w}_k \in \mathbb{R}^m$, where $k = 1, ..., C$, occupying the same feature space as the training data. The weight vectors are initialized randomly through a Gaussian sampling of the input features. Galaxies in the training sample are randomly shuffled, and then introduced to the SOM one at a time through an iterative process. A single training cycle is completed after every galaxy has been processed once. The maximum number of training cycles is $t_{\max}$.

At each iteration, the Euclidean distance between the input vector and the weight vectors of all the cells is calculated :

$$d(\boldsymbol{x}, \boldsymbol{w}_k) = \sqrt{\sum_{k=1}^{C} (\boldsymbol{x} - \boldsymbol{w}_k)^2} \,. \tag{2.26}$$

This is used to determine the best matching unit (BMU), c_b, or the cell whose weight vector is closest to the input vector in feature space:

$$c_b = \underset{k}{\operatorname{argmin}}\ d(\boldsymbol{x}, \boldsymbol{w}_k) \,. \tag{2.27}$$

Next, cell weights are adjusted to more closely resemble the input vector. Rather than update the weights of all cells on the grid after the introduction of each galaxy, only the weight vectors of the cells whose physical location on the map is within a certain distance of the BMU are modified to more closely resemble the features of the input galaxy. This is what helps to preserve the topological structure of the input space. The update function is also a function of the time step t in the training:

$$\boldsymbol{w}_k(t+1) = \boldsymbol{w}_k(t) + \eta(t)\, H_{b,k}(t)\, [\boldsymbol{x}(t) - \boldsymbol{w}_k(t)] \,. \tag{2.28}$$

The learning rate $\eta(t)$ is a monotonically decreasing function of the time step that determines the responsiveness of the map to new inputs, and becomes less sensitive as the number of iterations increases:

$$\eta(t) = \eta_0^{t/t_{\max}} . \tag{2.29}$$

The constant $\eta_0 \in [0, 1]$, and for our study we use $\eta_0 = 0.5$. The weight vectors that are selected to be updated lie within a neighborhood of the BMU. This area is determined by the neighborhood function $H_{b,k}(t)$, which is a Gaussian kernel centered on the BMU that also decreases over time:

$$H_{b,k}(t) = \exp[-D_{b,k}^2/\sigma^2(t)] . \tag{2.30}$$

The width of the kernel is given by:

$$\sigma(t) = \sigma_s^{1-t/t_{\max}} . \tag{2.31}$$

The initial value of the kernel σ_s is chosen to be equal to the mean of the physical separations between each cell and every other cell on the map. The distance between the BMU, c_b, and cell c_k is just the Euclidean distance:

$$D_{b,k}^2 = \sum_i^p \sum_j^q (c_{b_{i,j}} - c_{k_{i,j}})^2 . \tag{2.32}$$

As the training progresses and new galaxies are introduced to the map, the width of the neighborhood function shrinks so that, toward the end of each training cycle, it only includes the BMU and its closest neighbors.

The loss function and convergence

Termination of the training process is based on the value of a loss function, which is a measure of the "cost" of each successive cycle. The loss is defined as the sum of the smallest feature-space distances over all iterations in a given cycle. In other words, it is the sum of the distances in Eq. 2.26 between the input vectors and their associated BMUs in each cycle:

$$L = \sum_{t=1}^{t_{\max}} d(\boldsymbol{x}_t, \boldsymbol{w}_{c_b}) \,. \tag{2.33}$$

Training is said to have converged when the loss function has plateaued, such that additional training cycles no longer result in a significant decrease in the loss. At this point, the map has sufficiently defined phenotypic cells as combinations of galaxy features and can be used to classify new data. This mapping of new galaxies to cells is done by finding the BMU of each galaxy (using Eq. 2.27), which is the cell that minimizes the Euclidean distance between a given galaxy's features and the weights of all the cells.

2.6.3 Refining phenotypes using two SOMs

The toy model made use of three separate galaxy populations: a spectroscopic sample with accurate redshifts for the prior distribution, a tracer population containing galaxy clustering information, and a photometric sample of observed galaxies whose redshifts we were trying

to determine. As we move away from the toy model and develop a more sophisticated way of identifying galaxy phenotypes, we will need to introduce two additional galaxy populations.

The feature likelihood in Eq. 2.7 expresses the probability of observing galaxy i to have features F_i given that it belongs to phenotype t_i. We can construct a single SOM using the features in our photometric sample, but this will introduce photometric scatter into the map due to measured errors in the observed features. Instead, we will train two SOMs: one on the observed photometry, or the *wide* sample, and another on more accurate, multi-band photometry that is complete to faint magnitudes, hereafter referred to as the *deep* sample. The deep sample is assumed to be representative of the wide data, such that any galaxy observed in the wide data would be observed if its location were in the deep field. An *overlap* sample, defined as galaxies for which both wide and deep observations are available, will be used to generate a mapping from the wide SOM to the deep SOM. This essentially discretizes the likelihood function by expressing it in terms of discrete cells rather than continuous feature values.

2.6.4 Constructing the feature likelihood

The assumption that every observed galaxy in the wide sample would be observed in the deep sample if it were located within the deep footprint means that every wide galaxy can be mapped to a location in the deep feature space with some likelihood. This mapping is performed by training one SOM on the wide sample and another SOM on the deep sample, and finding a mapping from one to the other. Since the weight vectors of the wide cells will be subject to photometric noise, mapping the wide cells to a second, higher-resolution SOM trained on noiseless features will refine the classification of galaxy phenotypes, not only because there is lower variance in the cell assignments, but because discretizing galaxy features over a finer grid allows us to better resolve distinct type categories.

The feature likelihood in Eq. 2.7 can be redefined in terms of wide and deep SOM cells, $\hat{c}$ and c such that

$$\mathcal{L}_{it} \equiv p(F_i, s|t_i, \theta_i) = p(\hat{c}_i, s|c_i, \theta_i)\,. \tag{2.34}$$

In other words, given a galaxy's assignment in the deep SOM, c_i, we can determine the likelihood of it having the observed features in its wide-cell assignment, $\hat{c}_i$. Per Bayes' theorem, the likelihood in SOM parlance becomes:

$$p(\hat{c}, s|c) \propto p(c|\hat{c}, s)\, p(\hat{c}|s)\,. \tag{2.35}$$

The quantity $p(c|\hat{c}, s)$ is called the transfer function, and is the fractional occupation of a deep cell given a wide cell after all galaxies in the overlap sample have been mapped to both SOMs. The transfer function can be recast using the law of conditional probabilities:

$$p(c|\hat{c}, s) = \frac{p(c, \hat{c}|s)}{p(\hat{c}|s)}\,. \tag{2.36}$$

The numerator is the probability of a galaxy in the overlap sample belonging to the cell combination $(c, \hat{c})$, based on its BMU assignment in each SOM:

$$p(c, \hat{c}|s) = \frac{1}{n_s} \sum_{i \in s} \delta_{c,c_i} \delta_{\hat{c},\hat{c}_i}\,, \tag{2.37}$$

where n_s is the number of galaxies that were selected as part of the overlap sample, and δ is the Kronecker delta function. The probability of a galaxy belonging to wide cell $\hat{c}$ is given by:

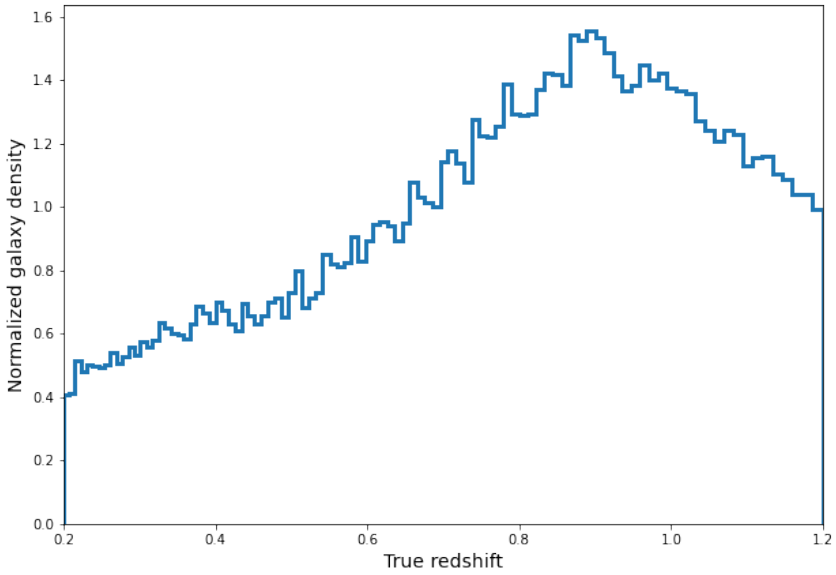


Figure 2.13: Redshift distribution of the 1×10^6 galaxies in the main sample.

$$p(\hat{c}|s) = \frac{1}{n_s} \sum_{i \in s} \delta_{\hat{c},\hat{c}_i} \, . \tag{2.38}$$

The transfer function then becomes the number of galaxies with unique combinations of $(c, \hat{c})$, divided by the number of galaxies in each wide cell $\hat{c}$:

$$p(c|\hat{c}, s) = \frac{\sum_{i \in s} \delta_{c,c_i} \delta_{\hat{c},\hat{c}_i}}{\sum_{i \in s} \delta_{\hat{c},\hat{c}_i}} \, . \tag{2.39}$$

2.7 Generating wide and deep samples

Since the catalog includes both true and observed magnitudes for every galaxy, we will consider the ideal case where we use the same 1×10^6 galaxies that have been randomly sampled from the truncated cosmoDC2 catalog (i.e. with the selection cuts in Table 2.4

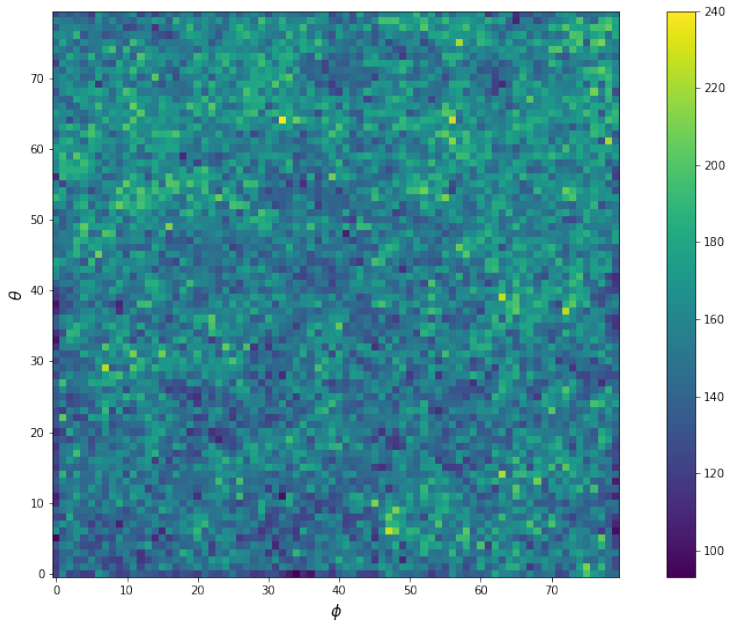


Figure 2.14: A two-dimensional histogram of angular bins for all 1×10^6 main sample galaxies.

applied) as the main sample. In this case, the overlap sample, wide, and deep samples are all from the main sample and consist of an identical set of galaxies. The analysis in this chapter was performed using the `SOMVIZ` package (Abolfathi 2022), which enables training and classification methods to generate self-organizing maps, as well as additional methods to calculate various metrics and produce visualizations of the maps.

The normalized redshift distribution of the main sample is shown in Figure 2.13. Galaxy redshifts were translated into indices belonging to bins of finite width. Discrete redshift indices are represented as 100 bins equally spaced in comoving distance between $0.2 \leq z \leq 1.2$. The distribution of redshifts with this binning scheme is shown in Figure 2.13.

For the clustering portion of the likelihood function, we use the true density field of the wide sample, rather than a separate tracer sample. Therefore, our density fluctuation field estimators $\hat{\delta}_z$ are the true fluctuations δ_z, and galaxy bias is implicitly accounted for. Angular positions (ϕ, θ) are defined based on the tangent plane to the (RA,DEC) coordinates by

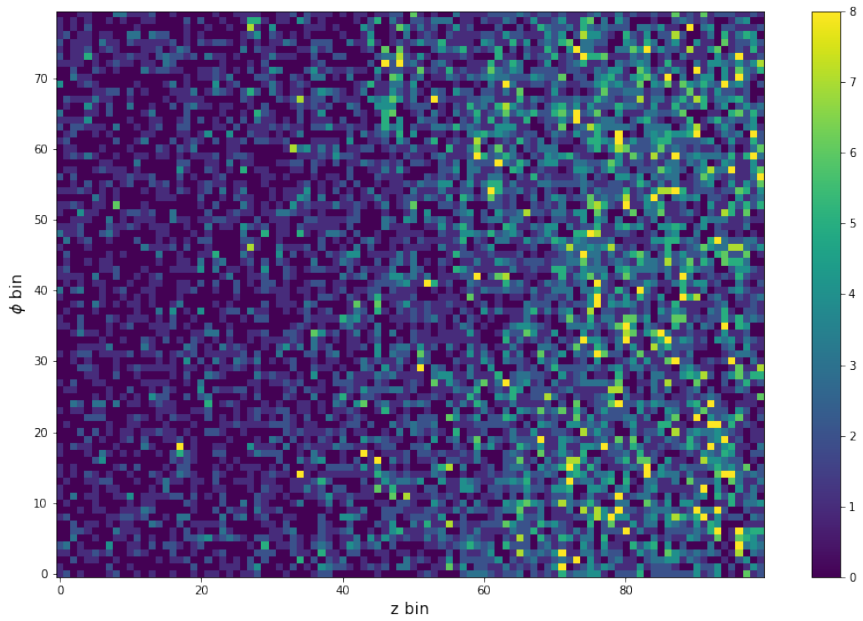


Figure 2.15: A two-dimensinoal histogram of ϕ and z bins for all galaxies in the main sample.

subtracting the mean RA/DEC of the entire galaxy sample from each galaxy coordinate. These adjusted coordinates are then binned into a 2D histogram consisting of an 80×80 grid of (θ, ϕ) bins (see Fig. 2.14). A histogram showing the density of galaxies in the main sample in (ϕ, z) is shown in Figure 2.15.

The features in the deep sample consisted of four colors (u-i, g-i, r-i, z-i), whereas the wide sample features contained one magnitude and three colors (i, g-i, r-i, z-i). The deep features were obtained using true magnitudes, and the wide features consisted of magnitudes with simulated 10-year errors. Each wide color was normalized to lie within the range $[0, 1]$, while the wide $i-$band magnitude was normalized to span the range $[0, 0.1]$, such that wide colors were given 10x more weight than the magnitude. Pairplots of the feature distributions used to train the SOM for the deep and wide sample are shown in Figures 2.16 and 2.17, respectively.

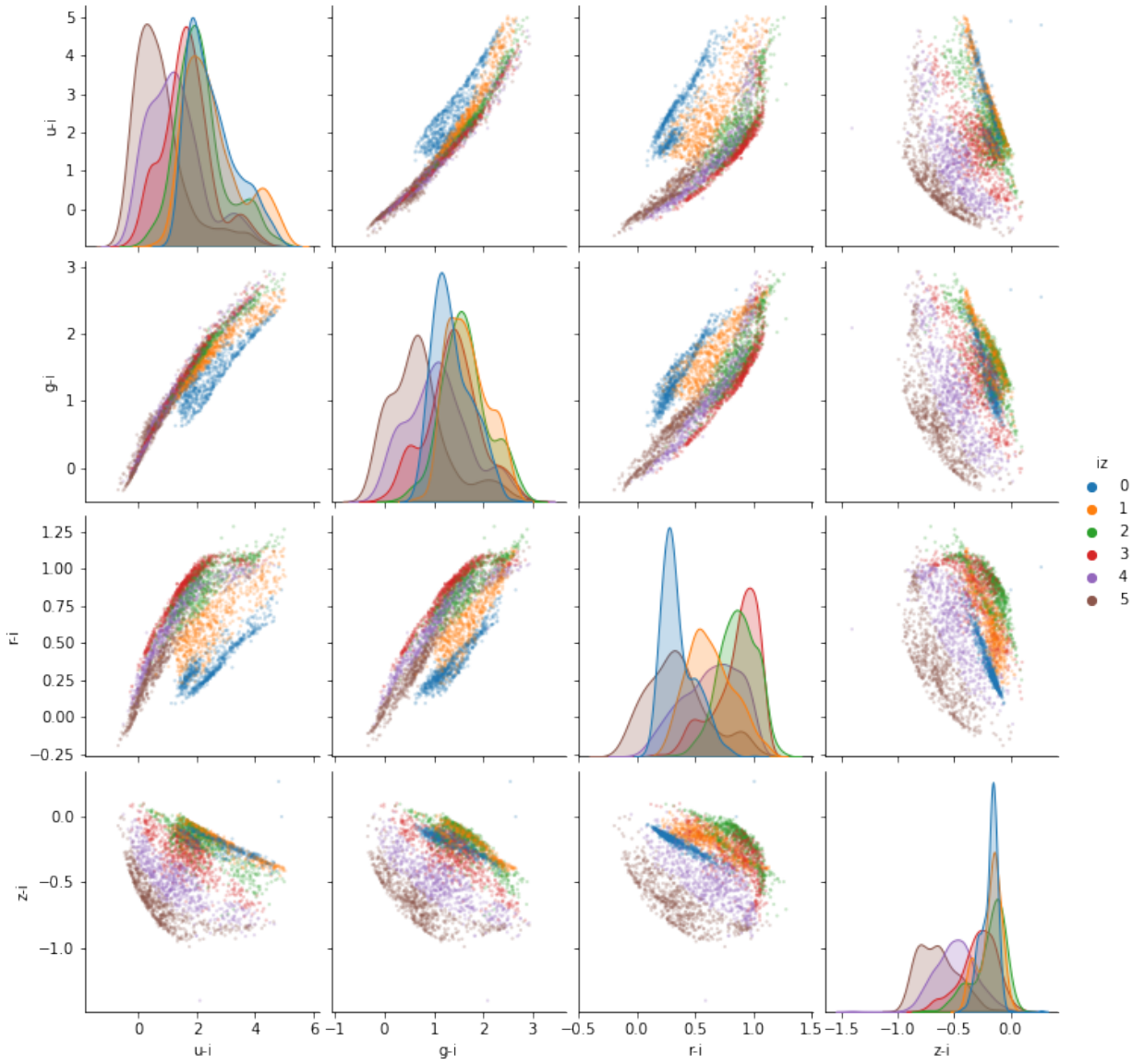


Figure 2.16: A pairplot of the deep features for the first 4×10^3 galaxies in the main sample, equally divided into six redshift percentile bins. The features consist of the four colors: u-i, g-i, r-i, and z-i.

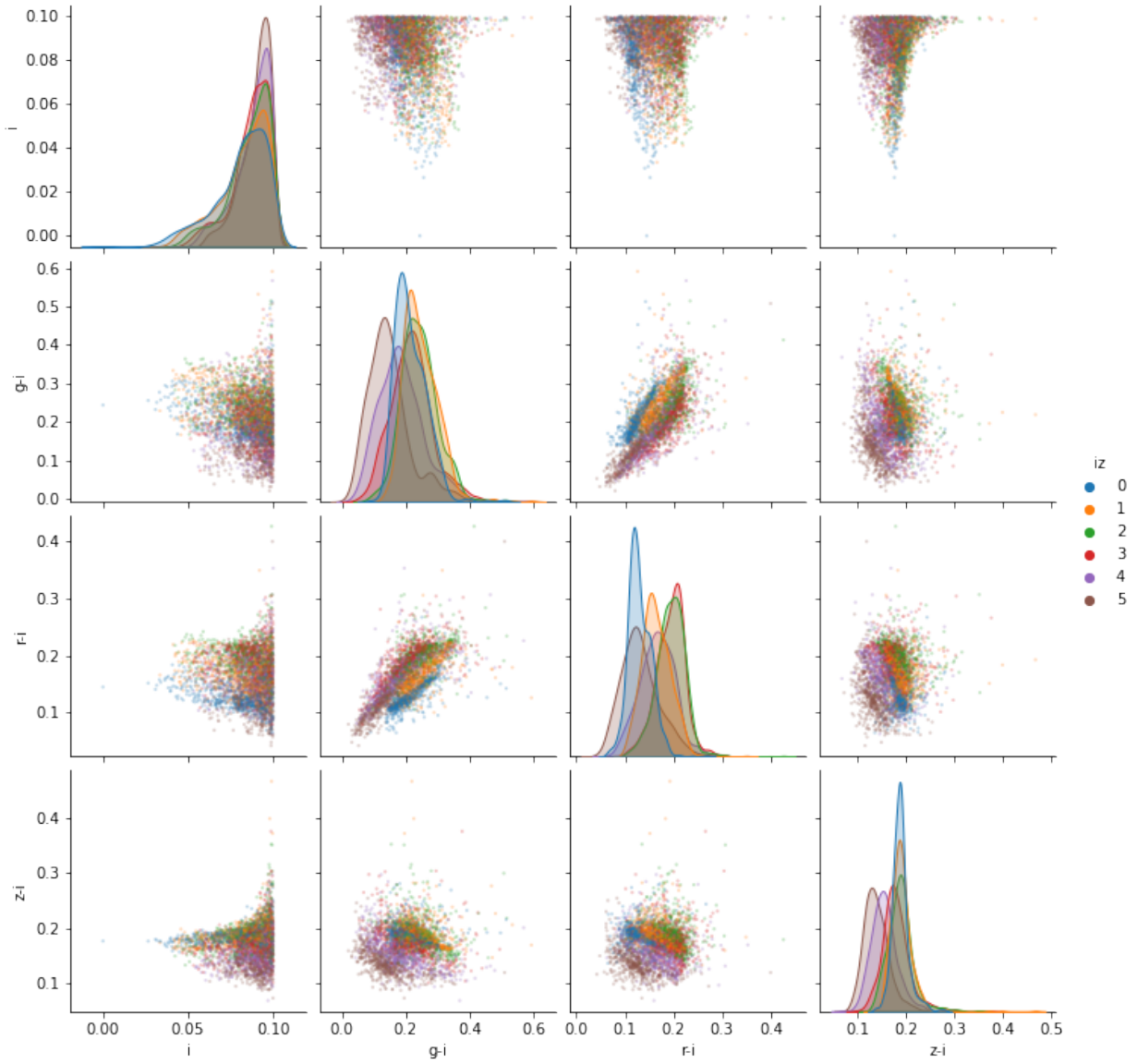


Figure 2.17: A pairplot of the wide features for the first 4×10^3 galaxies in the main sample, equally divided into six redshift percentile bins. The features consist of three normalized colors and one magnitude: i, g-i, r-i, and z-i.

2.8 Results of the Gibbs sampler on cosmoDC2

Both deep and wide SOMs were trained on a 14x14 grid for 100 iterations, in batches, where each batch consisted of a 1% randomly-sampled subset of each sample. The loss function for each is shown in Figure 2.18. The number of galaxies per cell in each trained SOM, obtained by mapping every galaxy in the main sample to its best-matching deep and wide cell, is shown in Figure 2.19. A plot of the predicted redshifts versus true redshifts for both deep and wide SOMs is shown in Figure 2.20. The mean redshift and standard deviation of the redshifts per cell are shown in Fig. 2.21 for both deep and wide SOMs, along with histograms of each. While the range of redshifts for both wide and deep SOMs overlap for the most part, the standard deviation is much better constrained by the deep SOM.

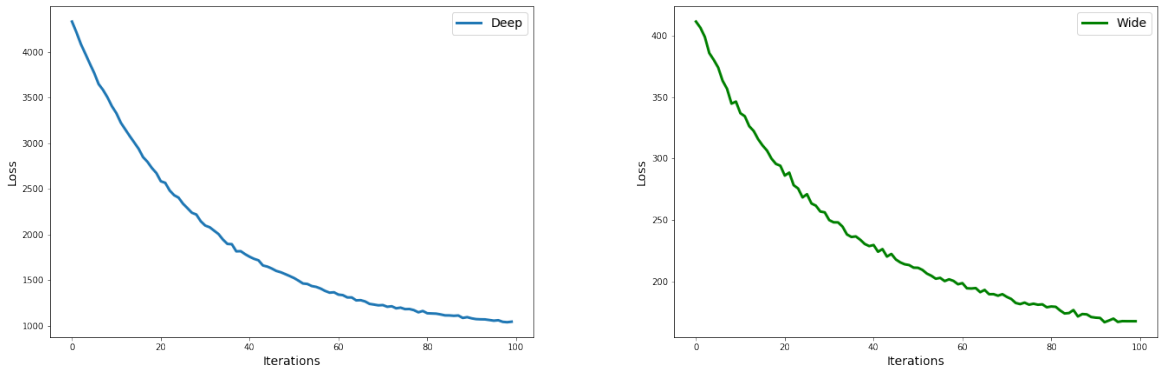


Figure 2.18: The loss function from training the deep (left) and wide (right) SOMs.

Next, the transfer function (Eq. 2.39) is calculated by finding the best-matching cell in the deep SOM for every wide SOM cell. A 1000-step Gibbs sampler was then run for four different cases: (1) prior-only, (2) feature likelihood and prior, (3) the same case as (2), but incorporating the clustering part of the likelihood function up to $z < 0.5$, and (4) the same case as (2) with the full clustering likelihood up to $z < 1.2$. Each of the four panes has a figure on the left showing a histogram of the sampled galaxies in blue, along with the likelihood and true redshift, and the sampled redshifts as a function of step number on the

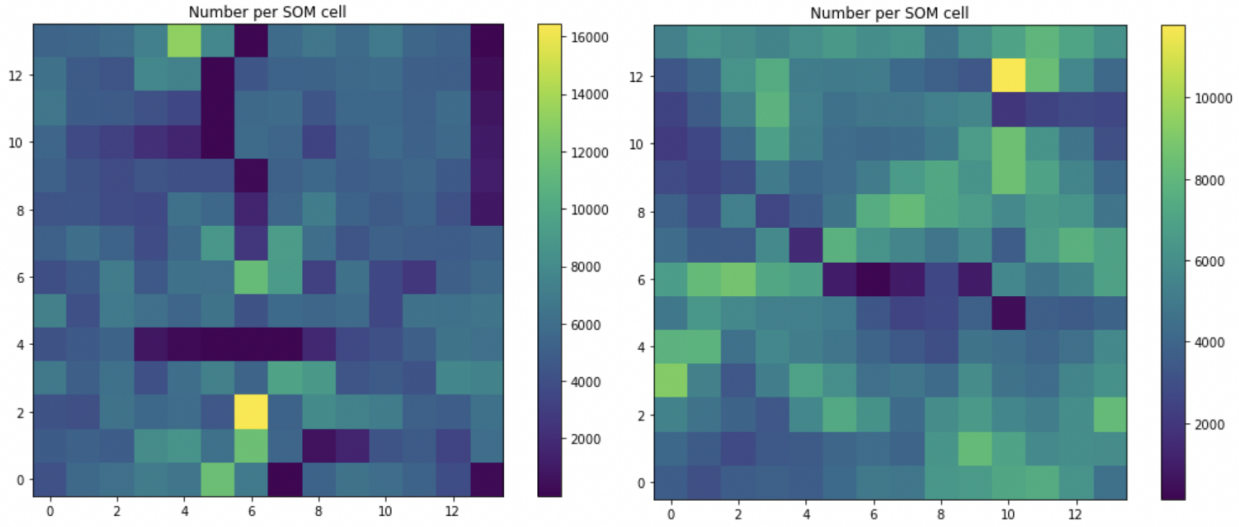


Figure 2.19: The number of galaxy counts per deep (left) and wide (right) SOM cell.

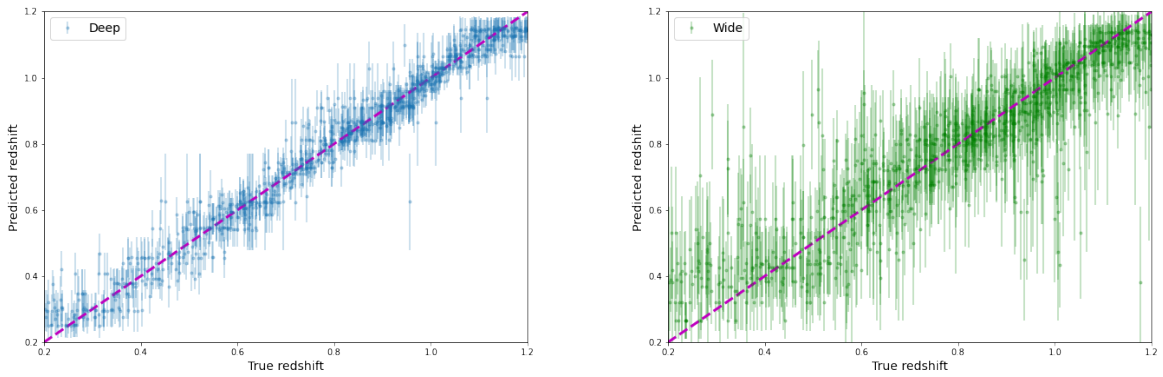


Figure 2.20: The predicted redshifts obtained by mapping each galaxy in the main sample to its best-match deep (left) and wide (right) SOM cell, plotted against their true redshifts. The error bars correspond to the standard deviation of the redshift distribution of each best-matching cell.

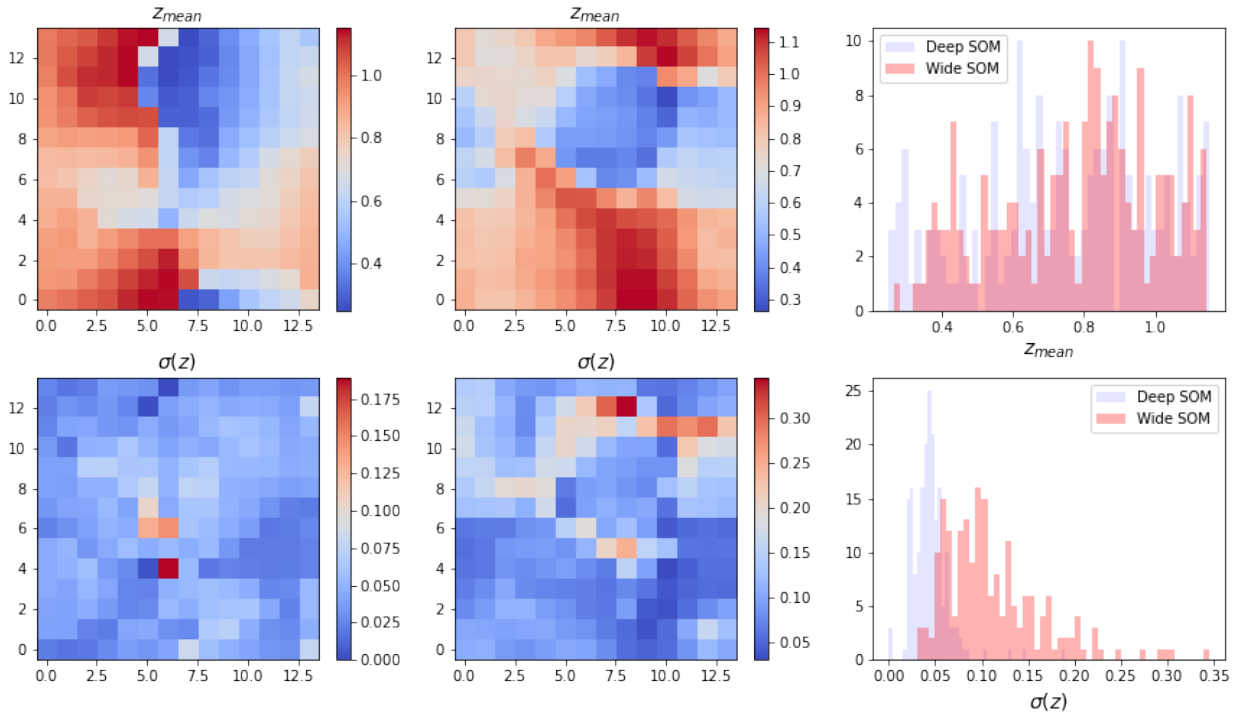


Figure 2.21: The mean of the per-cell redshift distribution for the deep (top left) and wide (top middle) SOMs, along with the histogram of the mean redshift for cells in each SOM (top right). Similar plots are shown in the bottom row for the standard deviation of the per-cell redshift distribution. The histogram on the bottom right clearly shows the deep SOM to be more robust.

right. The results for a single galaxy in the main sample are shown in Fig. 2.23. Including only the prior does very little to constrain the redshift. The histogram of sampled galaxies in pane (b) is slightly more peaked near the true redshift when adding the feature likelihood, but the redshift is still poorly constrained. In pane (c), adding the clustering portion of the likelihood decreases the probability of sampling a galaxy at lower redshifts. Since the sampled galaxy's true redshift is greater than $z = 0.5$, the constraining power of including the full clustering likelihood can be easily seen in pane (d), where the histogram is more sharply peaked near the true redshift.

Each of the cases above was run for four separate chains, each consisting of a 1000 cycles of the Gibbs sampler. The four chains were combined and the Gelman & Rubin metric (Eq. 2.24) was used to determine their convergence, with the results shown in Table 2.5.

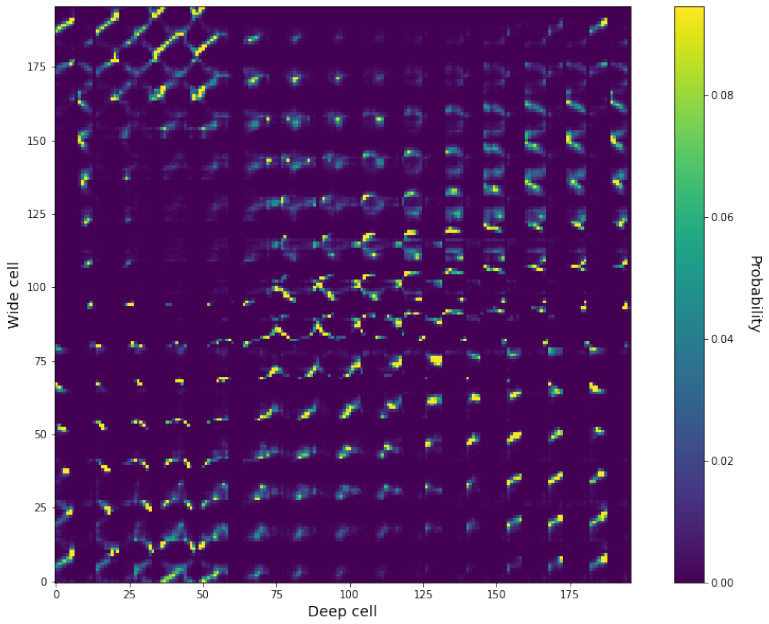
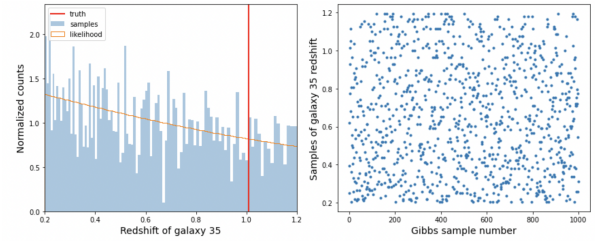


Figure 2.22: The transfer function, used to estimate the likelihood function, captures the probability of each wide cell being mapped to each deep cell.

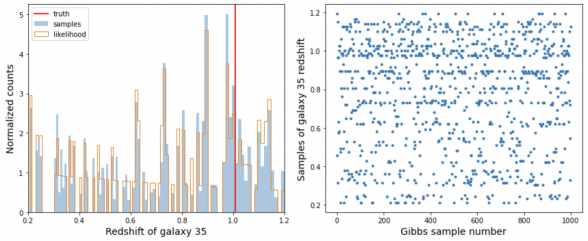
The Δz metric from Eq. 2.23 was used to determine the overall performance of the model. The result for the full posterior is shown in Fig. 2.24, with a distribution mean of around $\Delta z \approx 0.02$.

2.9 Discretizing types and redshifts with `zotbin`

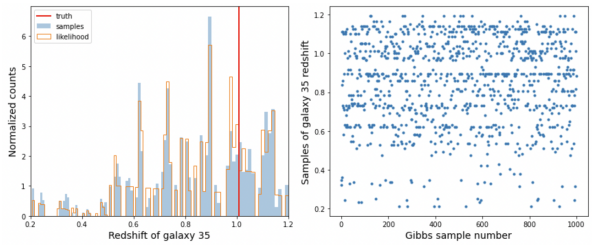
This section compares the performance of binning galaxies as a function of type and redshift via a self-organizing map with that of a method developed by David Kirkby that uses a generative model called a normalizing flow.



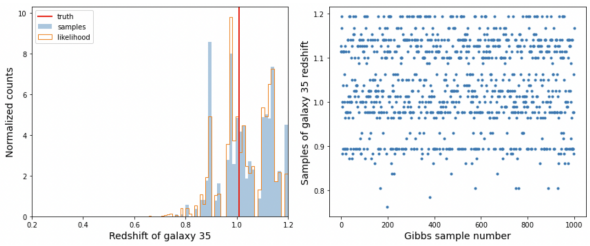
(a) Prior only.



(b) Prior + Feature likelihood.



(c) Prior + Feature likelihood + $\delta(z \leq 0.5)$.



(d) Prior + Feature likelihood + δ(All z).

Figure 2.23: Results of running a 1000-step Gibbs sampler for galaxy 35 for different combinations of (a) prior, (b) feature likelihood, (c) clustering up to $z \leq 0.5$, and (d) clustering up to $z \leq 1.2$.

Posterior condition	Description of posterior condition	R - 1
`prior`	Prior only	-0.0001
`noz`	Prior + Feature likelihood	-0.0002
`halfz`	Prior + Feature likelihood + $\delta(z \leq 0.5)$	0.0002
`fullz`	Prior + Feature likelihood + δ(All z)	0.0004

Table 2.5: Gelman & Rubin convergence results for the four different cases shown in Fig. 2.23.

2.9.1 Normalizing flows

While a SOM can be used to organize galaxies into discrete phenotypes, its performance depends on how well the features are separated to begin with. Figures 2.16-2.17 show a pairplots of the feature distributions used to train the SOM for the deep and wide sample, respectively. It is clear that the data occupy narrow regions of the feature space and show strong correlations. If we could somehow spread out these distributions so that they are nearly uniform and uncorrelated, we could more easily differentiate between individual features, leading to better resolved galaxy phenotypes. This transformation from the native,

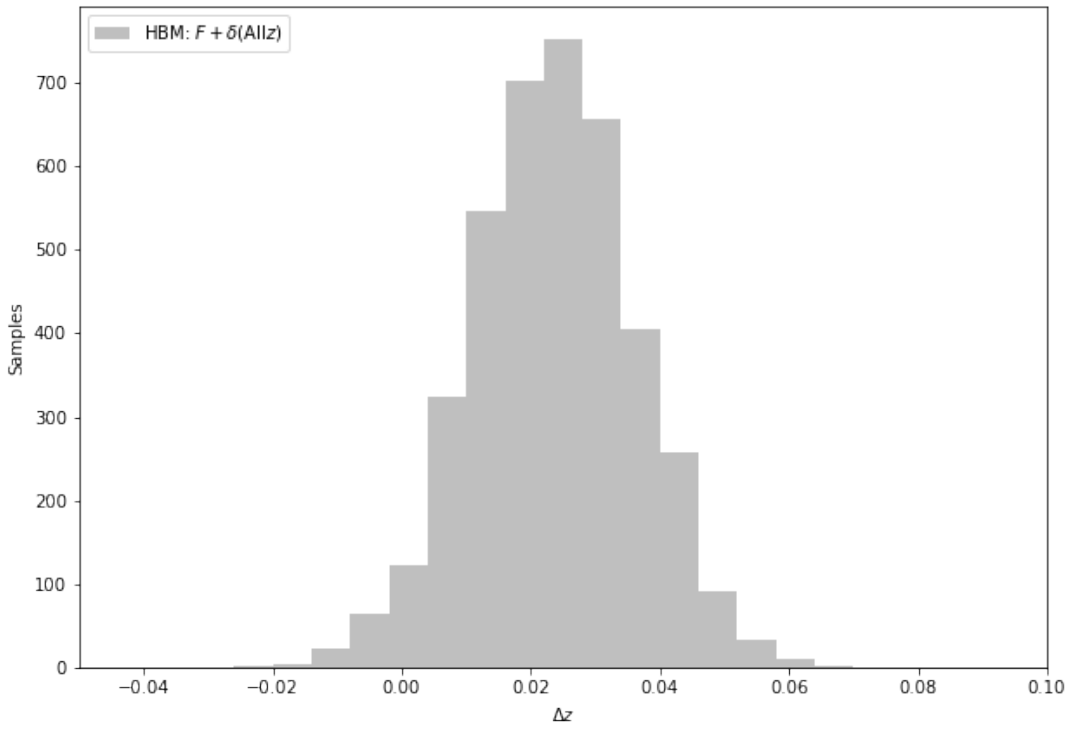


Figure 2.24: The Δz metric used to assess the performance of the hierarchical Bayesian model for the case of incorporating the full posterior.

m-dimensional feature distribution to a simpler, base distribution of the same dimensions is accomplished through the learning of a normalizing flow.

A normalizing flow is a type of generative modeling technique similar to that generative adversarial networks (GANs) Goodfellow et al. (2014) or variational auto-encoders (VAEs) Kingma & Welling (2014), wherein the goal is to characterize an unknown probability distribution given samples drawn from that distribution Kobyzev et al. (2020). However, unlike GANs and VAEs, normalizing flows allow for the exact and efficient evaluation of the probability distribution through a chain of parameterized invertible and differentiable functions. More specifically, they provide a mechanism to construct probability distributions by transforming a simpler, base distribution into a more complex distribution through a series of

diffeomorphisms. The complex distribution can then be sampled from by first sampling from the base distribution and then applying the transformations. The following section outlines the basic steps of this process.

Flow basics

Let $\mathbf{Y} \in \mathbb{R}^D$ be a random variable drawn from a complex, irregular distribution $p_{\mathbf{Y}} : \mathbb{R}^D \to \mathbb{R}$, and let the random variable $\mathbf{Z} \in \mathbb{R}^D$ belong to a simple, base distribution $p_{\mathbf{Z}} : \mathbb{R}^D \to \mathbb{R}$. We can express $\mathbf{Y}$ in terms of a forward mapping from the base distribution, represented by a differentiable and invertible function $\mathbf{g}(\mathbf{Z}) = \mathbf{Y}$. The transformation $\mathbf{g}$ can be thought of as reshaping a small neighborhood around $\mathbf{z}$ to mold $p_{\mathbf{Z}}(\mathbf{z})$ into $p_{\mathbf{Y}}(\mathbf{y})$. If the inverse mapping is given by the function $\mathbf{f}(\mathbf{Y}) = \mathbf{Z}$, where $\mathbf{f}$ is also differentiable and invertible, then the probability distribution of the random variable $\mathbf{Y}$ can be expressed as the product of the inverse-transformed sample, and the associated change in volume given by the determinant of the Jacobian matrix $\mathbf{J_f}$ for that transformation:

$$p_{\mathbf{Y}}(\mathbf{y}) = p_{\mathbf{Z}}(\mathbf{z}) \left|\det \mathbf{J_g}(\mathbf{z})\right|^{-1} \tag{2.40}$$

$$= p_{\mathbf{Z}}(\mathbf{f}(\mathbf{y})) \left|\det \mathbf{J_f}(\mathbf{y})\right| . \tag{2.41}$$

Because the transformations $\mathbf{g}$ and $\mathbf{f}$ are differentiable and invertible, they are also composable. In other words, if $\mathbf{g}$ represents the series of transformations $\mathbf{g} = \mathbf{g_N} \circ \mathbf{g_{N-1}} \circ ... \circ \mathbf{g_1}$, then $\mathbf{f} = \mathbf{g}^{-1}$ is given by $\mathbf{f} = \mathbf{f_1} \circ ... \circ \mathbf{f_{N-1}} \circ \mathbf{f_N}$. Moreover, following Eq. 2.41, the probability density of the base distribution can be expressed as:

$$p_{\mathbf{Z}}(\mathbf{z}) = p_{\mathbf{Y}}(\mathbf{g}(\mathbf{z})) \left|\det \mathbf{J_f}(\mathbf{g}(\mathbf{z}))\right|^{-1} . \tag{2.42}$$

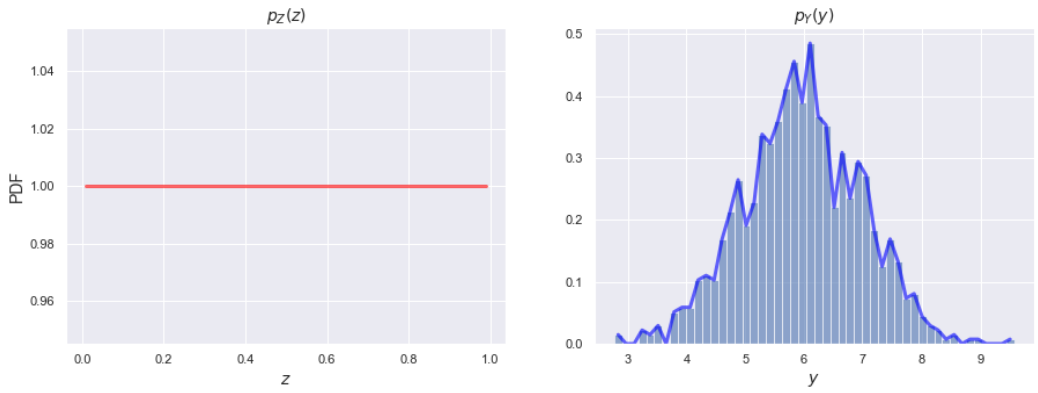


Figure 2.25: A flow in the generative direction, going from a uniform distribution to a Gaussian distribution.

The transformation from the simple, base distribution $p_{\mathrm{Z}}(\mathbf{z})$ to the more complex density $p_{\mathrm{Y}}(\mathbf{y})$ is called a *pushforward* of the density $p_{\mathrm{Z}}(\mathbf{z})$ in the *generative* direction. The inverse transformation, which molds the complex, irregular distribution into the base distribution, is in the *normalizing* direction. The latter mapping is what we will use to transform the cosmoDC2 data from a sparse feature space exhibiting complex correlations into a space where the galaxy features are more uniformly distributed.

A simple, one-dimensional demonstration of this is shown in Fig. 2.25 using a uniform distribution $\mathcal{U}(0,1)$ as the base distribution $p_{\mathbf{Z}}(\mathbf{z})$, and an empirically derived distribution sampled from $\mathcal{N}(6,1)$ as the complex distribution $p_{\mathbf{Y}}(\mathbf{y})$. We can transform in the normalizing direction, from $p_{\mathbf{Y}}(\mathbf{y})$ to $p_{\mathbf{Z}}(\mathbf{z})$, so that the cumulative distribution function of $p_{\mathbf{Y}}(\mathbf{y})$ matches that of a uniform distribution $\mathcal{U}(0,1)$ (see Fig. 2.26).

2.9.2 Calculating flows with `zotbin`

The `zotbin`[3]. package contains a set of algorithms used for the implementation of the ZotBin method, a binning scheme developed for the DESC 3x2pt Tomography Challenge Zuntz et al. (2021). ZotBin consists of a pre-processing stage, a grouping stage, and a stage

[3]`https://github.com/dkirkby/zotbin`

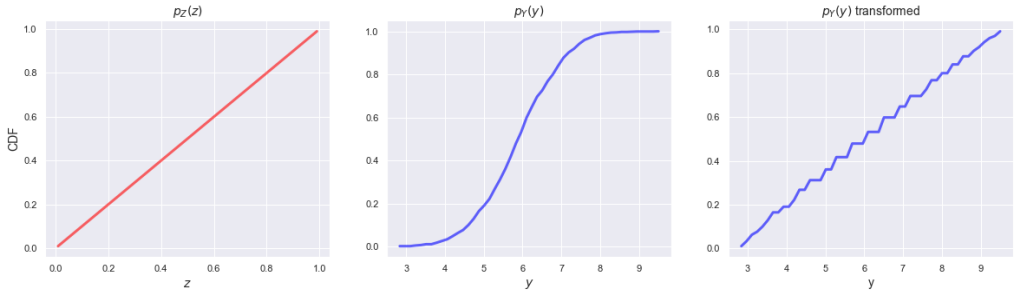


Figure 2.26: A flow in the normalizing direction. The cumulative distribution functions (CDF) for a uniform distribution and for a Gaussian distribution are shown on the left and middle, respectively. The transformed Gaussian CDF on the right resembles the original uniform CDF (left).

that optimizes redshift bin edges for tomographic weak lensing. Methods related to the first two stages were used to assign phenotypes to galaxies.

Bins in `zotbin` are initially defined as a rectangular grid in feature space, where each feature axis contains an equal number of samples (show plot). The pre-processing stage involves using a normalizing flow to transform the n-dimensional feature space into a new, simpler space with the same dimensions. This is done by first mapping the original probability density into a multivariate unit Gaussian, and then into a uniform distribution through a second transformation by applying the error function. This simpler distribution exhibits a nearly uncorrelated distribution on the unit hypercube $[0, 1]^n$.

The grouping stage is performed by dividing the unit hypercube into a regular lattice of $\mathcal{O}(10\text{k})$ cells. These cells are then merged into 100(?) groups in an iterative fashion. Cells are merged based on their "similarity", expressed as the product of their independent redshift and feature similarities. The feature similarity vector for a single cell i is defined in terms of the grid separations $\mathbf{dr}_i$ between itself and every other cell in the multidimensional feature space:

$$\mathbf{f}_i = \exp^{-(\mathbf{dr}_i/\sigma)^2} \ . \tag{2.43}$$

The hyperparameter σ determines the relative importance of the feature and redshift similarities when grouping cells. Redshift similarity is defined as the cosine of the angle between the redshift histograms of each cell. Grouping is performed iteratively by combining pairs of groups with the maximum similarity until a specified number of groups is reached. When two groups are merged, their redshift histograms are added, and their feature similarities are updated with the other remaining groups. The final groups represent the distinct galaxy phenotypes that will be compared with the results of the SOM in Section 2.8.

2.9.3 Grouping galaxies: `SOMVIZ` versus `zotbin`

A possible alternative to using a self-organizing map to build the feature likelihood is to implement the grouping scheme in `zotbin`. Whereas a SOM is ultimately two-dimensional, `zotbin` uses a non-linear transformation to transform from a highly complex, n-dimensional space to a new n-dimensional space where the data are uncorrelated. Figure 2.27 shows the four features in the deep sample before a normalizing flow has been applied, while Figure 2.28 shows the new feature space after this preprocessing with a flow.

The redshift-type histograms for the wide and deep SOMs are shown in Figures 2.29-2.30. The effect of using true versus measured fluxes is easily visible, as the deep histogram is clearly better at constraining the redshift-type relation over the wide histogram. Figure 2.31 shows a similar ridge plot using the grouping method in `zotbin`. Although the distribution is not as narrowly peaked as the histogram in Fig. 2.29, its performance over a SOM in estimating the feature likelihood still needs to be investigated.

2.10 Conclusion

A Hierarchical Bayesian model is able to successfully incorporate photometric, spectroscopic, and clustering information to constrain redshifts, as well as provide estimates of their errors. After demonstrating the success of this method on a toy model, the method was applied to simulated data from cosmoDC2, where results showed the added benefit of including clustering information on constraining the posterior. A self-organizing map was used to generate the feature likelihood, which gave probabilities of a galaxy having a particular redshift given its assignment to a discrete phenotype. This method of binning galaxies by redshift and phenotype was compared to a grouping method using the `zotbin` package. Future work would involve investigating the performance of each of these to find an optimal tomographic binning scheme for maximizing the figure-of-merit for the 3x2pt probe.

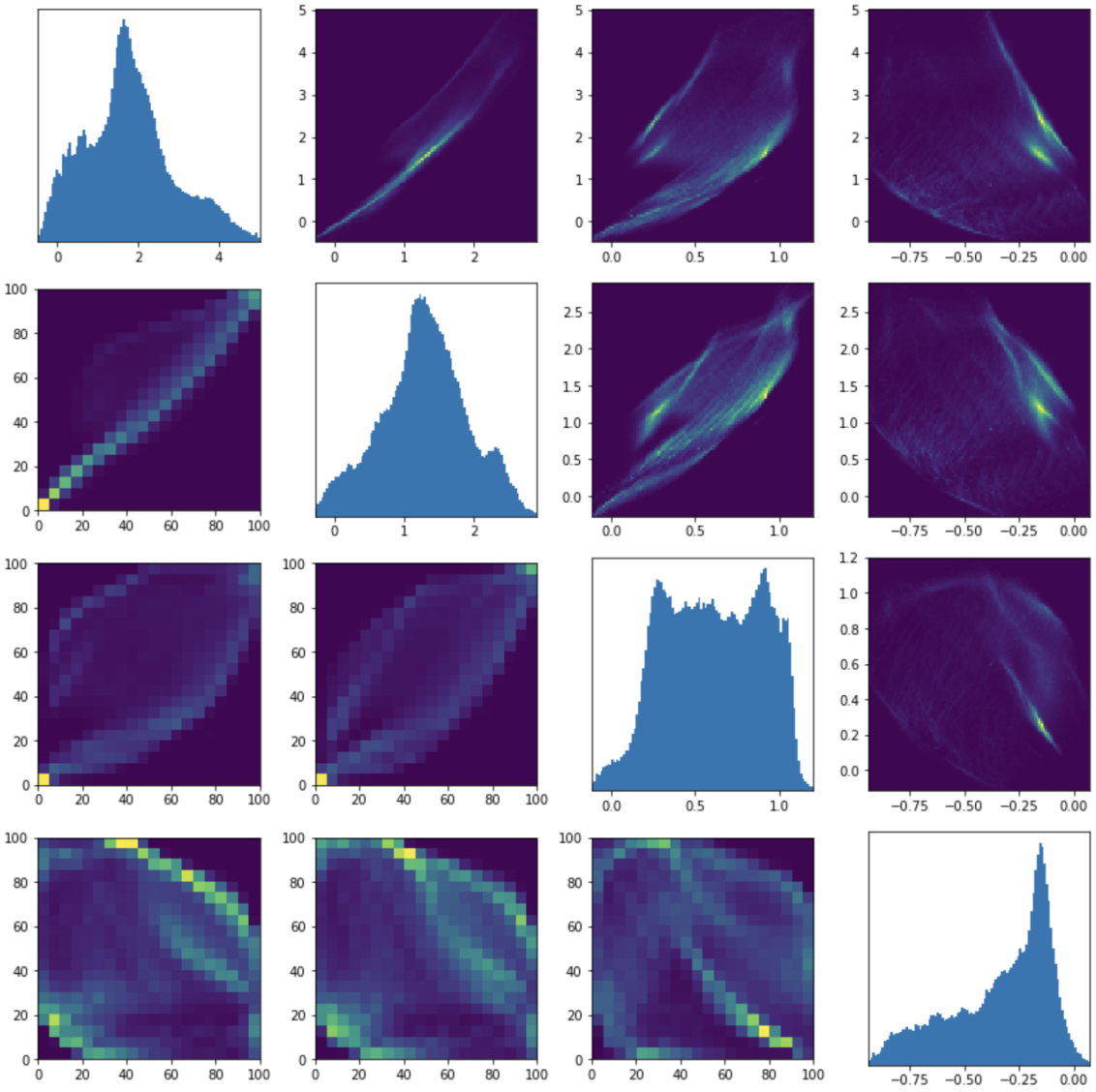


Figure 2.27: A pairplot showing the deep feature space (with axes similar to Fig. 2.16) before applying a normalizing flow.

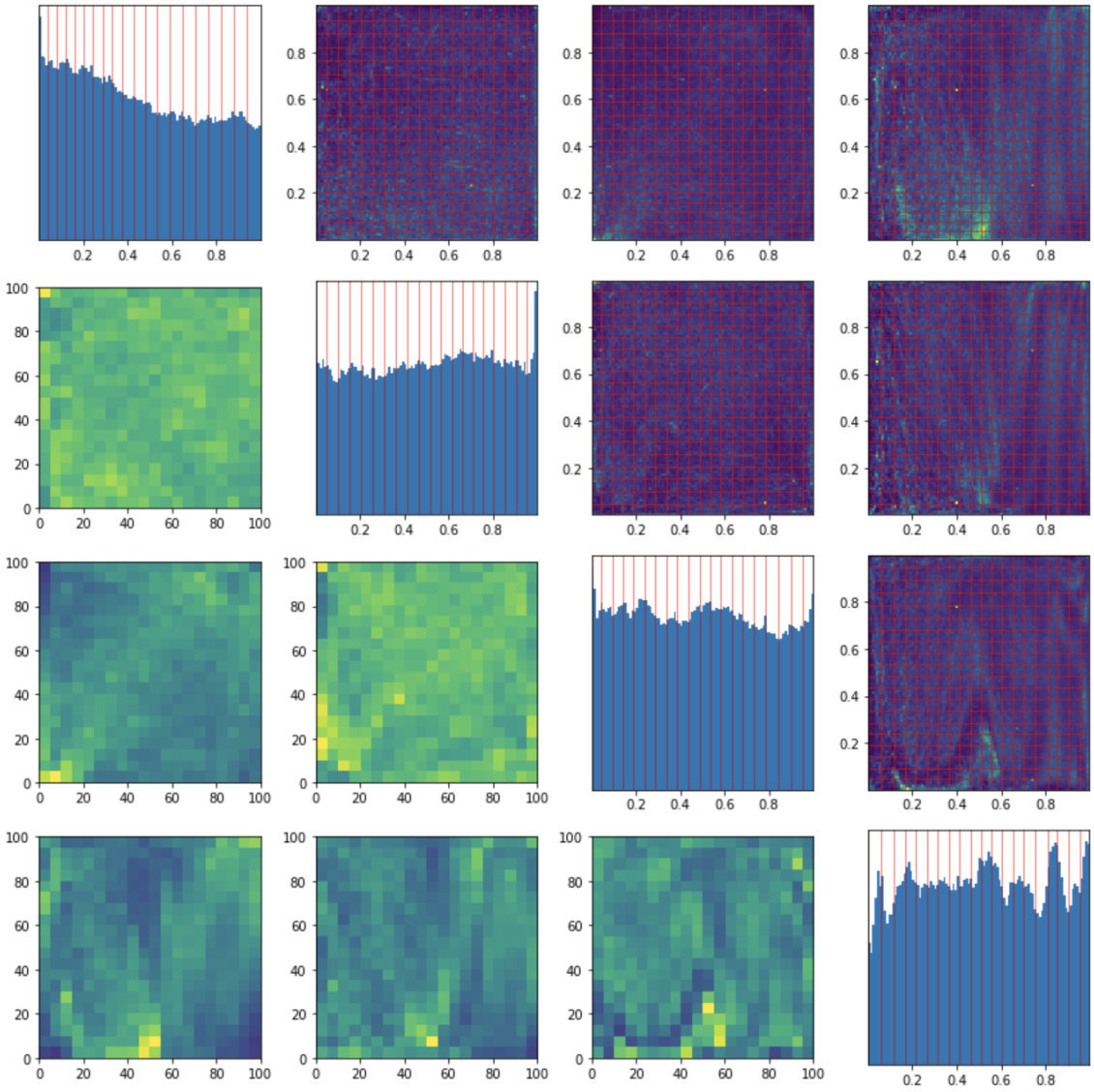


Figure 2.28: A pairplot showing the feature space after applying a normalizing flow. It's clear how the feature space is more uniform, especially from the histograms along the diagonal.

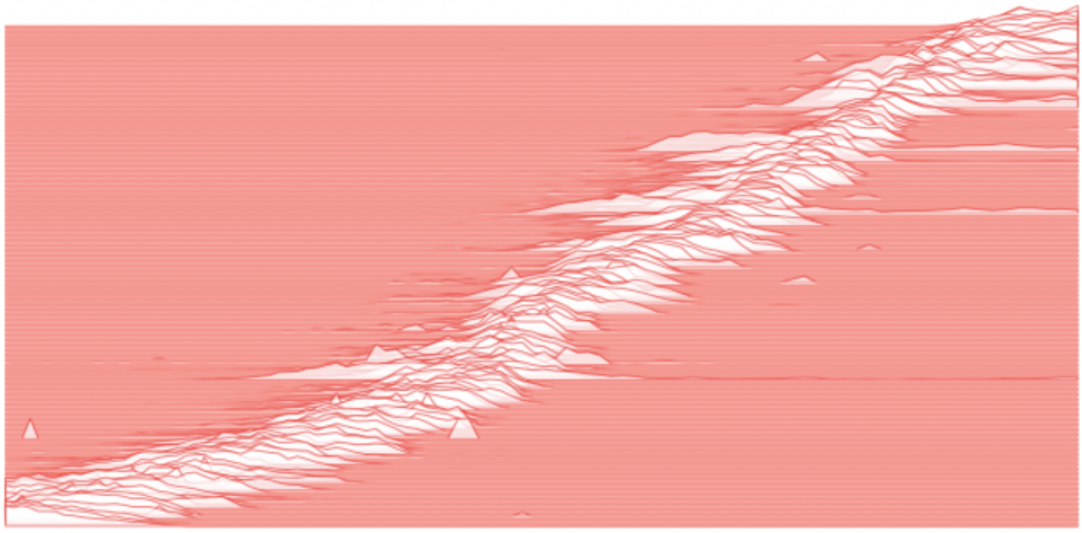


Figure 2.29: A ridge plot showing a histogram of the true redshifts of galaxies in the deep sample (x-axis) against their deep cell assignment, or phenotype (y-axis). The histogram is normalized by the maximum number of galaxies in each redshift bin.

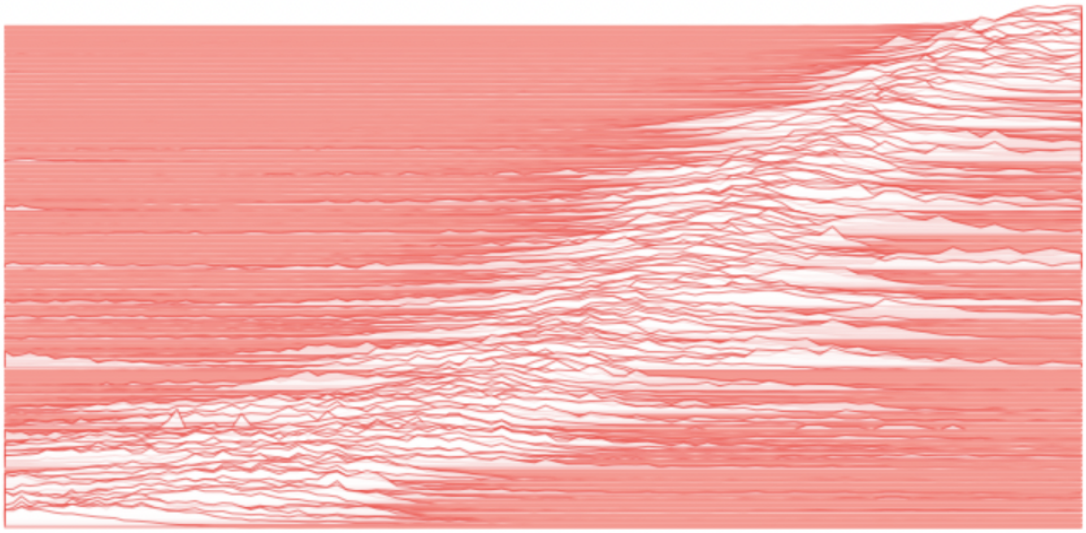


Figure 2.30: A ridge plot showing a histogram of the true redshifts of galaxies in the wide sample (x-axis) against their wide cell phenotype (y-axis). The histogram is normalized by the maximum number of galaxies in each redshift bin.

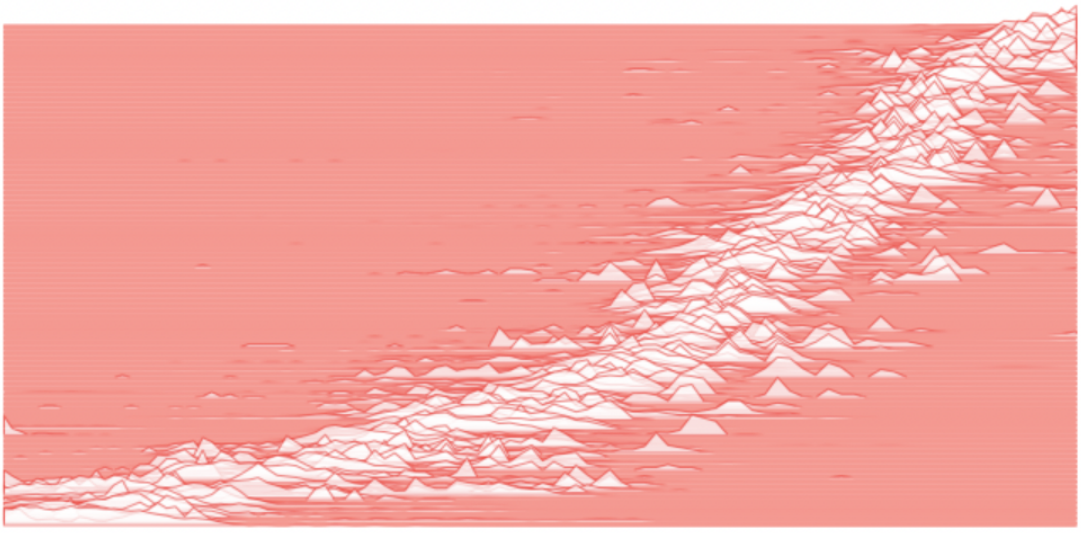


Figure 2.31: A ridge plot produced by the `zotbin` grouping method, with the same binning scheme on the x-axis (100 redshift bins equally spaced in comoving distance), and 100 phenotype cells on the y-axis.

Chapter 3

Sensor effects in LSST

3.1 Introduction

Addressing the science goals for LSST requires accurate measurements of cosmological parameters, which is partly limited by the systematics of the instrument response. The work in this chapter improves upon our understanding of the LSST Camera and sensors by investigating anomalous signatures in bias images taken during the Integration and Verification Testing phase of Camera assembly. In addition to looking at real Camera data, we will also discuss work related to simulating the sensors. The latter half of the chapter will describe the effects of varying the sky model across the focal plane in the `imSim` software package (Dark Energy Science Collaboration 2020).

3.2 LSST Camera design and system requirements

Measuring 1.65 meters high and 3 meters long, the LSST Camera is the largest digital camera ever constructed. Its focal ratio of f/1.2 and focal length of 10.31 meters yields a 3.5 degree

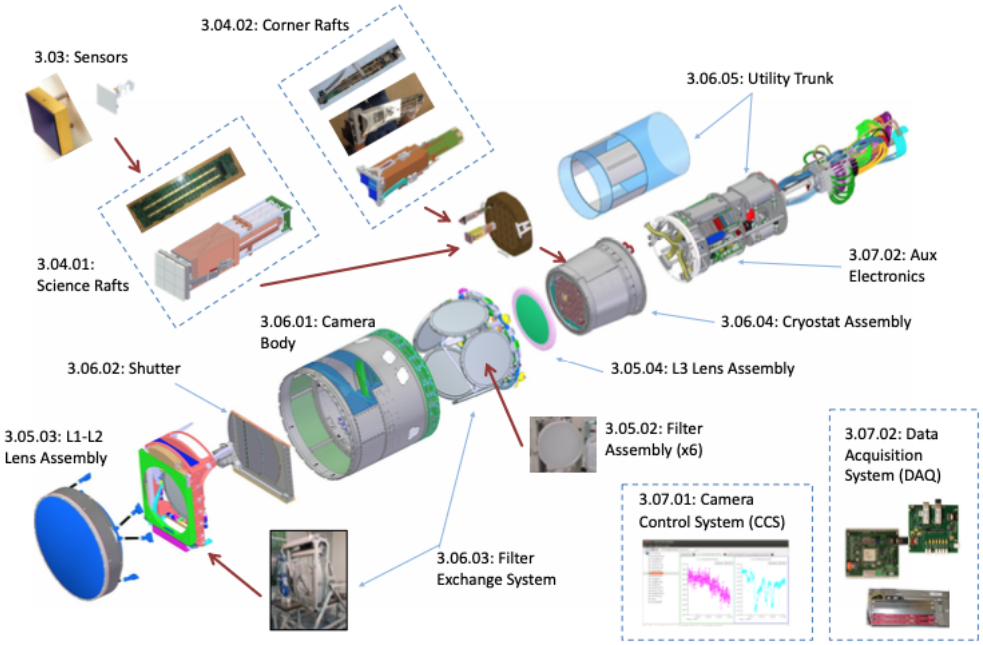


Figure 3.1: An exploded view of the LSST Camera. Figure is from Roodman et al. (2018).

field of view, allowing for a wide-fast-deep survey strategy that will cover a large portion of the sky with high cadence (Ivezić et al. 2019). The exploded view in Fig. 3.1 shows the components of the Camera structure, including the lens assembly, shutter, the Camera body, and the cryostat, which houses the focal plane (see Fig. 3.2).

The Camera design is guided by the LSST Camera system and Focal Plane requirements, and adherence to a strict delivery schedule. To this end, a modular design was adopted to enable assembly and testing of individual components in parallel in a test-as-you-build fashion, as well as to allow for the ability to quickly isolate any issues or defects without disturbing the rest of the focal plane. In keeping with this approach, the LSST focal plane consists of 189 CCDs arranged on 21 autonomous cameras called Raft Tower Modules (RTMs) (see Fig. 3.4). An RTM, along with its associated mechanical, thermal and electronic support components, consists of a square matrix of nine CCDs. In addition to the 21 science rafts, there are four additional corner rafts made up of wavefront sensing and guiding arrays used

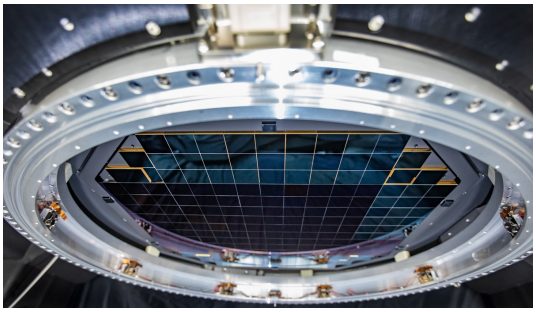
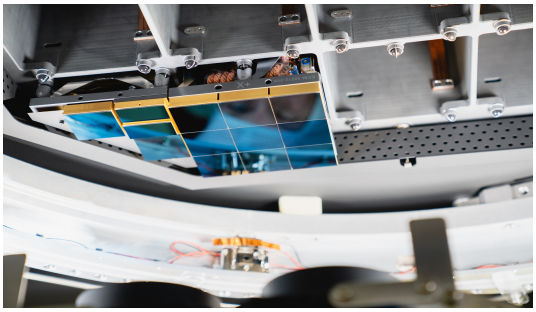



Figure 3.2: The LSST focal plane (left), containing all 21 science rafts and four corner rafts containing guide sensors and a split wavefront sensor for focusing, and a corner raft and full science raft tower (right).

Table 3.1: A subset of LSST Focal Plane requirements from Roodman et al. (2018).

Title	Requirement	Test
QE	$> 41\%$, 78%, 83%, 82%, 75%, 21% (u,g,r,i,z,y)	Cam. Calib. wide-beam
CCD Diffusion	$\sigma < 5\mu$m	^{55}Fe
Read Noise	$< 9e^-$	Bias images
Linearity	$< 3\%$	Flat-fields

for focusing and telescope guiding. Fig. 3.3 shows how the focal plane is subdivided into rafts, CCDs, and individual amplifiers.

A subset of the LSST Camera system requirements are shown in Table 3.1. The fully assembled Camera must satisfy these conditions in order to achieve the precision and accuracy needed for Stage-IV requirements. The survey's speed requirements include reaching the 5σ limiting magnitudes given in Table 3.2 in each of the six filter bands in each patch of the survey footprint, which will be imaged every three nights. It is therefore essential to minimize the amount of downtime between exposures by enabling a rapid readout of the focal plane. Thus, the entire 3.2 Gigapixel focal plane will be read out in two seconds. This is enabled by a CCD pixel architecture consisting of 4k x 4k 10μm square pixels and 16 independent amplifier segments, allowing for a pixel digitization rate of 0.5 MHz.

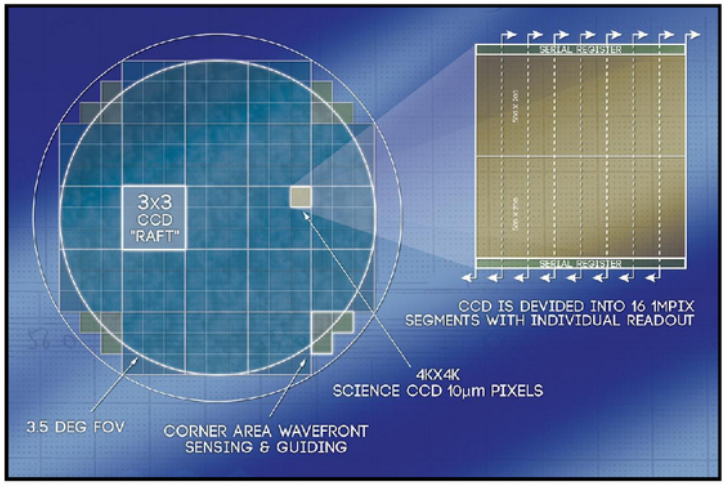


Figure 3.3: The LSST focal plane and blown-up image of an individual CCD with all 16 amplifier segments. Figure is from LSST Science Collaboration and LSST Project (2009).

Table 3.2: LSST survey depth requirements from Roodman et al. (2018).

Band	5σ depth per visit	5σ depth over 10 year survey
u	23.9	26.1
g	25.0	27.4
r	24.7	27.5
i	24.0	26.8
z	23.3	26.1
y	22.1	24.9

3.2.1 Camera workflow

The LSST Camera was assembled in the IR2 high bay in a clean room at the Stanford Linear Accelerator Center (SLAC). Although Camera assembly took place at SLAC, initial testing of the CCDs and raft assembly began at Brookhaven National Laboratory (BNL). Once the rafts were completed, they were shipped to SLAC where the results of a similar suite of tests were compared to those obtained at BNL for consistency. The bias studies in Section 3.3 took place at SLAC, prior to the delivery of the Camera to the telescope site on the Cerro Pachón ridge in north-central Chile.

Verification of Camera components was performed on a number of testing benches located in the SLAC Camera Clean Room. I specifically analyzed data from an optical testing bench

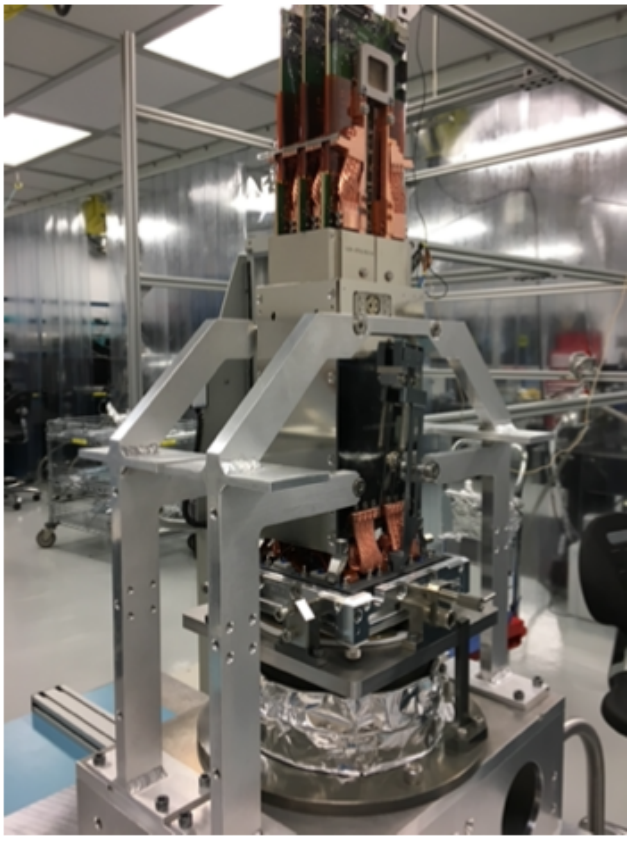


Figure 3.4: An individual science raft tower in the camera clean room. Figure is from Roodman et al. (2018).

called Test Stand 8 (TS8), which was used to perform acceptance tests on individual rafts. TS8 used a 300 Watt Xe arc lamp, a monochromator, a 6-position filter wheel, an off-axis parabolic feed optic, a 300 millimeter integrating sphere and a 1 meter baffled dark box to produce dark images, flat fields, ^{55}Fe X-ray sources, and multi-spot images to study noise, linearity, gain, cross-talk, dead pixels, as well as other unexpected instrument signatures (Bond et al. 2018). These electro-optical performance tests were run using a software package developed at SLAC called `eotest` (LSST Project 2019), which is a set of acquisition and analysis scripts used to test the electro-optical performance of the RTMs. I used the `eotest` package to study bias and offset corrections, as well as to characterize irregularities in the LSST sensors. Because they were produced by two different vendors, Imaging Technology Laboratory (ITL) and Teledyne e2v (e2v), the CCDs have their own unique signatures due to

differences in the manufacturing process. A careful study of sensor effects therefore involved evaluating data from TS8 at the raft, CCD, amplifier and vendor level.

3.3 Bias and offset studies

Bias frames are the first set of image sequences taken during pre-processing of Camera data. They are used to calibrate verification sequences related to a subset of the System-Level Verification Tests as laid out in the LSST Camera Verification Plan (LCA-283-F). Tests related to gain stability, linearity and dynamic range, cross-talk, noise, image quality, and diffusion requirements were performed at TS8. The verification sequences used for these studies included flat-fields (exposures of a uniformly illuminated focal plane) to measure gain and to study noise, ^{55}Fe X-ray sources to measure gain, dark images (non-zero exposures taken with the shutter closed) to study bad pixels and fixed-pattern noise due to the thermal properties of the sensors, and bias images (zero-second exposures taken with the shutter closed) to study the noise contribution of the sensor electronics. Because there is no charge integration in bias frames, these images do not contain any signal due to thermally excited or photo-electrons.

3.3.1 Offset correction

The pre-processing steps that involve bias images are bias and offset corrections. At the end of an exposure, the photo-electrons generated in the bulk silicon go through a capacitor, where their combined charge is registered as a voltage. Since this analog voltage cannot be measured to perfect accuracy, there is some degree of uncertainty when it is translated into digital counts via the analog-to-digital converter. If the signal in a pixel were smaller than the readout noise it could result in that pixel having negative counts. To avoid this, a bias

voltage must be applied to the capacitor to ensure that pixels with few or no photo-electrons still register non-negative counts. A bias image should, therefore, look flat overall, with an average count per pixel roughly equal to the offset level.

As a result, every CCD image contains a bulk offset level due to this bias voltage that must be subtracted prior to any scientific analysis (see Figure 3.5). This offset can be thought of as a "zero-point", and typically contains around 20,000 counts for LSST sensors. One can measure the offset in an image from the serial overscan region of a single channel in the CCD. The overscan region does not refer to a physical location on the sensor. Rather, this region contains "virtual" pixels that are the result of continuing the readout process after data from the last physical pixel in the serial register has been recorded. The overscan should theoretically not contain any signal from an external source. However, it does contain an artificially-induced offset signal, which guarantees that the signal processed by the analog-to-digital converter (ADC) is always positive.

I analyzed the overscan regions of bias images from TS8 to investigate the performance of various methods of offset correction at the raft, sensor and amplifier level. I then integrated two different offset-correction methods into the `oetest` software package. At the start of my project, the EOTest package had two ways to calculate the offset level. One was to measure the offset as the mean of all the pixels in the serial overscan region, and the other was to fit a low-order polynomial function to the mean number of counts per row in the overscan region. I compared these two methods to three other offset-correction methods: calculating the offset as simply the mean-per-row in the overscan, as a linear fit, and as a cubic spline fit to the mean-per-row in the overscan region. The cubic spline fit is computed using a weighted least-square spline fit, where the weights are given as $1/(rms \times \sqrt{nx})$, where $rms \approx 10$ and $nx = 50$ is the number of columns per row in the overscan. These comparisons can be seen in Figs. 3.6-3.7 for two different e2v CCDS on RTM 012 and RTM 005, respectively.

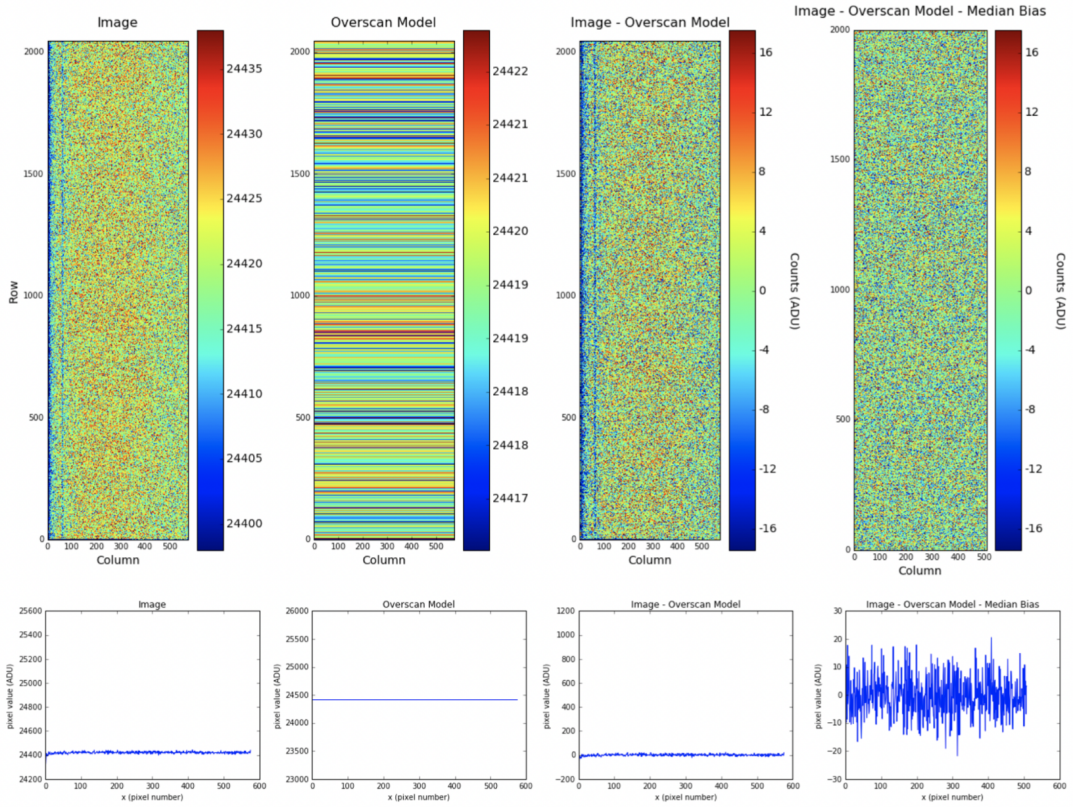


Figure 3.5: The process of performing bias and offset corrections for an e2v sensor on RTM 005 using the row-by-row mean method. A bias image is taken, then an overscan model image is generated from the mean-per-row in the serial overscan region. This overscan model is subtracted from the original bias image. Finally, the bias level is subtracted out by taking the difference between the image minus the overscan model and the median image of a stack of offset-corrected bias images. Under each image is a slice across an arbitrary row.

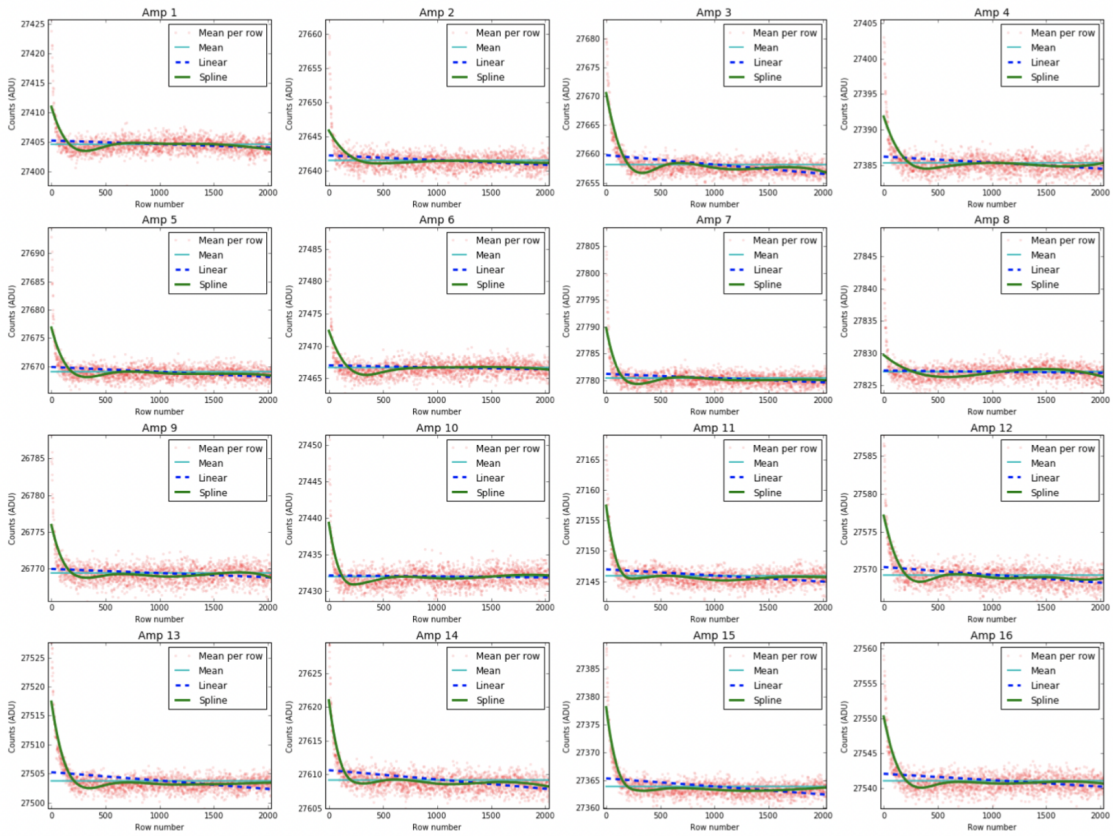


Figure 3.6: The performance of different offset models as fits to the mean-per-row in the serial overscan region of all 16 amplifiers on an e2v sensor on RTM 012.

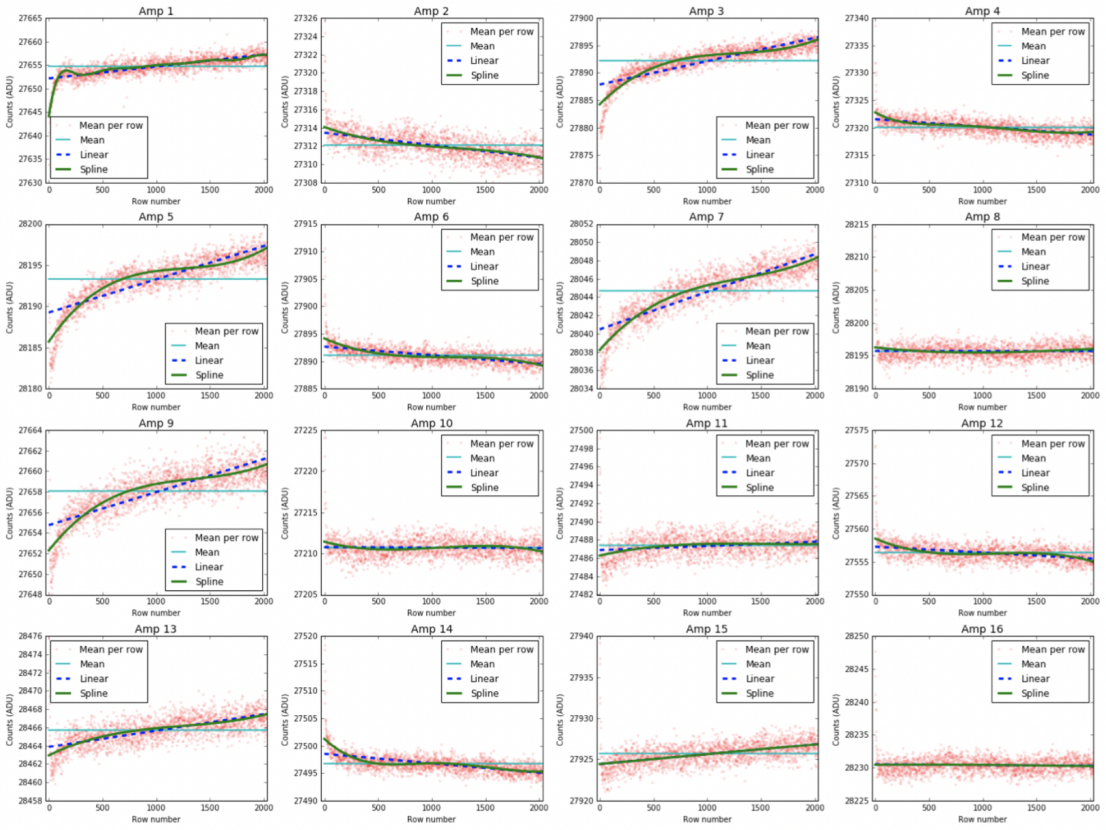


Figure 3.7: The performance of different offset models as fits to the mean-per-row in the serial overscan region of all 16 amplifiers on an e2v sensor on RTM 005. Comparing this to Fig. 3.6, it is clear that the shape of the mean-per-row curve varies across sensors even if they are from the same vendor.

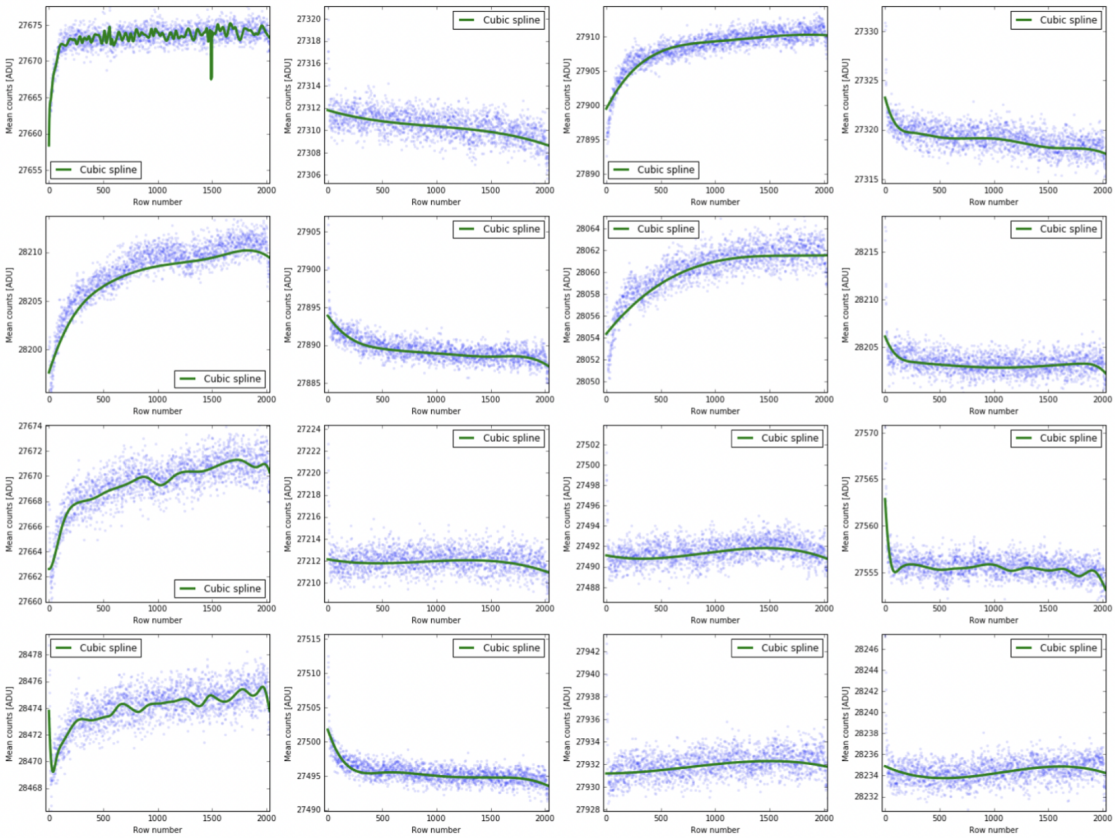


Figure 3.8: The performance of a cubic spline fit to the mean-per-row on all amplifiers on an ITL sensor on RTM 004. The fit is more sensitive to certain parameters in the spline fit, such as the smoothing factor, as seen in the first amplifier.

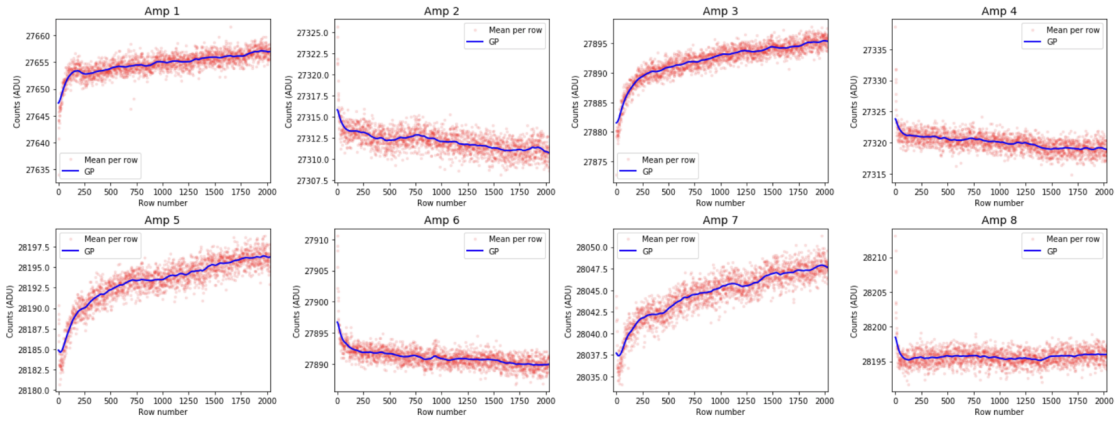


Figure 3.9: The performance of a Gaussian process fit to the mean-per-row on eight amplifiers on an e2v sensor.

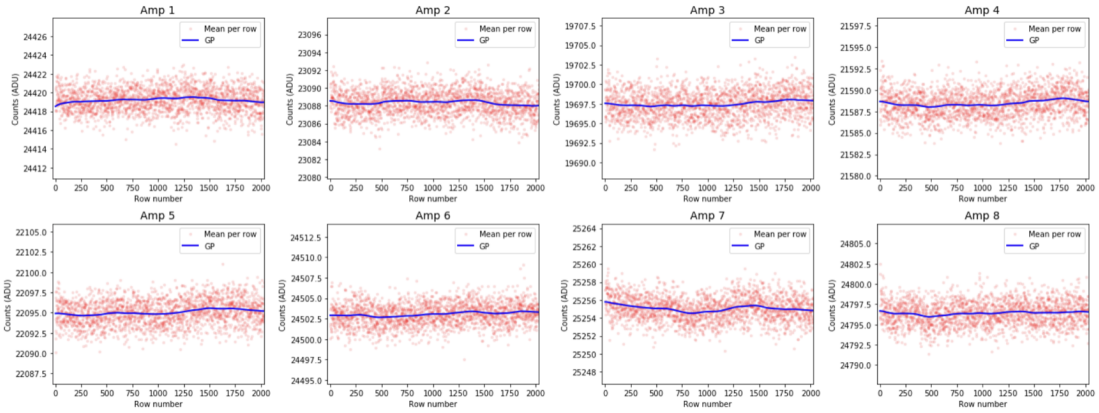


Figure 3.10: The performance of a Gaussian process fit to the mean-per-row on eight amplifiers on an ITL sensor.

While the row-by-row correction did the best at characterizing variations in the offset level, it introduced correlated noise of around $10/\sqrt{50}$ 1.4 counts per pixel. The cubic spline fit generally did well at fitting the shape of the mean-per-row without adding a significant amount of extra noise. However, its performance varied depending on the sensor vendor as it tended to perform better on e2v sensors, as well as on the specific parameters of the spline fit. Results of the cubic spline fit for an ITL CCD on RTM 004 is shown in Figure 3.8. In some cases, a cubic spline would tend to overfit the data for certain amplifiers, as in the top- and bottom-left panes of Fig. 3.8 but when the smoothing factor was increased it would come at the cost of losing peaks or dips in the mean-per-row, particularly in the first 50 or

so rows. Modeling the offset as a Gaussian process showed an improvement, but at the cost of higher complexity, as well as not being able to properly quantify the introduction of noise (see Figs. 3.9-3.10).

I modified the `eotest` package to allow the option to apply a mean-per-row or cubic spline offset correction in addition to the existing mean and polynomial fit functions. After carefully studying these methods and applying them to images from TS8, it was decided that the best way to proceed in terms of the current testing software would be to use the mean-per-row method because, even though it introduces slightly more noise, it is significantly more accurate in the lower rows and would suffice for general performance testing. This became the default offset-correction method for all electro-optical performance tasks in `eotest`.

3.3.2 Bias level

Once the offset level is subtracted and the image is trimmed to remove the overscan region, the next step is to remove the bias level. The bias level corresponds to the pixel-to-pixel variations in the read noise in an image. This structure varies from amplifier to amplifier, as well as across a single amplifier on a CCD. To correct for the bias, a "super bias" is generated for each amplifier by stacking many bias frames that have been offset-corrected and trimmed. This super bias is then used to de-bias other exposures. I added the functionality in `eotest` to be able to stack a set of images according to a statistic (for example, median-stacking or taking a clipped mean of the stacked images), as well as a method to create a super bias file, which generates a FITS file containing a super bias for each amplifier for a given raft, sensor and run number (see Figure 3.11).

Next, I performed a study of how well the super bias corrected for the bias level. This initially required a visual confirmation of each bias image that would be included in the super bias as a function of raft, sensor, and run number. Referring to an offset-corrected

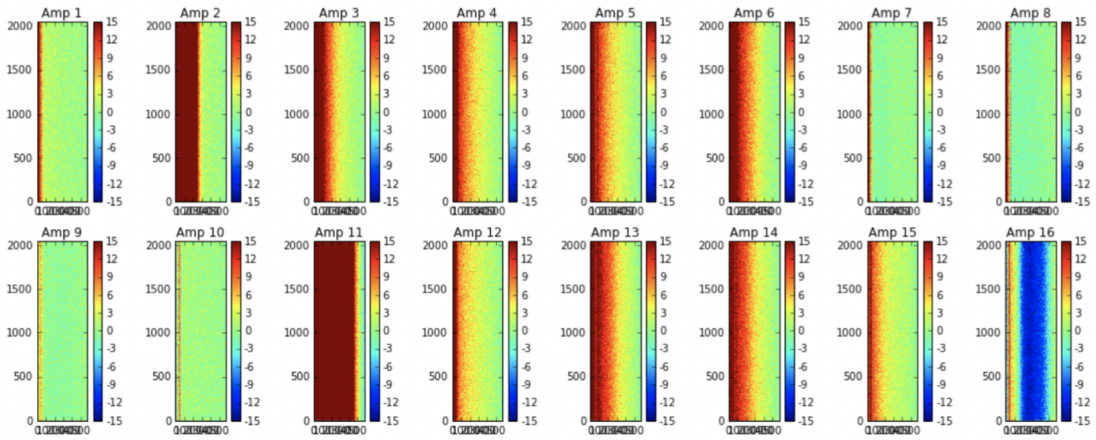


Figure 3.11: The output generated by the "super bias file" method added to the `eotest` package. Each super bias corresponds to an amplifier on sensor S00 on RTM 004. Rows are shown on the y-axis and columns are on the x-axis, with the number of counts per pixel shown in the color bar to the right of each image.

and de-biased image will as a corrected image from this point forward, all bias images taken during each acquisition mode, which included flat fields, superflats, dark images, quantum efficiency, and an Fe55 source, were verified by plotting serial and parallel projections of each corrected image. This amounted to the number of counts as a function row and column in every corrected bias frame (see Fig. 3.12). The mean and standard deviation over all corrected bias frames were also plotted as a function of row and column. The mean over all corrected bias images is expected to be zero and the standard deviation is expected to be approximately constant. Bias structure would therefore appear as the mean deviating from zero or the standard deviation varying as a function of row or column.

Most of the corrected bias images were consistent with no structure, however there was some peculiar behavior in the standard deviation as a function of row where I observed random oscillations that varied slowly in time (bottom left panel of Fig. 3.12). This is something I did not have a chance to investigate thoroughly, but one potential cause to be investigated is banding, which has been observed in some ITL sensors as a row-wise effect where one can see bands going across the image that vary in bias level. Projecting the bias in this way also revealed an effect called persistence (see Fig. 3.13). Several of the corrected bias frames

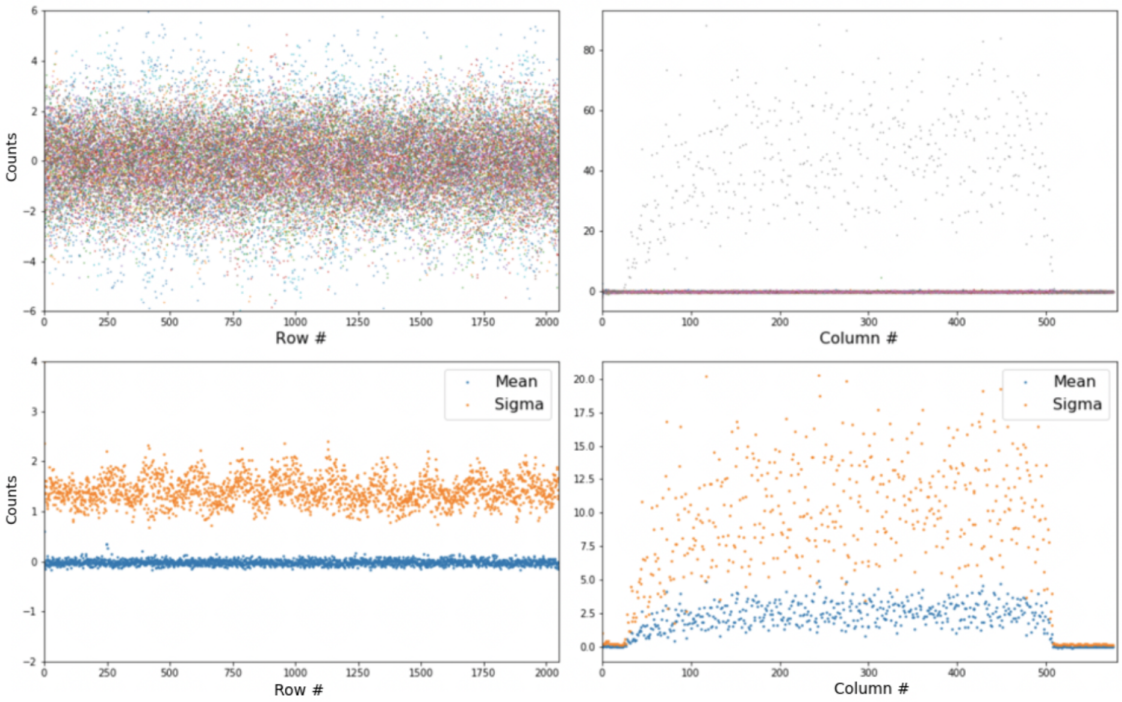


Figure 3.12: Projections of corrected bias images across the rows and columns for 18 bias images taken during the flat pair acquisition are shown in the top row. The bottom row shows the mean and standard deviation over all 18 images, where the oscillation in the mean is clearly seen.

taken during the flat-field acquisition registered counts as high as 80 ADUs. A flat-field is an image taken when the sensors are uniformly illuminated. Since flat-field sequences at TS8 involve taking flats at increasingly longer exposure times, they register more light the longer they are exposed. Bias images are taken after every two or three flats in this sequence. Plotting the timestamp of each flat and bias image showed the outlier bias image as the final bias frame taken during this acquisition, as seen in Fig. 3.14. Looking at a projection plot of the flat taken just before this bias image showed the flat to have saturated the sensors. It turned out that the excessively high mean value in the outlier bias image was due to charge that had persisted from the previous flat image due to insufficient clearing of the CCD. In this case, more time was needed between parallel pixel transfers during readout.

Once all bias images that showed anomalous behavior were removed, I created a super bias for all amplifiers on a specific sensor from 50 bias images. To study the performance of these

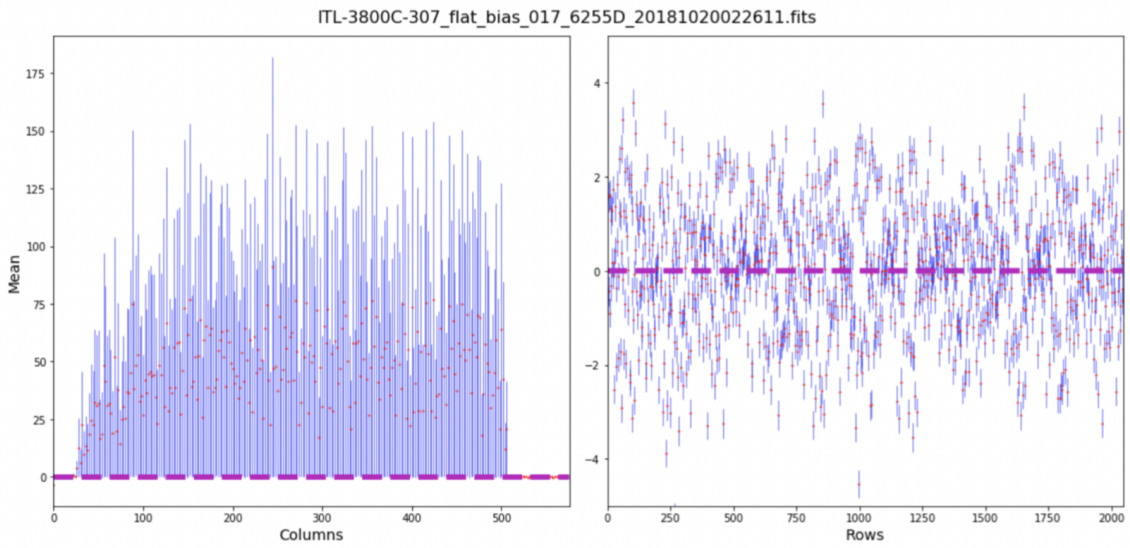


Figure 3.13: A bias frame taken during the flat pair acquisition that has persisting charge from a saturated flat taken immediately before it. This is seen here as an increased value in the mean over all rows for each column in the image.

super biases, I conducted a 'ratio test' using images taken from the superflat acquisition. Superflats were taken in two modes, with low superflats taken to have around 1000 counts, and high superflats having around 50,000 counts.

The ratio I used to study the super biases was defined as the sum over all corrected low superflats divided by the sum over all corrected high superflats:

$$R = \frac{\sum_{images} LSFLAT - superbias}{\sum_{images} HSFLAT - superbias} \tag{3.1}$$

Because this is essentially a ratio of images, it is immune to any effects like bad pixels or quantum efficiency, which would be divided out. I made a series of two-dimensional histograms of this ratio against the super bias for a number of ITL and e2v rafts. Since the same super bias image is being subtracted from the numerator and the denominator, we would expect the mean of the ratio to be around $1000/50{,}000 = 0.02$ counts, with a standard deviation of around 0.008. It is also expected that the histogram of the super bias image

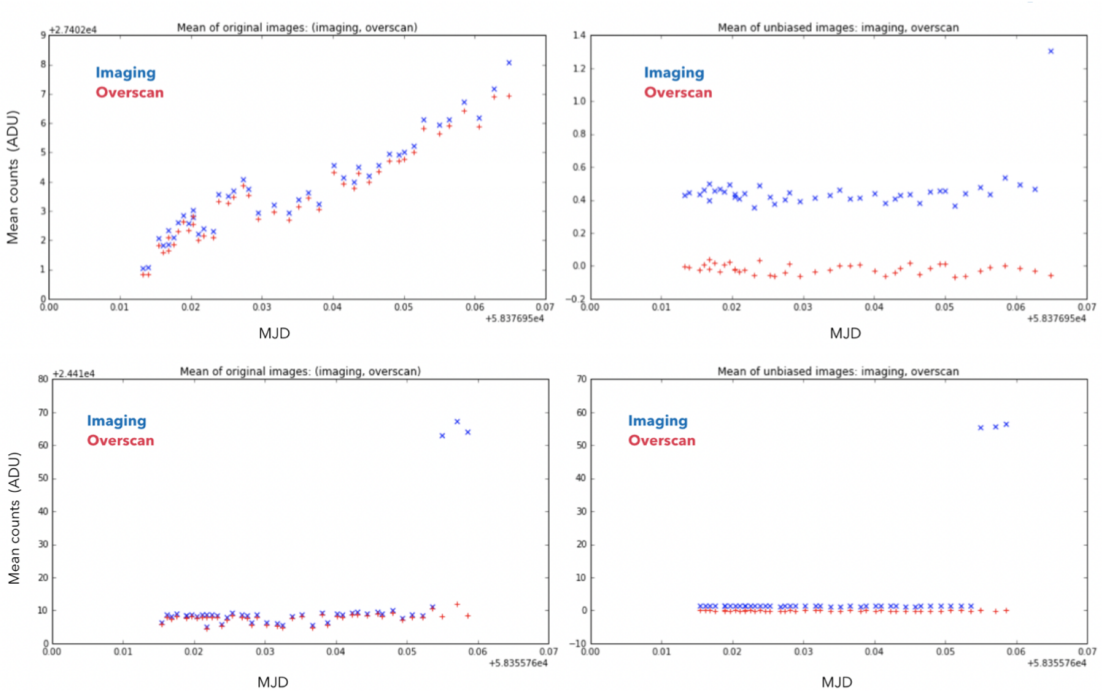


Figure 3.14: An example of persisting charge in bias frames taken on a single amplifier on two different sensors: sensor 00, amplifier 1 on RTM 012 (top), and sensor 00, amplifier 1 on RTM 004 (bottom). The mean counts in the imaging (blue) and overscan (red) regions are plotted on the y-axis against the Modified Julian Date (MJD) on the x-axis.

would have the highest density around zero counts, since the super bias is offset-corrected and should only contain pixel to pixel variations in the noise level.

The results showed a surprising amount of structure in the super biases, particularly in the ITL sensors (see Figs. 3.15-3.17). For example, there were "clumps" of points in the super bias that were concentrated in circular lobes at various intervals along the super bias axis. There were also large concentrations of points in the super bias that had values that deviated significantly from zero. On the ratio axis, there were also separate clumps of points, most of which were still centered around 0.02, but some of which had slightly lower or higher mean values. After mapping these questionable pixels onto their physical location on the sensor, most either corresponded to bad columns or to the first few columns of an amplifier. This behavior clearly indicates that something is not correct in the way the super bias is being subtracted, and that there may be pixels that should be masked but are being excluded by

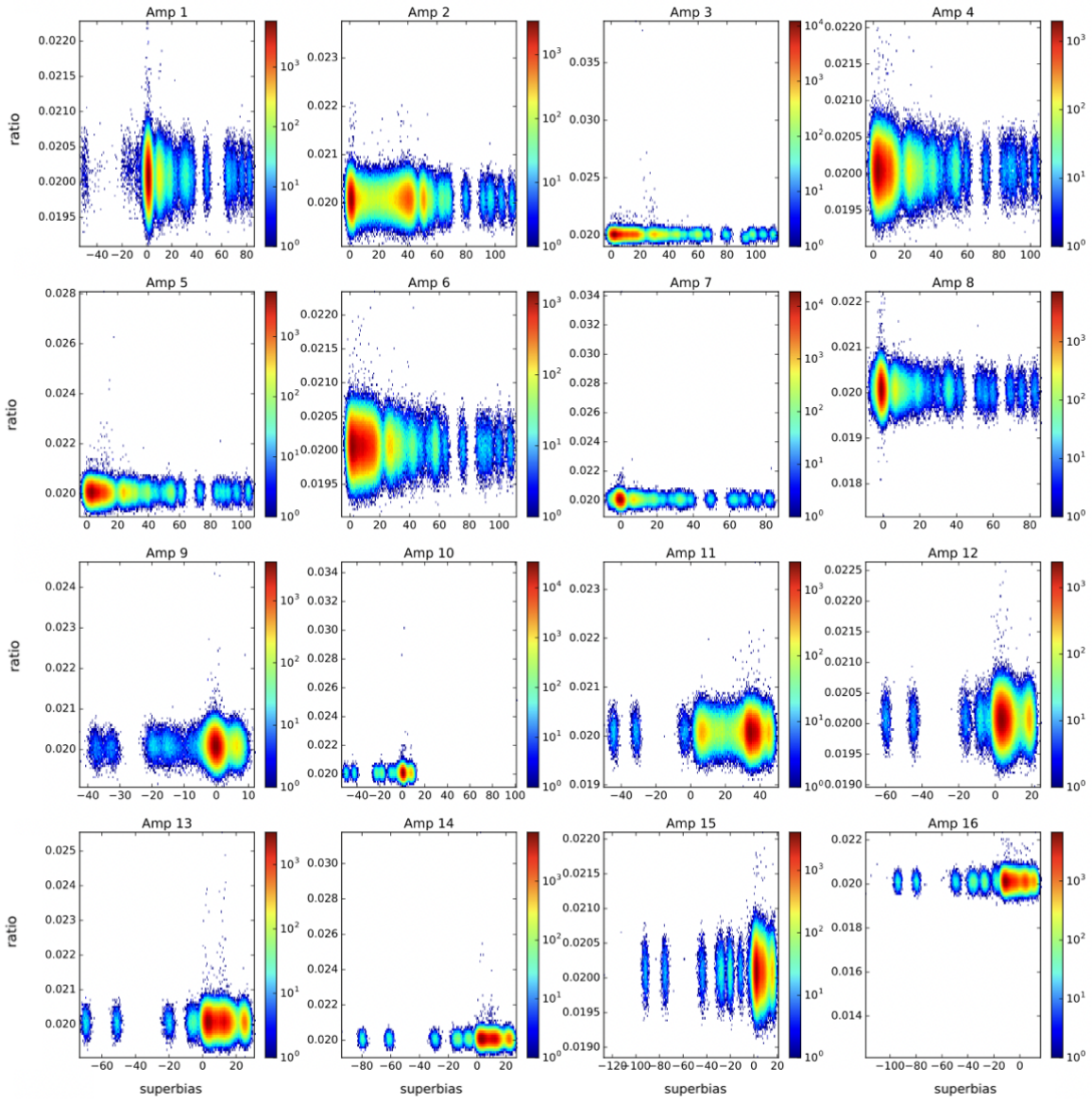


Figure 3.15: Two-dimensional histograms of the number of counts in each pixel in the ratio image plotted against the number of counts in each pixel in the super bias for all amplifiers on sensor 00 on RTM 014.

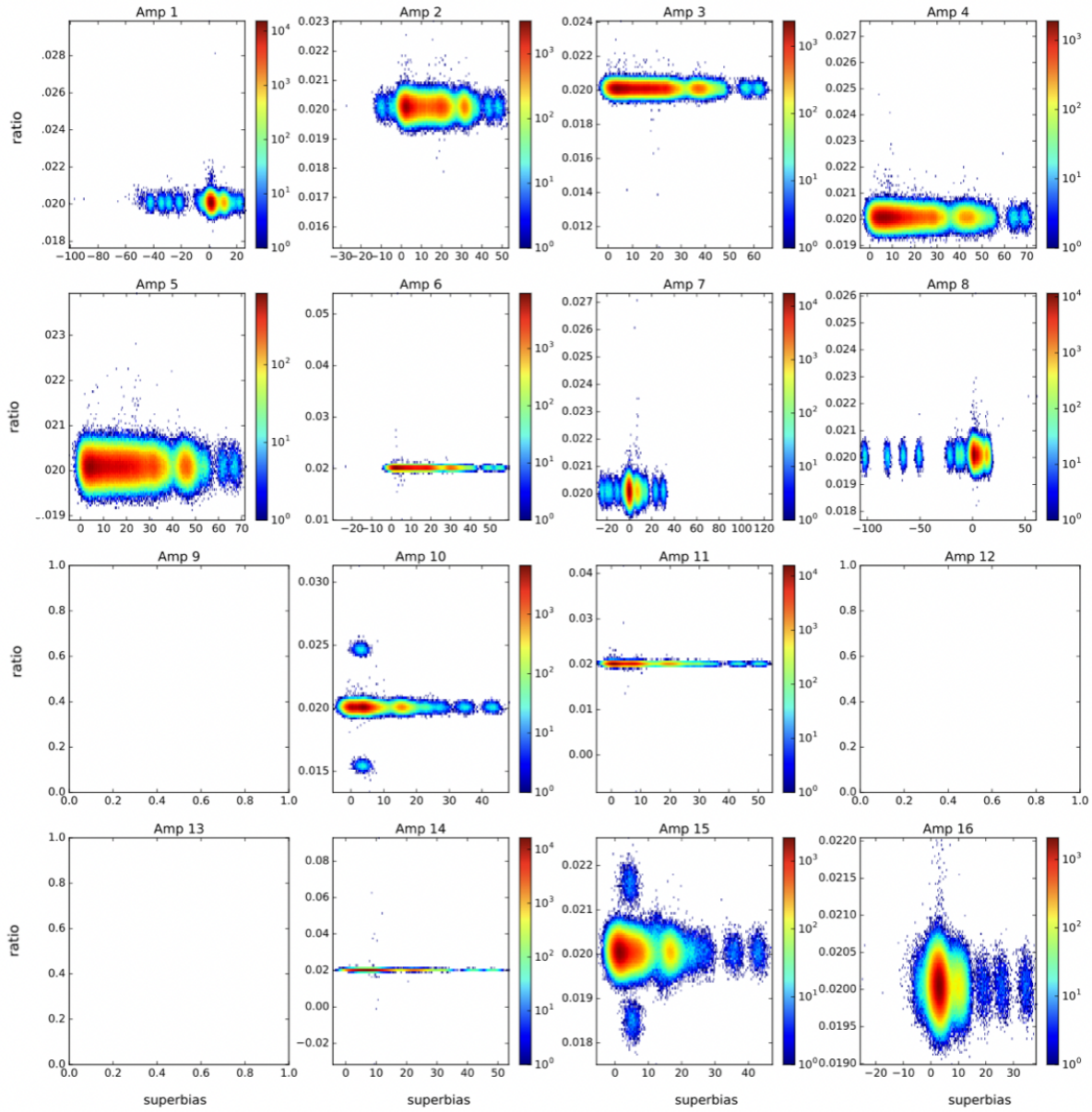


Figure 3.16: Two-dimensional histograms of the number of counts in each pixel in the ratio image plotted against the number of counts in each pixel in the super bias for all amplifiers on sensor 20 on RTM 014.

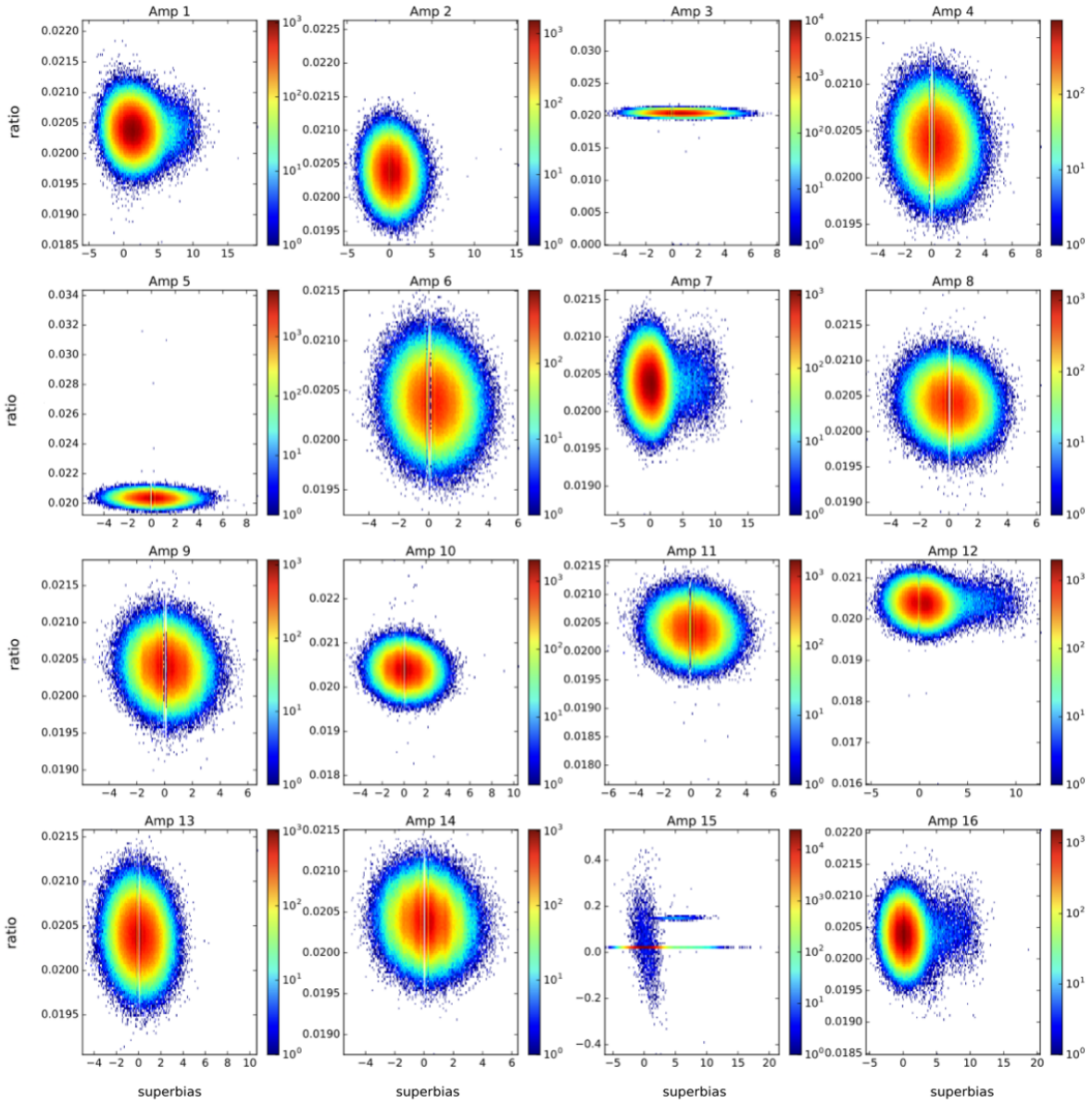


Figure 3.17: Two-dimensional histograms of the number of counts in each pixel in the ratio image plotted against the number of counts in each pixel in the super bias for all amplifiers on sensor 01 on RTM 010.

the current masking selection process in the `eotest` code. The e2v sensors were in general more well-behaved, except for a vertical gap around a super bias value of zero, where pixels in the super bias were not being plotted. Since these super bias images were masked, the gaps do not seem to be due to improper masking. One possible course of investigation is to see whether stacking the super bias using a clipped mean rather than the median would fix this issue.

Lastly, I studied another issue known as bias trending, which was observed by plotting the mean of the imaging section of all bias images in a run over time. Theoretically, a bias image should only contain the offset level, the bias level and the read noise, the last two of which are sub-dominant. This should make the mean stable around 20,000 counts. However this was not the case and there were significant fluctuations in the mean. My work showed that doing a proper bias and offset correction resolved this instability.

3.4 Modeling sky brightness in ImSim

3.4.1 ImSim image simulator

The success of cosmological surveys depends in large part on the generation of realistic simulations. Synthetic images are used to test software infrastructure such as data processing and analysis pipelines in preparation for the survey, allowing for the study of systematic effects as well as performance testing. Since there is only one observable universe and cosmological parameters cannot be repeatedly measured under exactly identical conditions, we also need simulated data to test different models of the universe once on-sky data is available.

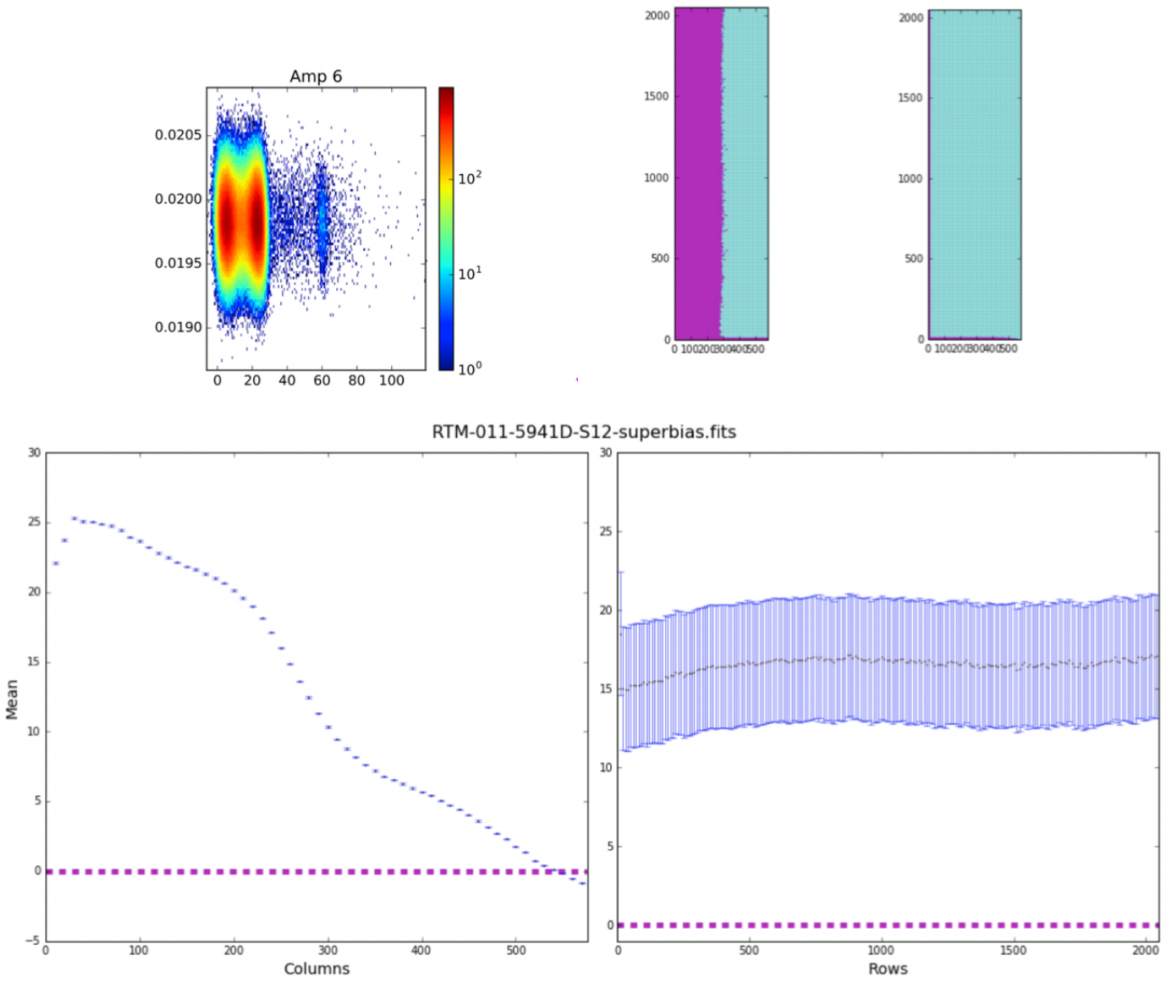


Figure 3.18: Mapping anomalous pixels in the ratio test to their physical location on amplifier 6 of sensor 12 on RTM 011. The magenta regions of the amplifier on the top right correspond to two separate cuts: values of the superbias between 15 and 25, and superbias counts above 35. The bottom two plots show projections of the mean over the rows and columns.

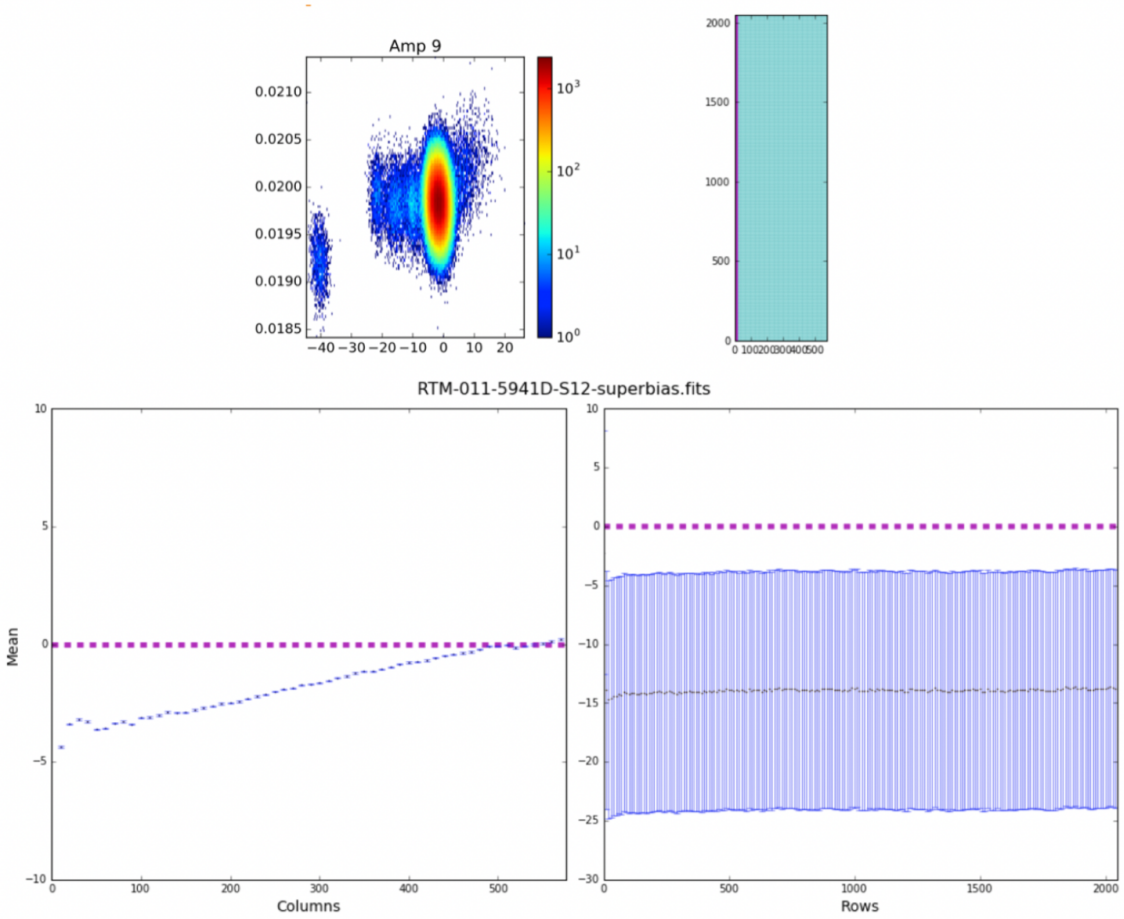


Figure 3.19: Mapping anomalous pixels in the ratio test to their physical location on amplifier 6 of sensor 12 on RTM 011. The magenta regions of the amplifier isolate regions of the ratio plot on the top left with superbias counts less than -10. The bottom two plots show projections of the mean over the rows and columns.

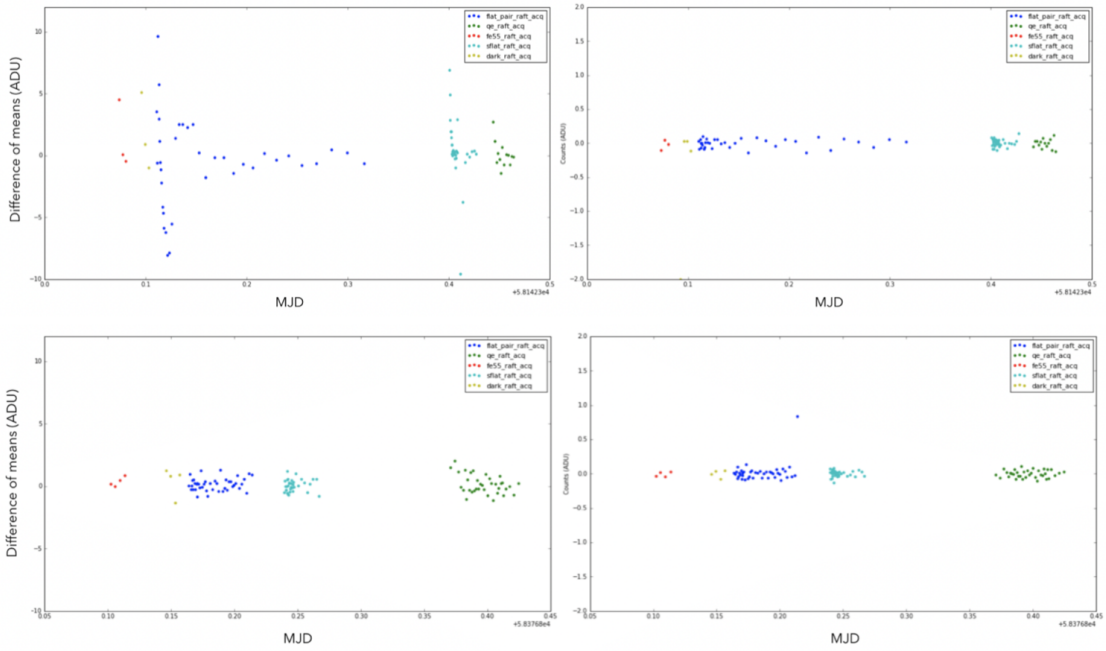


Figure 3.20: The difference between the mean of all pixels in sequential bias images taken during each acquisition mode for sensor 00, amplifier 1 on RTM 005 (top) and sensor 00, amplifier 1 on RTM 012 (bottom). The mean value of a bias image should not change significantly from exposure to exposure, however there is still some instability. Doing a proper bias and offset correction corrects for this trending, as seen in the right column.

`imSim` is a software package developed to fill the need for high-fidelity simulations of Rubin exposures. It is built on top of the GalSim software library to render astrophysical sources as well as the night sky. It also uses telescope and hardware specifications from LSST software libraries describing the telescope PSF and the Rubin focal plane such as CCD geometry and electronics readout. `imSim` takes this information as input, as well as descriptions of source positions, spectral energy distributions, magnitudes, cosmological shears and morphologies to produce raw pixel-level data resembling actual data from the Rubin telescope.

Each visit on the sky uses a catalog of sources provided by the `CatSim` package (Connolly et al. 2010, 2014) and observing conditions given by the OpSim database[1]. The observing metadata, such as sky brightness and atmospheric seeing, is used by `imSim` to calculate the sky background level.

[1] `https://github.com/lsst/sims_featureScheduler`

The initial configuration in `imSim` was to calculate the sky level at the center of the focal plane and then use this value for every pixel in the image. This was reasonable as a first-pass, however there was a need to investigate whether this approximation was sufficient enough for capturing variations in the sky level across the focal plane. I studied whether there was a significant enough level of sky variation to warrant a more detailed approach by comparing the existing implementation to a model where the sky level would vary per chip as opposed to maintaining a constant value over every sensor.

To study variations in sky counts per chip I modified the LSST `skymodel` module, based on the model[2] used by the European Southern Observatory[3], to give a different sky background for each individual chip. This involved creating a new function that returned the name and pixel coordinates of the centers of each chip. I added a new `chipName` parameter to the existing sky noise model and converted each pixel coordinate on a chip to its sky coordinates in (ra, dec). Finally, the on-sky chip center positions were passed to the `skymodel` module to estimate the number of photons incident on the chip for a 30 second exposure (corresponding to two back-to-back 15 second exposures for LSST). Finally, I implemented this new sky model into the `ImSim` package.

The results in the r-band are shown in Fig. 3.21 for relatively good observing conditions where there is very little contamination from the moon. Each point in the plot represents an individual sensor on the LSST focal plane. The color corresponds to the mean sky level per chip in electrons per exposure. Because sky noise is Poisson distributed, the scale for variations across the focal plane is given by the square root of the number of counts at its center. This scale was used to compare the implementation of a per chip sky model versus a model that uses a single sky level for the entire focal plane.

[2] `https://github.com/lsst/sims_skybrightness`
[3] `https://www.eso.org/public/`

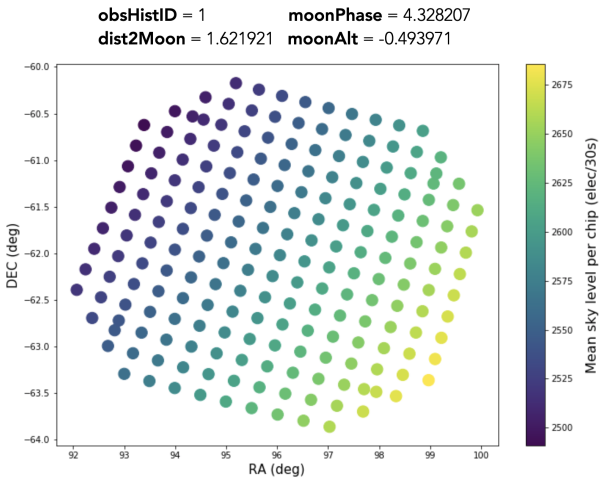
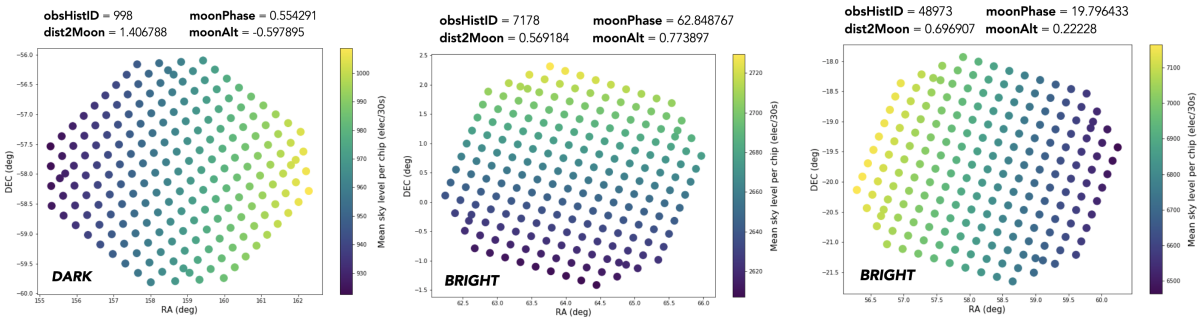


Figure 3.21: Sky level in electrons per 30-second exposure for a single ship with little moon contamination. The observing parameters are obsHistID (a unique observation identifier for this chip), moonPhase (the phase of the moon at the time of the observation, with 0=new, 100=full), dist2Moon (the angular distance to the moon in radians), and moonAlt (the moon altitude in radians).



(a) Dark (b) Bright (c) Bright

Figure 3.22: The sky level for three different observing conditions.

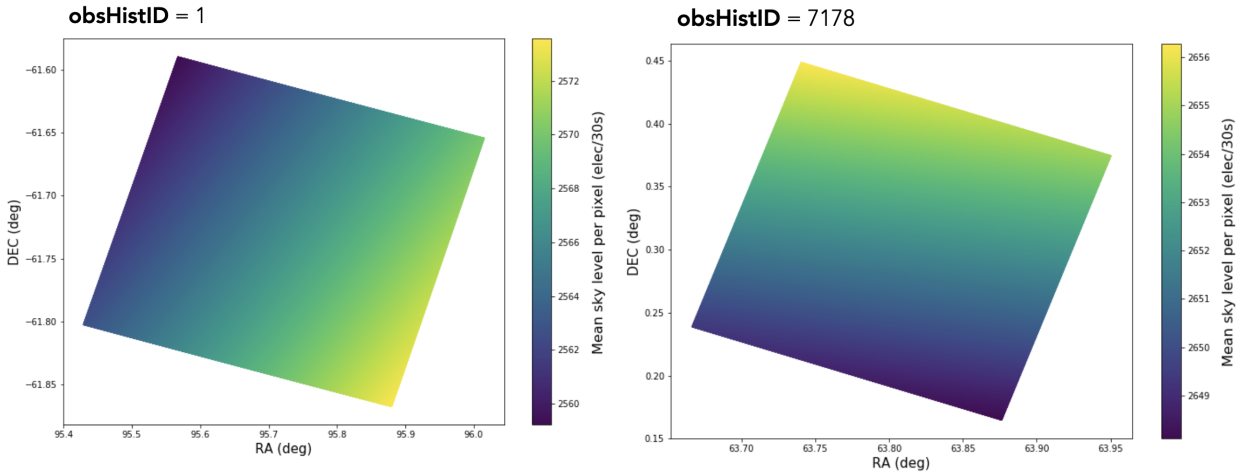


Figure 3.23: The sky model when varying the sky level per-pixel for two separate observations on the same chip, showing per-pixel variation to be sub-dominant.

In this case, the per chip variation is around 175 sky counts, which is significant compared to the spread around the number of counts at the focal plane center, which is approximately $\sqrt{2600} = 50$. This is based on the uncertainty on the mean number of counts in a single exposure, given by $\sqrt{N}$, where N is the number of counts at the center of the focal plane. Fig. 3.22 shows three more examples in both bright and dark conditions. Once again, in all three cases it is clear that the range in sky counts across the focal plane per chip is dominant over the scale for variations based on the number of counts at the center.

A natural follow-up was to investigate whether it was necessary to add the sky level per-pixel as opposed to per chip. Fig. 3.23 shows the per-pixel sky counts in the r-band for two different observations. It's clear that the per-pixel variation is sub-dominant and that accounting for the sky level based on the center of each pixel is sufficient.

3.5 Conclusion

The work I completed during my award period primarily focused on analyzing bias images, which are zero second exposures taken with the shutter closed. One would think these are

fairly well understood images, however my work showed that they still possess a great deal of structure. My work highlighted this structure and explored various ways of characterizing it so that it can be corrected for when performing image calibrations. I also did extensive studies on the structure of super biases, which are used to correct for variations in the noise level in CCD images. I wrote a module that can be run on a single sensor to show how bias subtraction performs for each amplifier on the CCD. This task can be used on data from any raft to identify problematic amplifiers, sensors or rafts.

The work I conducted for my offset-correction studies was eventually integrated into the camera testing software as part of the suite of electro-optical verification tests that are run on test stand data. This work could also be useful for developers in LSST Data Management, who are writing the software pipeline that will be used to process LSST data. I also wrote a module for the ratio test to be included in the `eotest` package that will be run to generate sensor-level ratio plots in order to test bias subtraction performance.

`imSim` is a software package used to simulate realistic images of galaxies that captured by LSST. In order to make these synthetic images as faithful as possible, it is necessary to not only model the photometric output of observed galaxies, but also contamination due to the sky. The initial configuration of `imSim` modeled the sky brightness level to be the same value across the focal plane. I changed the functionality of `imSim` to make the sky level vary per chip. I verified that making the sky level vary per-pixel was unnecessary, as the spread of the sky level per pixel was significantly less than the spread around the number of counts at the center of the focal plane.

Chapter 4

Magnitude and color distributions of DESI targets for generating mocks

4.1 Introduction

DESI will extract spectra from over 35 million astrophysical sources to measure the expansion rate over the past 11 billion years. Prior to obtaining spectra, it will need to assign objects to each of its 5,000 optical fibers by pre-selecting targets from three photometric surveys, collectively referred to as the DESI Legacy Imaging Surveys[1] (Dey et al. 2019). These wide-field surveys provide photometry that is both uniform and spatially dense, ultimately pushing to depths around two magnitudes deeper than the Sloan Digital Sky Survey[2] (SDSS). Compared to SDSS, the Legacy Surveys detected over 15 times the number of $z > 0.5$ galaxies and over 200 times more $z > 1.0$ galaxies by reaching a 5σ depth of $z \approx 22.8$ AB mag. The utility of these surveys goes beyond target selection for DESI, as they can be used in tandem with overlapping spectroscopic data in a variety of contexts such as photometric

[1] `https://www.legacysurvey.org`
[2] `https://www.sdss.org`

redshift estimation through the use of spectroscopic priors, or to study dark matter halos by cross-correlating spectroscopic and imaging maps.

4.1.1 DESI Legacy Imaging Surveys

The Legacy Surveys completely overlap with DESI's approximately 14,000 deg^2 footprint, which is divided into separate 9,900 deg^2 and 4,000 deg^2 patches in each of the Northern and Southern Galactic Caps, respectively (Fig. 4.1). The Dark Energy Camera (DECam; Flaugher et al. (2015)), which is installed on the 4-meter Blanco telescope at the Cerro Tololo Inter-American Observatory in Chile, imaged the entire SGC footprint and part of the NGC in the (grz) bands as part of the Dark Energy Camera Legacy Survey[3] (DECaLS). The majority of DESI's northern footprint was imaged at Kitt Peak National Observatory. Photometry in the g- and r-bands was taken by the Prime90 instrument on the Bok Telescope as part of the Beijing-Arizona Sky Survey[4] (BASS; Zou et al. (2019)), and additional z-band photometry was provided by the Mayall z-band Legacy Survey[5] (MzLS) using the Mosaic-3 camera installed on the Mayall Telescope. These optical data were combined with mid-infrared photometry from the Wide-field Infrared Survey Explorer[6] (WISE) satellite to help differentiate between targets and to characterize galaxy morphologies. See Table 4.1 for more info about survey coverage and depths.

4.1.2 Target classes

DESI will observe luminous red galaxies (LRG), emission line galaxies (ELG), quasars (QSO) and bright galaxies (BGS) to study cosmic expansion and the growth rate of structure

[3] `https://www.legacysurvey.org/decamls/`
[4] `https://www.legacysurvey.org/bass/`
[5] `https://www.legacysurvey.org/mzls/`
[6] `https://www.nasa.gov/wise`

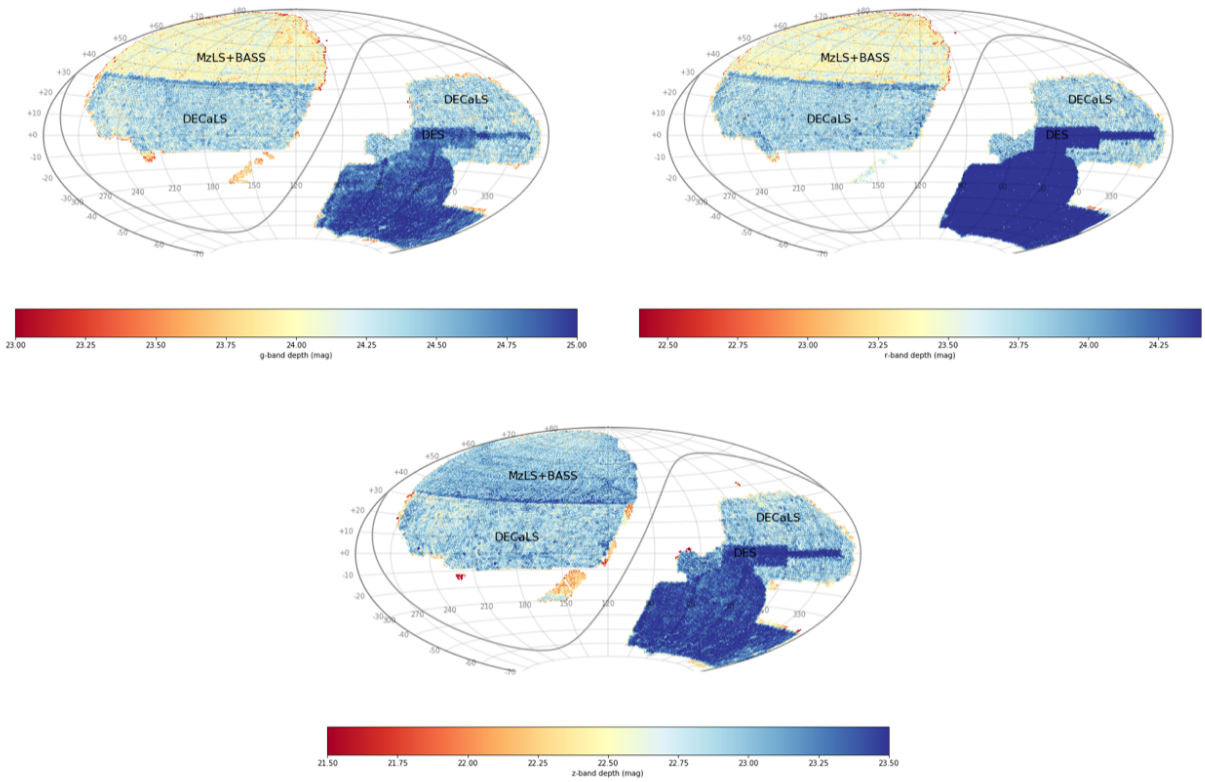


Figure 4.1: The footprint of each Legacy Survey as of Data Release 9 (DR9), color-coded by the extinction-corrected depth. The Galactic plane is shown as the solid black line. Figure is from `https://www.legacysurvey.org/status`.

Table 4.1: DESI Legacy Surveys: Areas and Depths from Dey et al. (2019).

Survey Name	Telescope/Instrument	Galaxy Depth (mag)			Area (deg^2)	Location
		g	r	z		
DECaLS	Blanco/DECam	23.72	23.27	22.22	9000	NGC($\delta \leq +32°$)+SGC
		g	r			
BASS	Bok/90Prime	23.48	22.87		5000	NGC($\delta \geq +32°$)
		z				
MzLS	Mayall/Mosaic-3	22.29			5000	NGC($\delta \leq +32°$)

through the imprints of baryonic acoustic oscillations (BAO; see Section 1.4.2) and redshift-space distortions (RSD; see Section 1.5.2). Fainter, higher-redshift targets such as LRGs, ELGs, and QSOs will be observed during a Dark Time Program, when there is little light contamination from the moon. Low redshift BGS targets will fall under the Bright Time Program, where they will be able to achieve DESI's signal-to-noise requirements despite significant moon illumination.

During its Bright Galaxy Survey, DESI will obtain spectra of 10 million low redshift galaxies below $z < 0.4$ to study BAO and RSD via galaxy clustering. At slightly higher redshifts, between $0.3 < z < 1.0$, it will target at least 8 million LRGs. These galaxies are characterized by a strong 4000Å break, a signature of the absorption of high-energy radiation by metals in cooler stellar atmospheres that have ceased star formation. ELGs will comprise at least half of all DESI targets. These bluer, star-forming galaxies will probe the $0.6 < z < 1.6$ universe and will reveal identifiable emission lines such as the OII doublet. Finally, DESI will target QSOs in an even higher redshift regime. QSOs that lie between $0.9 < z < 2.1$ will be used as direct tracers of dark matter, and beyond $z > 2.1$, Lyα quasars will study large scale structure through the absorption of neutral hydrogen in the intergalactic medium (see Table 1.2 for target densities and redshift ranges).

Objects from these target classes were selected from the imaging surveys based on their specific photometric properties. After reduction with the `legacypipe` pipeline[7], the `desitarget`[8] package was used to choose targets from the processed images based on a set of unique magnitude and color cuts that captured distinguishing features of each class. The baseline fiducial cuts for each class in the Dark Time Program are shown in Tables 4.2-4.5. The resulting target catalogs were released progressively as a series of data releases. The studies in this chapter were performed on imaging from Data Release 5[9] (DR5).

[7] `https://github.com/legacysurvey/legacypipe`
[8] `https://github.com/desihub/desitarget`
[9] `https://www.legacysurvey.org/dr5`

Table 4.2: LRG selection cuts in the Northern DESI footprint.

Selection	North
Non-stellar cut	$(z - W1) > 0.8 \times (r - z) - 0.6$
Faint limit	$z$fiber < 21.61
Low-z cut	$(g - W1 > 2.97)$ OR $(r - W1 > 1.8)$
Double sliding cuts	$((r - W1 > (W1 - 17.13) \times 1.83)$ AND $(r - W1 > (W1 - 16.31))$ OR $(r - W1 > 3.4)$

Table 4.3: LRG selection cuts in the Southern DESI footprint.

Selection	South
Non-stellar cut	$(z - W1) > 0.8 \times (r - z) - 0.6$
Faint limit	$z$fiber < 21.6
Low-z cut	$(g - W1 > 2.9)$ OR $(r - W1 > 1.8)$
Double sliding cuts	$((r - W1 > (W1 - 17.14) \times 1.8)$ AND $(r - W1 > (W1 - 16.33)$ OR $(r - W1 > 3.3)$

Table 4.4: ELG selection cuts for both the Northern and Southern footprints.

Selection	All
Bright cut	$g > 20.0$
Faint cut	$g$fiber < 24.1
Blue cut	$rz > 0.15$
Star/low-z cut	$gr < 0.1 + 0.5 \times rz$
OII cut (fiducial)	$gr < -1.20 \times rz + 1.3$
OII cut (low-priority)	$gr > -1.20 \times rz + 1.3$ AND $gr < -1.20 \times rz + 1.6$

Table 4.5: QSO selection cuts for the entire DESI footprint.

Cuts
$17.5 < r < 22.7$
$grz > 17.0$
$(g - r) < 1.3$
$-0.4 < (r - z) < 1.1$
$W1 - W2 > -0.3$ AND $W1 - W2 > -0.4$

4.2 Generating inputs for mocks

Aside from target selection, another use for these imaging catalogs is to generate realistic synthetic spectra, or mocks. DESI used the `desisim`[10] software to simulate DESI-like spectra for a variety of purposes, including identifying and accounting for systematics, testing science pipelines, as well as survey strategy and design. `desisim` uses a function called `select_mock_targets` to randomly assign fluxes and colors to galaxies based on representative distributions of those features. I generated flux and color distributions for `select_mock_targets` using mixtures of Gaussians to characterize galaxies from each target class that passed selection cuts. These models were stored as `.fits` files in `desitarget` and sampled from using `desisim`.

4.2.1 Gaussian mixture models

A mixture model is a form of density estimation that can be used to create generative models of multi-dimensional data. This is done by assuming that the data come from a mixture of multivariate Gaussian components whose parameters are optimized using an expectation-maximization (EM) algorithm. The model is then used to assign probabilities to new data based on whether they belong to a given component, and to generate new synthetic data.

4.2.2 Model selection

Gaussian mixture models (GMM) are useful in characterizing sub-populations within a data set in order to generate new random data that resembles the input data. Properties of the data set are learned through training a number of models with a varying number of components in an unsupervised way. The optimal number of components in the model

[10] `https://github.com/desihub/desisim`

Table 4.6: Number of sources for each morphological type in DR5.

Type	Number of sources
PSF	371,088,269
REX	222,184,611
DEV	22,036,854
EXP	61,380,049
COMP	3,066,121

balances model complexity (the number of degrees of freedom or components in the model) and model performance (a maximum-likelihood estimate of the data given the model). The model parameters for a GMM are the means and covariances of the Gaussian components, which are also included in the model optimization process. A common model selection method is the Bayesian Information Criterion (BIC), which is expressed analytically as

$$BIC = k \ln n - 2 \ln \hat{L}, \tag{4.1}$$

where k is the number of model parameters, n is the sample size, and $\ln \hat{L}$ is the maximum log-likelihood of the model. Once the BIC is evaluated for a set of models, the model with the lowest BIC is selected as the most optimal. Flux and color distributions were generated using a Gaussian mixture model with the optimal number of components given by the minimum BIC.

4.2.3 Magnitude and color distributions in DR5

Rather than model sizes, shapes, colors and magnitudes simultaneously, colors and magnitudes were modeled separately for each morphological type. This is a reasonable way to factorize the model and is motivated by the observation that different morphological components often correspond to different galaxy sub-populations. Morphological classifications were assigned according to each target's surface brightness profile, as inferred by the Trac-

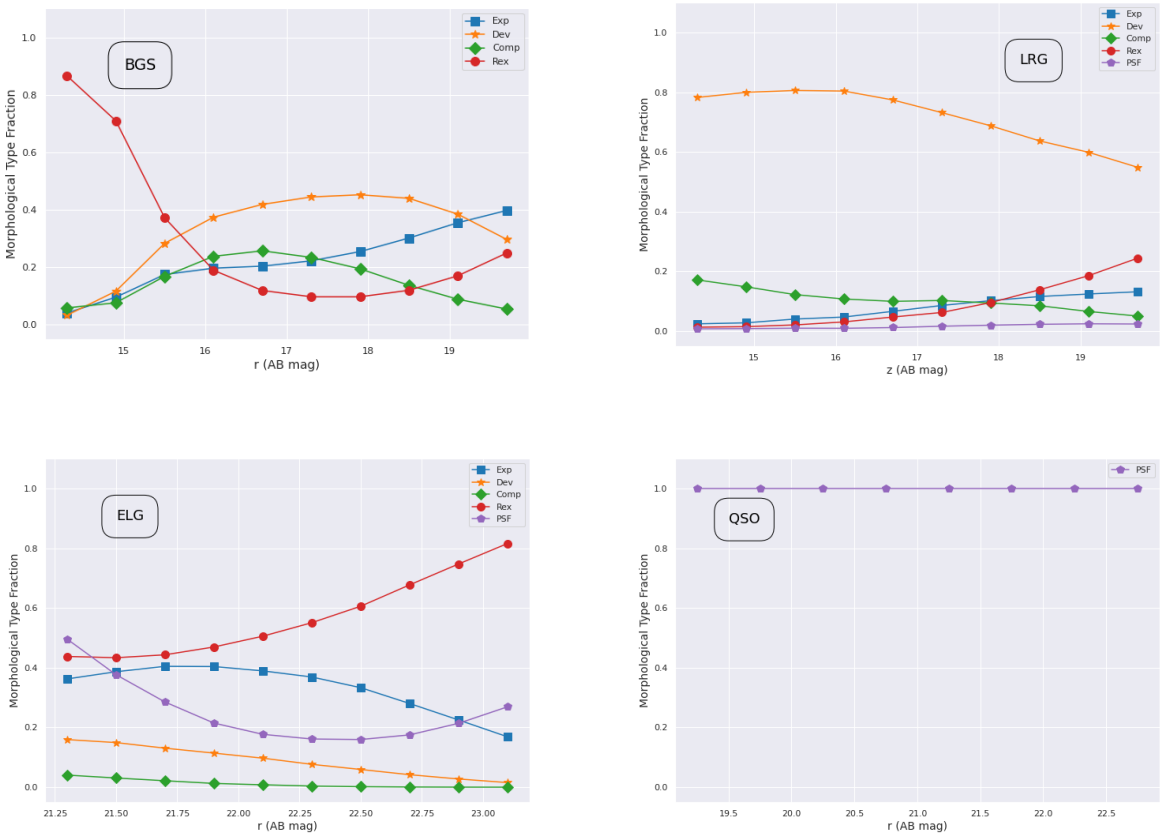


Figure 4.2: Morphological type fraction for each class in the DR5 sample.

tor[11] fitting algorithm. The morphological types used in DR5 are point sources (PSF), round exponential galaxies with a variable radius (REX), de Vaucouleurs profiles (DEV), exponential profiles (EXP), and composite profiles (COMP), which are a combination of deVaucouleurs and exponential profiles. Table 4.6 shows the distribution of types in DR5 in the absence of selection cuts, and the plots in Fig. 4.2 show the fraction of targets belonging to each morphological classification in ten bins equally spaced in magnitude.

The features used to build the GMMs were based on the magnitudes and colors used for selection cuts in each class. GMMs were first trained on the photometry, and then the model parameters and weights were used to generate synthetic distributions of the data with the same features. The data were split into a training set used to train the model, and a validation set used to assess the performance of the model on data that were not used for training.

Plots of the optimal number of Gaussian components determined by the BIC are shown in Figure 4.3 and Figure 4.8. Figures 4.4-4.9 show the 2-dimensional histograms for training and validation sets are shown in Figures 4.3 for LRGs with deVaucoulours profiles and QSOs. The sampled data from the trained model are overlaid in green. This process was extended in a similar fashion for the remaining combinations of classes and morphologies. The contours in the histograms show the 1-, 2-, and 3-σ contours containing 39.3%, 86.5%, 98.9% of the data, respectively.

4.2.4 Extreme deconvolution

The GMM implementation in Section 4.2.3 performs a density estimate of the underlying color-magnitude distribution, but it does not properly deal with noise. Extreme deconvolution Gaussian mixture modeling (XDGMM) incorporates a noise model to simultaneously

[11] `http://thetractor.org`

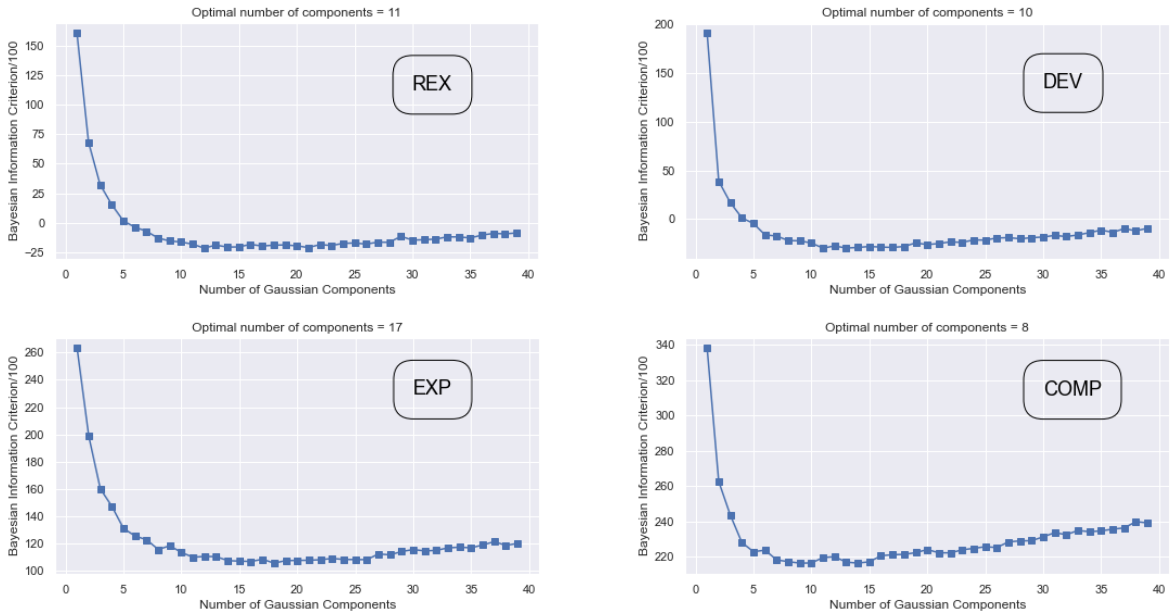
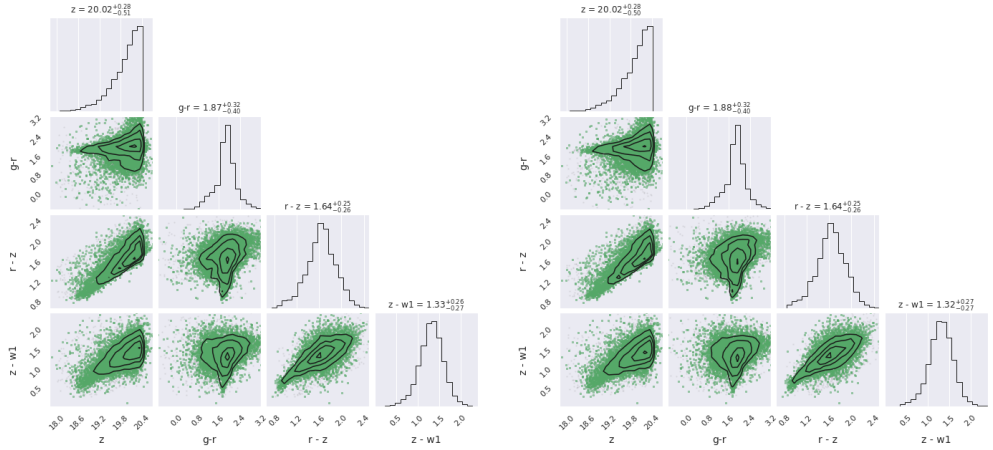
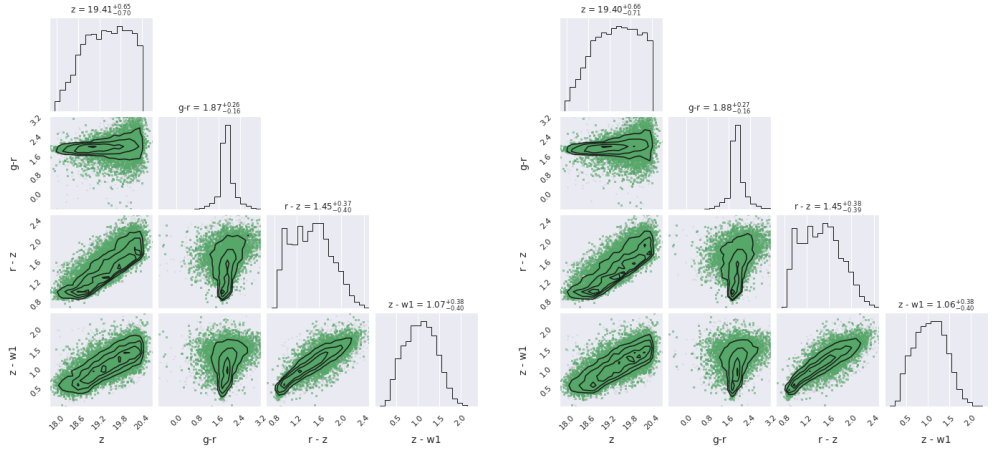


Figure 4.3: Bayesian Information Criterion for each LRG morphological type, with the optimal number of components corresponding to the lowest value of the BIC.



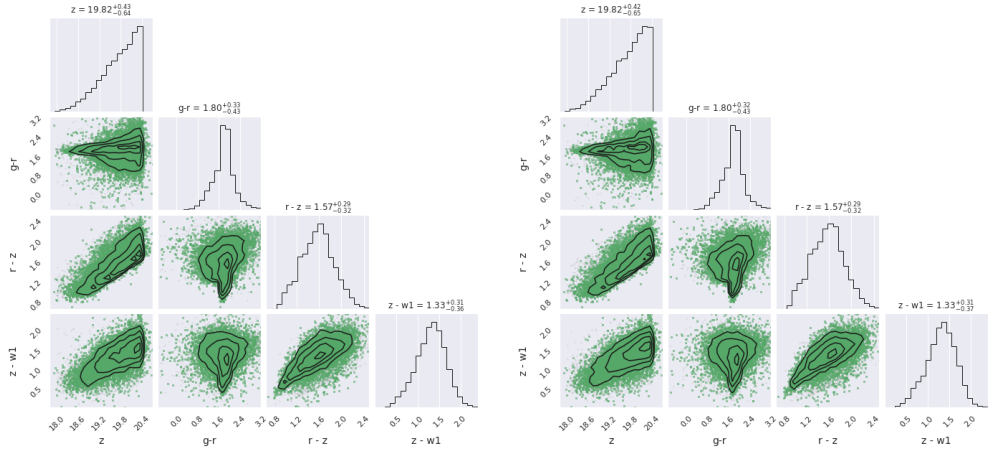
(a) Training (b) Validation

Figure 4.4: Corner plot showing 1-, 2-, and 3-σ contours of the training and validation data in black, along with sampled galaxies from the GMM in green, for LRGs with REX morphological profiles.



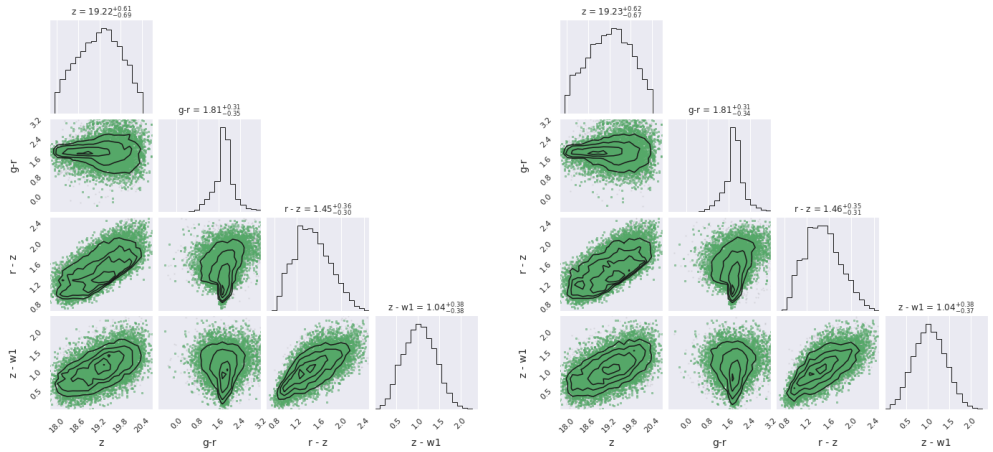
(a) Training (b) Validation

Figure 4.5: Corner plot showing 1-, 2-, and 3-σ contours of the training and validation data in black, along with sampled galaxies from the GMM in green, for LRGs with DEV shape profiles.



(a) Training (b) Validation

Figure 4.6: Corner plot showing 1-, 2-, and 3-σ contours of the training and validation data in black, along with sampled galaxies from the GMM in green, for LRGs with EXP profiles.



(a) Training (b) Validation

Figure 4.7: Corner plot showing 1-, 2-, and 3-σ contours of the training and validation data in black, along with sampled galaxies from the GMM in green, for LRGs with COMP profiles.

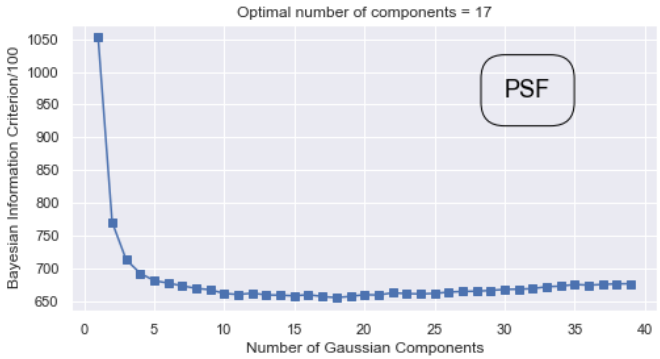
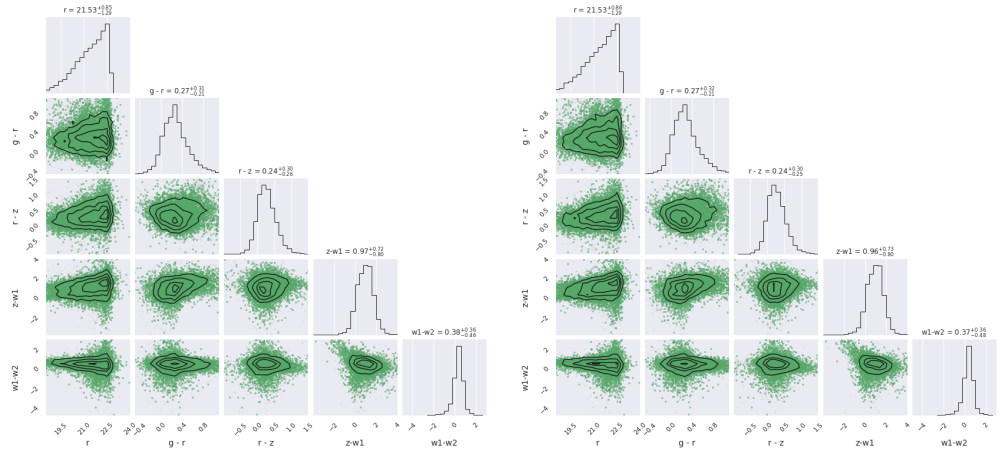


Figure 4.8: Bayesian Information Criterion for QSOs.



(a) Training (b) Validation

Figure 4.9: Corner plot showing 1-, 2-, and 3-σ contours of the training and validation data in black, along with sampled galaxies from the GMM in green, for QSOs.

model the data and the noise . This is done by convolving the underlying distribution with a covariance matrix containing the variance of each Gaussian in the mixture model, resulting in an error-deconvolved estimate of the underlying distribution function. The reader may refer to Bovy et al. (2011), Ivezić et al. (2014), and Holoien et al. (2017) for additional background on this technique.

I investigated the impact of using extreme deconvolution in reconstructing the underlying flux distribution for each target class. The results for LRGs with exponential and de Vaucouleurs profiles, and for QSOs are shown as dotted red lines in the histograms in Figures 4.10-4.11. XDGMM was performed using the `astroML`[12] package (Vanderplas et al. 2012). Components in each mixture model were assumed to be uncorrelated, and therefore a diagonal covariance matrix was used for each application. The blue histograms represent the error-deconvolved distribution, with Gaussian noise added back in. This was done by drawing errors for each data point from a Gaussian with zero mean and variance equal to its true error. This should most closely match the true distribution, shown in green. Finally, the histogram representing density estimates from a GMM, without extreme deconvolution, is shown in orange.

[12] `https://www.astroml.org`

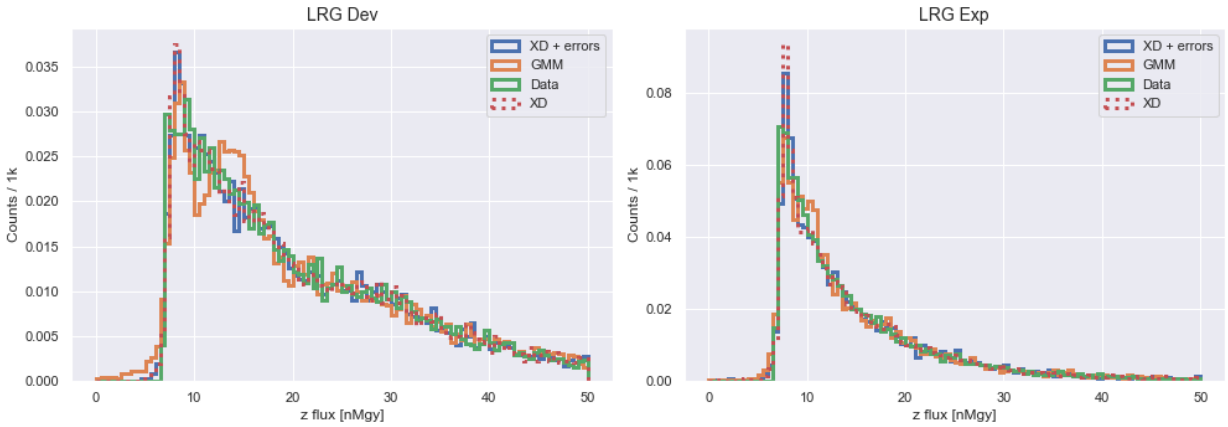


Figure 4.10: The distribution of true z-band fluxes for two LRG morphological types, with the true fluxes in (green), compared to those sampled from a straightforward GMM (orange), from an XDGMM (dotted red), and from an XDGMM with the errors added back in (blue).

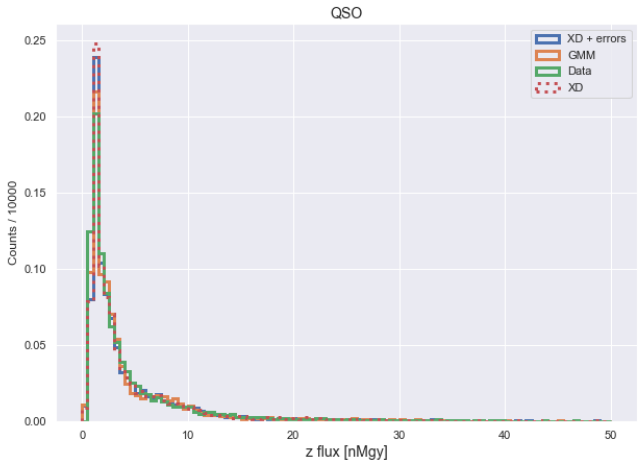


Figure 4.11: The distribution of true z-band fluxes for QSOs, with the true fluxes in (green), compared to those sampled from a straightforward GMM (orange), from an XDGMM (dashed red), and from an XDGMM with the errors added back in (blue).

The results did not show any significant advantage to using extreme deconvolution to model these flux distributions. For LRGs, the GMM seemed to perform slightly better closer to the peaks of the true distribution, while the XD results more tightly traced the data away from the peaks. For QSOs, the difference was negligible. Given these results, it was decided that a straightforward GMM was sufficient to use over the XDGMM technique.

4.3 Conclusion

DESI target selection involves pre-selecting a list of galaxies and quasars to observe from imaging over its 14,000 $\deg^2$ footprint. Photometry from the Legacy Surveys was not only used for targeting, but as input for generating mock spectra for DESI simulations. A Gaussian mixture model was used to characterize color and magnitude distributions as a function of target class so that they could be sampled from for simulations. Extreme deconvolution was used to de-convolve the noise from the data, however there was added benefit of doing so over a straightforward mixture model.

Chapter 5

Simulating eBOSS spectra with `specsim`

5.1 Introduction

One of the primary requirements for DESI is to be able to generate mocks in order to compare and validate our models with the data we observe (see Chapter 4). This not only includes creating faithful representations of source spectra, but accurately modeling the degrading effects of the sky, atmosphere, telescope, and instrument on these spectra. This section is concerned with modifying an existing software package intended to simulate the DESI instrument to instead produce realistic eBOSS spectra. The existence of having real eBOSS data on-hand provided a useful way of validating the results of `specsim` and debugging any potential issues.

5.2 The `specsim` software package

`specsim`[1] (Kirkby et al. 2020) is a software package developed by David Kirkby to simulate the response of a multi-fiber spectrograph. Although it was originally created to generate realistic synthetic spectra for the DESI instrument, it can be reconfigured to simulate any fiber-fed spectrograph, provided that the accompanying instrument parameter specifications and data are provided.

A single simulation consists of three components: a source spectrum, a model of the sky and atmosphere, and a model of the instrument. A schema of where each component comes into play is shown in Fig. 5.1, which shows the journey of photons as they are emitted from a galaxy to the moment they are read out by the detector.

5.2.1 Astrophysical source

Spectra of astrophysical sources are simulated individually, beginning with a spectral energy distribution (SED) in its rest frame, $s(\lambda)$, and its morphological profile. Configuration parameters for a source profile include the bulge/disk fraction and shape parameters such as the half-light radius, position angle, and the ratio of the semi-minor to semi-major axis. The source profile is assumed to be independent of its SED, and will come into play later on in Section 5.3.3 to account for fiberloss when modeling the effects of the instrument.

5.2.2 Sky model

Given a source, the next component of the `specsim` pipeline is the sky model, which incorporates the effects of the atmosphere. The three elements that make up the sky model are

[1] `https://specsim.readthedocs.io/`

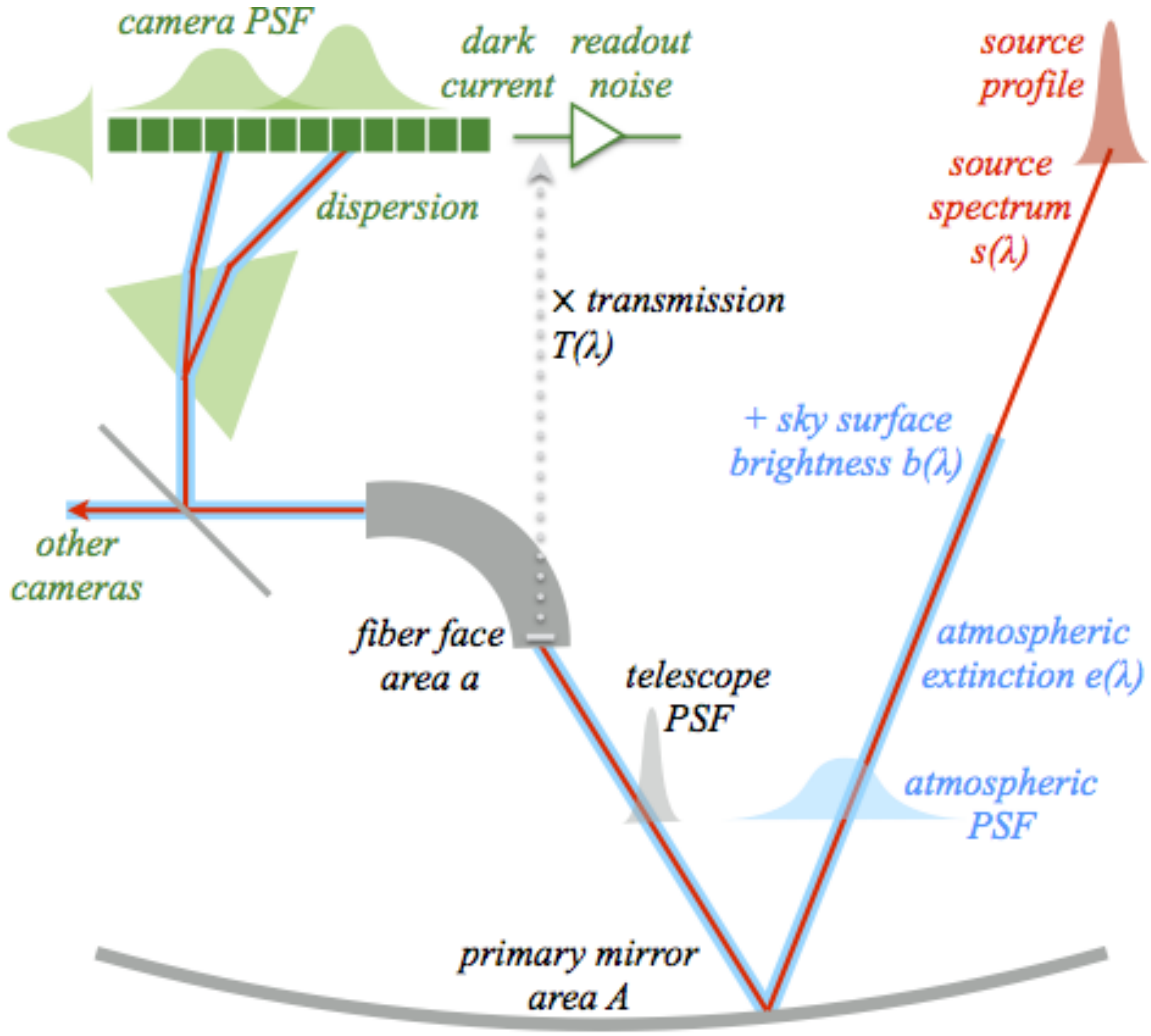


Figure 5.1: Schematic showing the main components of `specsim`. Light originating from a source above the atmosphere is modeled as it travels through the atmosphere, telescope optics, spectrograph, and is finally registered by the CCDs. Figure is from Kirkby et al. (2020).

Table 5.1: `specsim` main simulation parameters.

Parameter	Component	Description
Δt	Instrument	Exposure time
X	Atmosphere	Observing airmass
λ_i	Configuration	Fine wavelength grid
$s(\lambda)$	Source	Source SED
$b(\lambda)$	Atmosphere	Sky surface brightness
$e(\lambda)$	Atmosphere	Atmospheric extinction
A	Instrument	Primary unobscured area
a	Instrument	Fiber entrance area
$f_S(\lambda)$	Source & Instrument	Fiberloss fraction
$T_i(\lambda)$	Camera	Transmission throughput
$d_i(\lambda)$	Camera	CCD row width
$\sigma_i(\lambda)$	Camera	CCD resolution
n_{ip}	Camera	CCD trace width
$I_{dk,i}$	Camera	Sensor dark current
G_i	Camera	CCD gain
$\sigma_{ro,i}$	Camera	Readout noise

a sky emission spectrum, the atmospheric point spread function, or PSF, and atmospheric extinction.

The sky model used for both DESI and eBOSS configurations is virtually identical. The sky emission surface brightness $b(\lambda)$ for the DESI configuration contains data for three different types of sky spectra corresponding to dark, bright and grey conditions, whereas the eBOSS sky emmission spectrum is just the DESI dark sky extrapolated to cover the wider wavelength range of the eBOSS spectrograph.

The atmospheric PSF results in the blurring of an image of a source as its photons encounter varying indices of refraction when traveling through different layers of the Earth's atmosphere. Differences in temperature, pressure, density and molecular composition in each layer, as well as turbulence in the air, cause photons to slightly deflect from their original paths, resulting in a "smeared" image. Seeing is defined as the full width at half maximum (FWHM) of a radial profile of a point source, typically a star, from its peak value.

The effects of atmospheric seeing in `specsim` can be applied in one of three ways. The fastest mode assumes a fixed 1.1" seeing for each source type (LRG, ELG, QSO, etc.), and does not account for any additional information about surface brightness profiles. A slower, but more flexible mode uses `GalSim`[2] (Rowe et al. 2015), a software package that simulates high-fidelity images of astrophysical sources, to convolve the source profile with the atmospheric PSF. `GalSim` requires additional information in the `specsim` configuration file such as FWHM_{ref} and λ_{ref}, which are used to estimate the PSF as a function of wavelength in the following way:

$$\mathrm{FWHM}(\lambda) = \mathrm{FWHM}_{ref}\Big(\frac{\lambda}{\lambda_{ref}}\Big)^{-0.2}. \tag{5.1}$$

The reference values used in the configuration file are $\mathrm{FWHM}_{ref} = 1.1$ arcsec and $\lambda_{ref} = 6355$Å. Since `GalSim` uses a Moffat profile to model the PSF, a fixed value for the β parameter (with a typical value of 3.5), which determines the shape of the PSF, must also be specified. The Moffat distribution is predominantly used for modeling PSFs as it is more successful at capturing tails compared to a Gaussian or Lorentzian distribution.

The final component in the sky model is atmospheric extinction, which could be caused by Rayleigh scattering of air molecules or particulate matter such as aerosols, or by telluric absorption due to the Earth's atmosphere. This is provided in the configuration file as a list of tabulated values for the extinction coefficient $e(\lambda)$ at zenith as a function of wavelength. The amount of extinction at other airmasses is determined by multiplying the extinction coefficient by the airmass at that particular pointing.

Together, the three components of the sky model are combined to model the attenuation of the source spectrum and the sky brightness in terms of the extinction factor $e(\lambda)$ and

[2] `https://github.com/GalSim-developers/GalSim`

the airmass X due to the effects of the atmosphere. The spectral flux density $f(\lambda)$ at the entrance of the telescope then becomes:

$$f(\lambda) = 10^{-e(\lambda)X/2.5}[s(\lambda) + ab(\lambda)]. \tag{5.2}$$

`specsim` allows the option to include a scattered moonlight component as part of the sky brightness spectrum, which is determined for each observation by a solar SED[3], the moon phase, the moon zenith angle and the separation angle between the telescope pointing and the moon. Verification of the `specsim` sky model used will be discussed further in Section 5.4.

5.2.3 Telescope and instrument

Telescope

Once the light from a source has completed its passage through the atmosphere, it finally begins the final leg of its journey when it enters the telescope and is registered by the detector.

Photons that are incident on the telescope are impacted by the total collecting area of the primary mirror, A, used to normalize both the source flux and the sky brightness. Unlike the source spectrum, the sky level is proportional to the size of the fiber, and therefore must also be normalized by the fiber face area a. The telescope has an optical PSF which is due to the diffraction of light by the telescope aperture and potential aberrations in the lens or mirror. In a diffraction-limited system, the telescope will have reached its theoretical limit and the only contribution to the optical PSF will be due to the finite size of the aperture. This PSF is convolved with the atmospheric PSF and the source profile to determine the

[3] `https://www.nrel.gov/grid/solar-resource/spectra.html`

fiberloss $f_S(\lambda)$, or the fraction of photons incident on the fiber face that make it through the fiber (see Section 5.2.3). The flux density of the source is transformed to give the flux in units of erg/Å up to the point just before the photons enter the camera:

$$F(\lambda) = 10^{-e(\lambda)X/2.5}\Big[s(\lambda) + ab(\lambda)\Big] f_S(\lambda) A \Delta t, \tag{5.3}$$

where Δt is the length of a single exposure. The last step after combining the impacts of the atmosphere and telescope is to simulate the effects of the camera.

Instrument

The `specsim` package is able to simulate the response of a spectrograph and one or more cameras. This involves characterizing the effects of the instrument once photons have exited the fiber, as shown in the green portion of Fig. 5.1, beginning with the dispersion $d_i(\lambda)$, indexed by camera i, as incoming light is separated into its constituent wavelengths by the spectrograph. This is also referred to as the row width, and gives the conversion from wavelength in Angstroms to size in pixels.

Once the light has been dispersed to produce a spectrum, it reaches one or more detectors, or charge-coupled devices (CCDs) (see Chapter 3), where it is registered as a function of fiber number along the vertical direction, and wavelength or pixel number along the horizontal axis. If one were to plot a distribution of where photons from a single source landed on the detector, as a function of wavelength, the CCD resolution $\sigma_i(\lambda)$ would correspond to the Gaussian sigma of the best-fit Gaussian to that distribution. Values for the trace width, which gives the thickness of the spectrum along the fiber axis, must also be provided in the configuration file. Additional values for the amount of dark current, read noise, gain and

throughput for each camera must also be specified. Further details regarding each of these parameters will be given in the following section.

The calculations used to determine the camera response are performed on a simulation grid λ_j (indexed by bin j) whose resolution is smaller than the wavelength extent of a single camera pixel by at least a factor of five. If a photon with wavelength $\lambda_j < \lambda < \lambda_{j+1}$ enters a fiber and lands in a bin of width $\Delta\lambda_j$, where

$$\Delta\lambda_j \equiv \lambda_{j+1} - \lambda_j \,, \tag{5.4}$$

then the energy E_j associated with this photon in ergs can be derived from the Planck equation:

$$E_j = \frac{hc}{\Delta\lambda_j} \,, \tag{5.5}$$

where h is Planck's constant and c is the speed of light. This can be used to convert from the flux density in Eq. 5.3 at the center of each bin to the number of photons entering the fiber N_j^γ:

$$N_j^\gamma = \frac{\Delta\lambda_j}{hc\bar{\lambda}_j} F(\bar{\lambda}_j). \tag{5.6}$$

The bin centers are given by:

$$\bar{\lambda}_j \equiv \frac{\lambda_{j+1} + \lambda_j}{2} \,. \tag{5.7}$$

Each camera's throughput $T_i(\lambda)$ is defined as the probability that a photon incident on a fiber is converted into a photo-electron and registered by the CCD readout. The resolution effects in the wavelength direction σ_i are given by a convolution matrix R_{jk}, taking us from the simulation grid indexed by j to a detection grid indexed by k. Adding the effects of camera throughput and summing over the simulation pixels gives the number of detected electrons for camera i as a function of pixels in the detection grid:

$$N_{ik}^{e} = \sum_{j} R_{jk} N_{j}^{\gamma} T_i(\lambda_j)\,. \tag{5.8}$$

Next, the dispersion function $d_i(\bar{\lambda}_k)$ is determined as a function of Angstroms in the detection grid in order to re-bin the number of detected counts from the detection wavelength grid to continuous pixel coordinates N_{ip}^{e}.

Folding in sensor effects such as the camera gain G_i and dark current $I_{dk,i}$ gives the signal in terms of the number of detected electrons:

$$N_{ip}^{det} = G_i N_{ip}^{e} + I_{dk,i} n_{ip} \Delta t\,, \tag{5.9}$$

where n_{ip} is the trace width in pixels for wavelength pixel p. The variance of the resulting signal is given by:

$$V_{ip}^{det} = N_{ip}^{det} + \sigma_{ro,i}^{2} n_{ip}\,, \tag{5.10}$$

where $\sigma_{ro,i}$ is the readout noise in pixels. The read noise is assumed to be uncorrelated between pixels.

5.3 Simulating the eBOSS instrument response

The `specsim` package was developed with the flexibility to simulate the response of any multi-fiber spectrograph, given that the accompanying configuration data is provided. There were several motivations for reconfiguring `specsim` to simulate the eBOSS instrument and cameras. Firstly, an eBOSS configuration would serve a critical purpose by allowing for the validation of `specsim` itself, since eBOSS data already exists. From a software engineering perspective, it would confirm the extent to which the package is in fact generalizeable. Although the package should in theory be able to extend to other instruments, this had never actually been done before – `specsim` had only been set up to produce synthetic DESI spectra, and it was unclear how difficult it would be to accommodate two spectrographic arms for eBOSS instead of three for DESI, or to use different wavelength binnings, among other things. Another useful application, and the primary motivation for this undertaking, was to generate realistic simulations of eBOSS spectra for the eBOSS Lyα working group.

5.3.1 Generation of mocks for Lyα studies

Baryon acoustic oscillations (BAO; see Section 1.4.2) originating from the pre-recombination universe are a cosmological "standard ruler" whose imprints are seen as peaks in the matter correlation function at the sound horizon. Measurements of the BAO scale can be measured from the clustering of galaxies for redshifts below $z < 2$, but there are simply not enough tracers at redshifts above this limit for high-precision clustering measurements. The alternative is to use opacity fluctuations due to the absorption of neutral hydrogen in the Lyα forest of background quasars. The final eBOSS data release, Data Release 16, includes all quasars observed by both BOSS and eBOSS, including 210,005 quasars above $z > 1.10$ (du Mas des Bourboux et al. 2020).

The eBOSS Lyα working group rely on faithful representations of quasar spectra to validate analysis pipelines and study potential sources of systematic effects (Farr et al. 2020). Mock quasar spectra specifically generated with the eBOSS configuration of `specsim` were used to estimate the detection efficiency of Damped Lyα systems using a convolutional neural network, ultimately revealing the purity of the quasar sample used for BAO studies (Chabanier et al. 2022).

The production of mocks begins with the generation of a Gaussian random field on a box of length 10 h^{-1}Gpc. The peaks of this physical density field are then populated with tracer quasars via Poisson sampling, based on an input bias and number density. Next, skewers for each quasar are generated by interpolating the density field along the line-of-sight, and then processed to convert from fluctuations in the density field to a transmitted flux fraction for each Lyα spectrum (du Mas des Bourboux et al. 2020). These processed skewers are the source inputs to `specsim`, as shown in the red portion of Fig. 5.1.

5.3.2 Instrument parameters

A new configuration file was produced for the eBOSS instrument and named `sdss.config`. Values for the fixed telescope parameters were added, which included the size of the primary mirror, the obscuration diameter (the size of the gap in the primary mirror where light reflected off of the secondary mirror passes through), the size of the fibers, and the field radius (the radius of the plate). The read noise, dark current and gain for each camera, which contribute to added noise in the camera, were also specified. Values for both the DESI and eBOSS telescope and camera parameters as they appear in their respective configuration files are shown in Tables 5.2-5.4.

Table 5.2: Telescope parameters for DESI and eBOSS.

Telescope Parameters	DESI	eBOSS
Field of view (sq. deg.)	8.0	3.0
Collecting area (sq. deg.)	11.4	
Primary mirror diameter (m)	3.797	2.5
Obscuration diameter (m)	1.8	0.625
Fiber diameter (μm)	107.0	120.0
Field radius (mm)	414.0	325.0
Number of fibers	5000	1000

Table 5.3: DESI camera parameters.

Camera parameters	b	r	z
Read noise (e^-/pixel^2)	3.0	2.9	2.9
Dark current (e^-/hour/pixel^2)	3.0	2.0	2.0
Gain (e^-/ADU)	1.0	1.0	1.0

5.3.3 Camera

Camera data was derived separately for the red and blue cameras using data from real eBOSS spectra. The `bossdata`[4] software package (Kirkby 2015) was used to download data from the Sloan Digital Sky Survey (SDSS) Sky Server, and was part of version `v5_9_0` of Data Release 13, the first data release of the fourth phase of SDSS (SDSS-IV). Data was pulled from 20 random plates, each observed on a different day, for all four cameras (two blue cameras, b1 and b2, and two red cameras, r1 and r2).

Wavelength binning

The first task was to generate a canonical wavelength grid for each camera. Unlike the two other wavelength grids used in `specsim` (the wavelength solution for the source spectrum, controlled by the user, and a finely binned simulation grid with a resolution of 0.1Å is used to perform calculations for the instrument response), this canonical grid is the final wavelength binning of the simulated output. It is also the common grid over which the rest of the

[4] `https://bossdata.readthedocs.io/`

Table 5.4: eBOSS camera parameters.

Camera parameters	b	r
Read noise (e^-/pixel^2)	2.0	2.75
Dark current (e^-/hour/pixel^2)	2.1	4.25
Gain (e^-/ADU)	1.02	1.66

Table 5.5: eBOSS PLATE and MJD data for the 20 plates used in the `specsim` reconfiguration.

PLATE	MJD
7027	56448
6963	56724
7301	56746
6759	56416
6002	56104
6178	56213
6626	56330
6882	56541
7389	56769
7453	56749
7517	56772
6472	56362
6660	56370
6877	56544
6970	56444
6122	56246
7456	56727
7377	56741
7454	56751
7564	56804

camera data in this section was interpolated over to be used in the configuration file. This canonical grid was derived by combining the wavelength data for the two cameras in each channel separately (b1 with b2, and r1 with r2), and generating a linear grid with equal 1Å binning with endpoints determined by the minimum and maximum wavelength values of all combined fibers on all plates. The wavelength solutions for all fibers on each camera are shown in Fig. 5.2, and the final result, shown in Fig. 5.3, displays the output grids for each camera for both eBOSS and DESI configurations. Since the eBOSS pipeline stores spectra

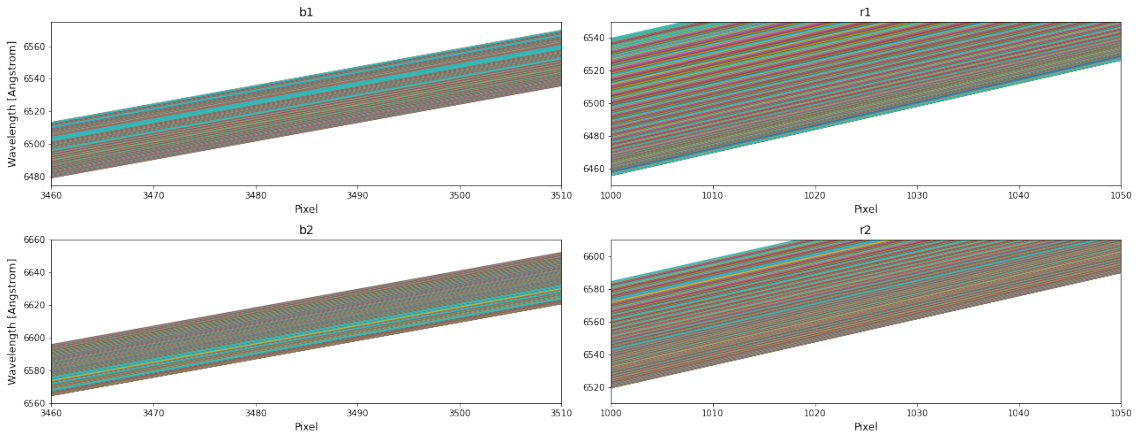


Figure 5.2: Wavelength solutions for all 500 fibers on each camera over 20 plates.

in units of log10(Å), the final output employed a similar logarithmic wavelength sampling. This is discussed further in Section 5.3.4.

Spectral resolution

The wavelength solution, dispersion and FWHM resolution for each fiber were accessed from spCFrame[5] files. These files contain flux-calibrated spectra for a single camera of a single exposure. The dispersion was determined by taking the gradient of the wavelength solution with respect to pixel number. This is often referred to as the "row width", and is shown in the top panel of Fig. 5.4 in units of Angstrom/pixel, alongside the analogous result for DESI. Obtaining the resolution along the wavelength axis required first converting the values for the Gaussian sigma along the wavelength axis from units of 10^{-4} log10(Å) to Angstroms, and then converting from a standard deviation to a FWHM by the standard conversion:

$$\mathrm{FWHM} = \sigma\sqrt{8\ln 2} \tag{5.11}$$

[5] `https://data.sdss.org/datamodel/files/BOSS_SPECTRO_REDUX/RUN2D/PLATE4/spFrame.html`

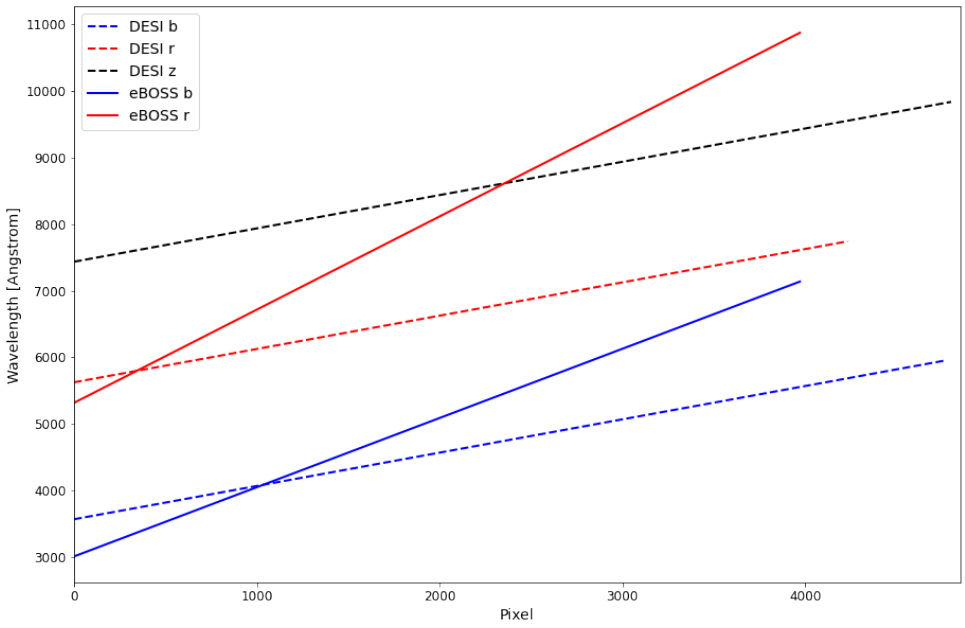


Figure 5.3: Canonical wavelength grids used by `specsim` for DESI (dashed curves) and eBOSS (solid curves).

Spatial resolution data was obtained from spFlat[6] files as the first-order corrected profile width for each fiber bundle. The Gaussian sigma for this profile, given in units of pixels, came in the form of a trace solution, or table of coefficients as a function of $x-$position on the CCD, that were then fed to a fitting function. This gave the spatial width as a function of wavelength.

Data for the dispersion, FWHM resolution, and spatial resolution were combined in the same fashion. After combining the data for each blue or red camera, these quantities were interpolated separately for all fibers over the canonical grid, and the median over all fibers and all plates was taken. The final results, along with DESI values for comparison, are shown in Figs. 5.4-5.6.

[6] `https://data.sdss.org/datamodel/files/BOSS_SPECTRO_REDUX/RUN2D/PLATE4/spFlat.html`

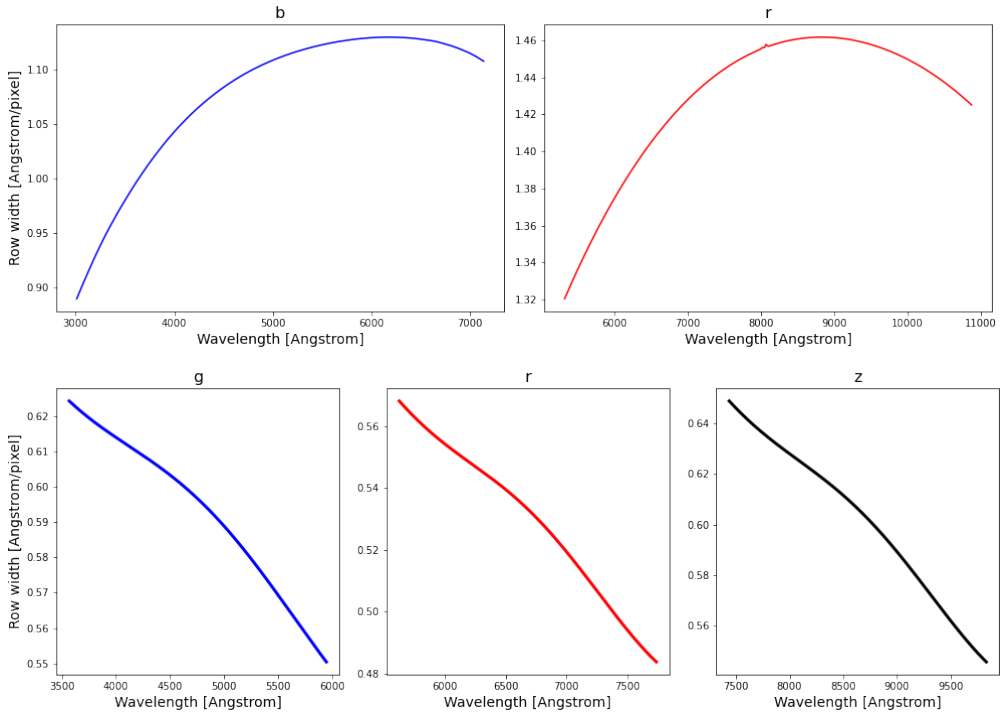


Figure 5.4: Dispersion solutions for the eBOSS (top) and DESI (bottom) configurations.

Throughput

Throughput data was initially generated for eBOSS quasars by working backwards from flux-calibrated SEDs to detected electrons. The SEDs were derived from all spectrophotometric standard stars on the plates listed in Table 5.5. Standard stars were used because, like quasars, their profiles are PSF-like. Since they are brighter, they also have a higher signal-to-noise ratio.

Figures 5.7-5.8 show the process of going from the final, calibrated flux, to the product of fiberloss and throughput for a single standard star. This product is the ratio of detected electrons on the CCD to photons entering the telescope. The top four panels trace the process of obtaining the number of photons incident on the primary mirror using data from spFrame files, which contain non-calibrated spectra for a single exposure of an individual camera. The next three panels show how to obtain the number of detected electrons using data from spCFrame files, while the last panel shows the ratio of the two quantities.

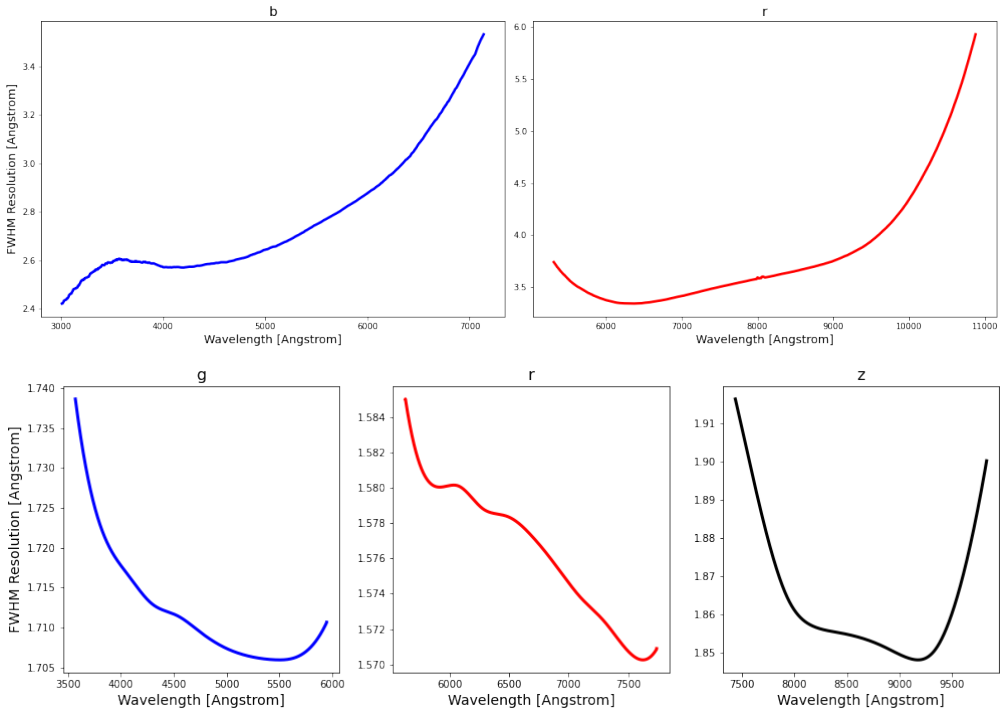


Figure 5.5: FWHM resolution for the eBOSS (top) and DESI (bottom) configurations.

We begin with a flux density in the upper left panel, which has been interpolated over the nominal wavelength grid. We then convert from a density to the number photons above the atmosphere by multiplying by the area of the primary mirror, the exposure time, and dividing by the energy of each photon at the midpoints of the canonical wavelength bins. Next we must remove the extinction correction, shown in the third panel after being interpolated over the nominal grid. This is done by multiplying the photons entering the atmosphere by the extinction to give the number of photons entering the telescope.

To derive the number of electrons detected by the camera, we start with the flux in flat-fielded electrons and remove the flat-fielding correction by multiplying by the superflat and the fiberflat (see the third row of Figs. 5.7-5.8). This result is divided by the number of photons entering the telescope to yield the product of the fiberloss and the throughput, as expressed in the last panel.

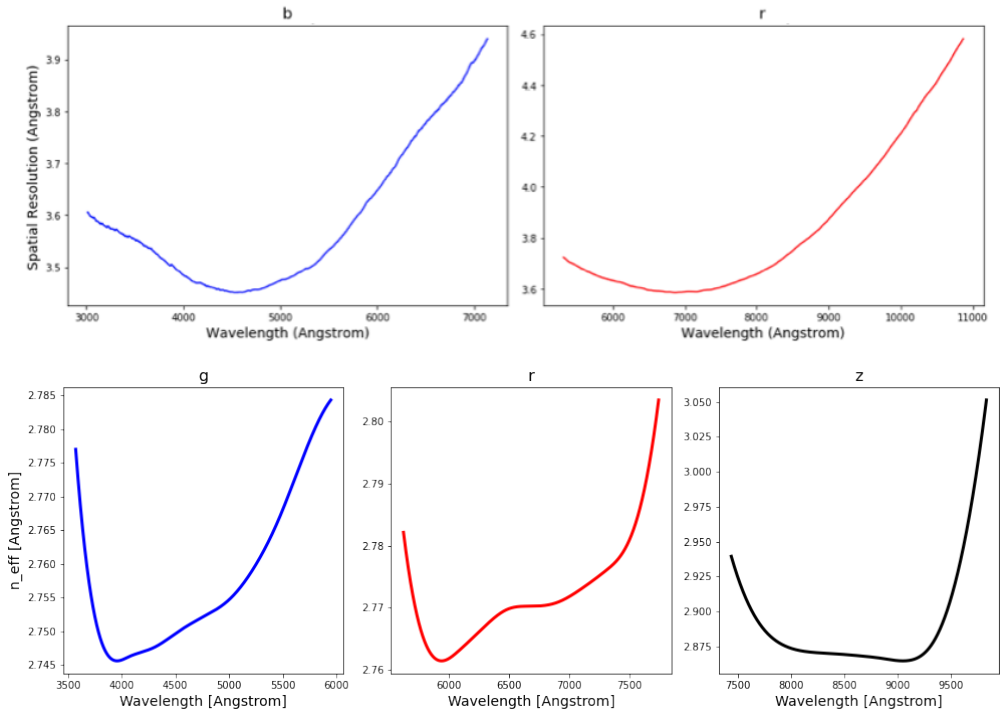


Figure 5.6: Spatial resolution for the eBOSS (top) and DESI (bottom) configurations.

This quantity was derived for all standard stars on plate 7027, the result of which is shown in Fig. 5.9. The vertical streaks in the red camera are most likely due to a failure in the sky subtraction, since sky emission lines show up predominantly at redder wavelengths. Next, the median of the 20 standard stars on this plate was taken, and smoothing was applied by downsampling by a factor of 100. The product of the fiberloss and throughput after smoothing is shown in Figures 5.10-5.11 for the blue and red cameras.

The final eBOSS quasar throughput data was obtained by dividing out a constant value for the fiberloss. Fig. 5.12 shows fiberloss values as a function of wavelength for 500 simulated DESI quasars, with the median over all fibers shown as the black solid line and the tabulated quasar fiberloss values shown as the dashed red line. The range of fiber colors correspond to their respective distance from the center of the focal plane. It was decided that a fiber acceptance value of 0.55 would be sufficient to use, as this seemed to be about the average across all wavelengths.

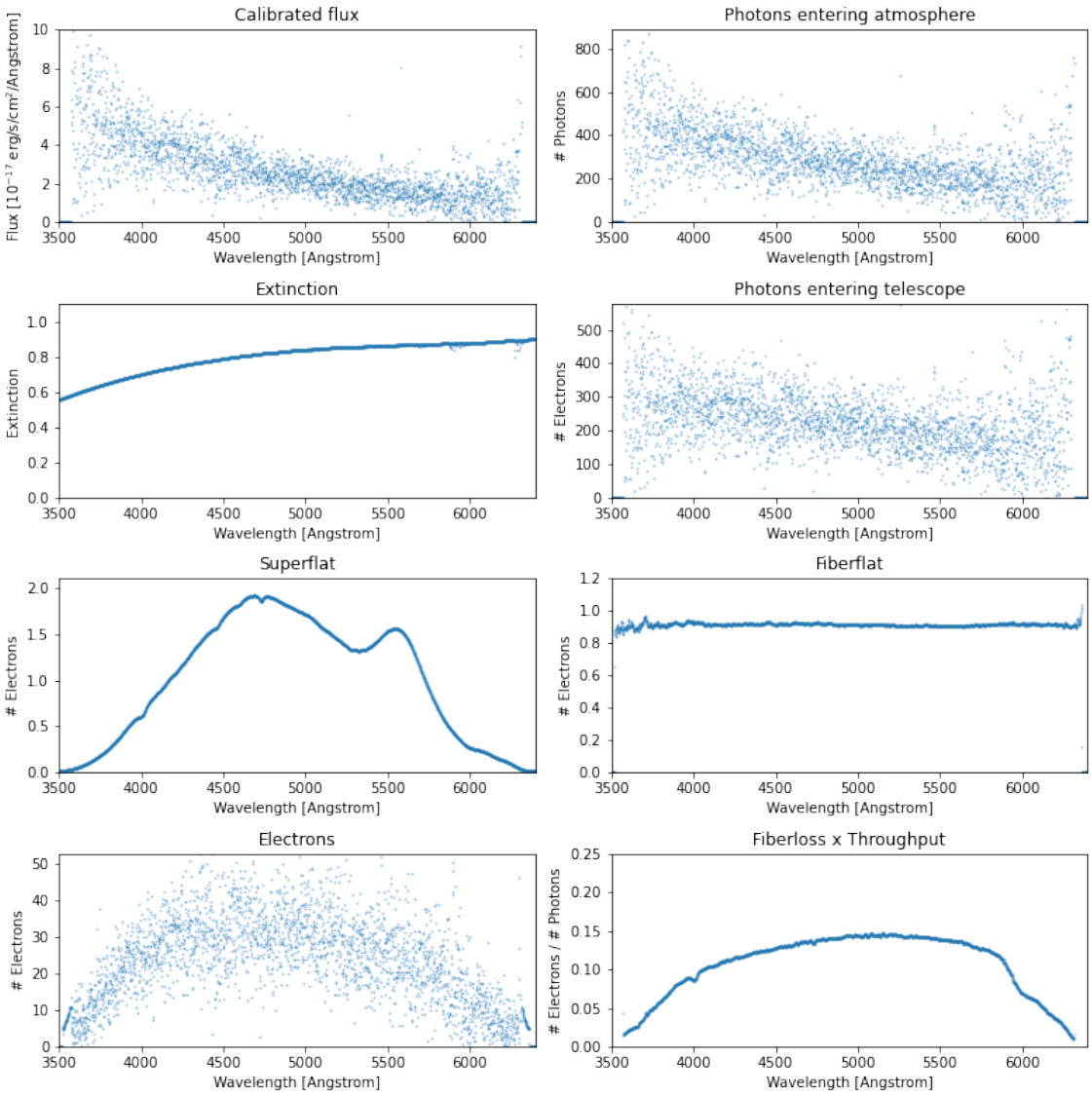


Figure 5.7: Flow chart showing the stages involved in converting from the calibrated flux of a standard star corresponding to PLATE 7027, MJD 56448, and FIBER 97 to detected electrons for the blue camera.

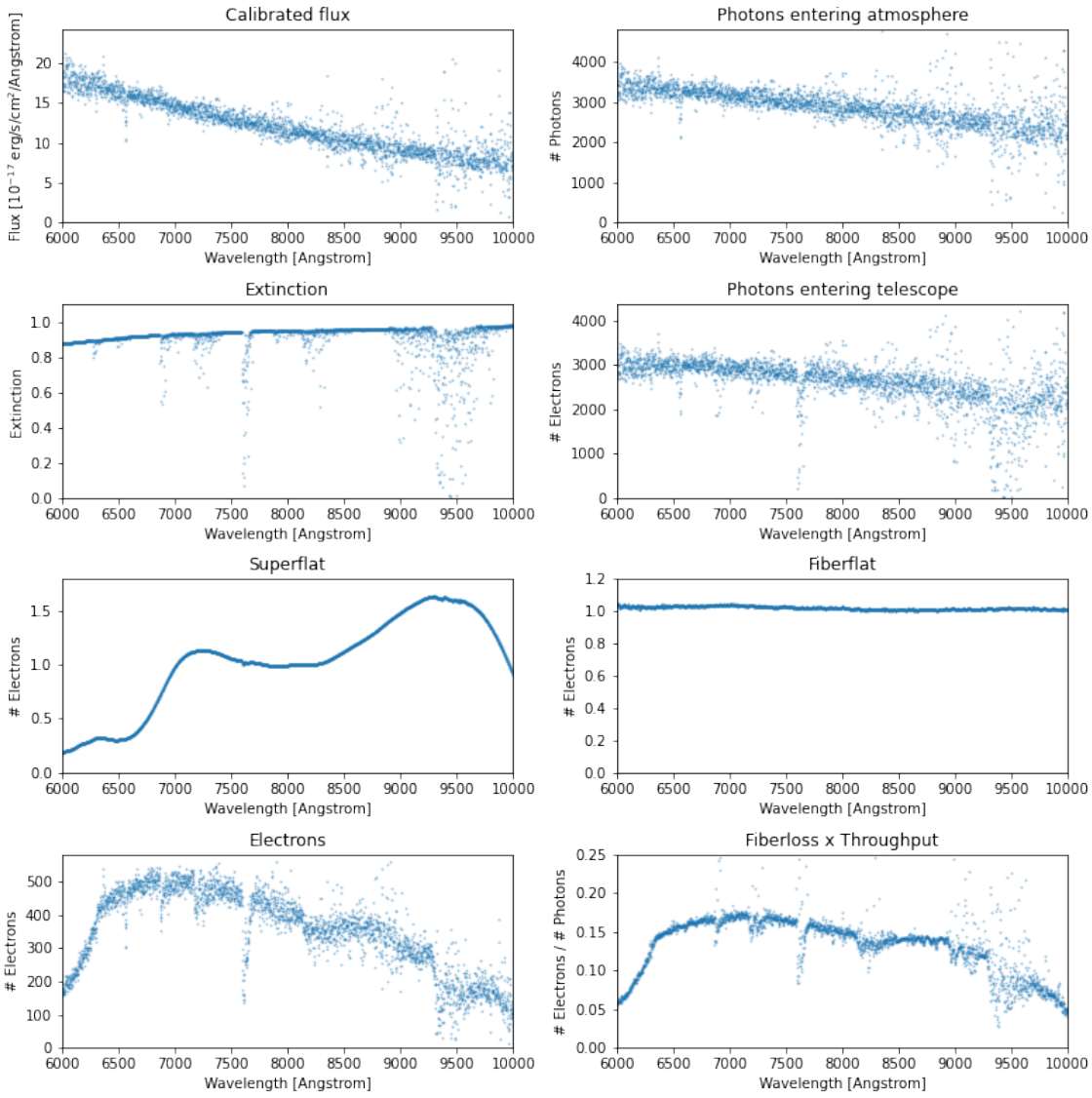


Figure 5.8: Flow chart showing the stages involved in converting from the calibrated flux of a standard star corresponding to PLATE 7027, MJD 56448, and fiber 97 to detected electrons for the red camera.

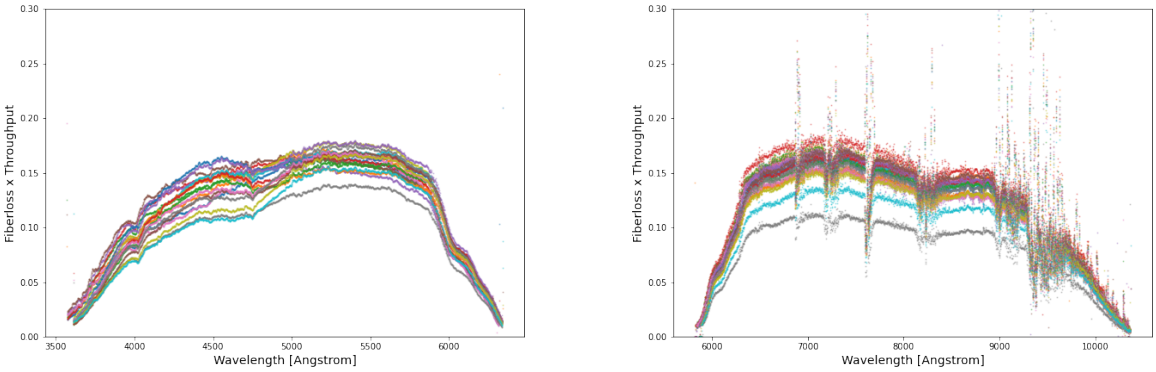


Figure 5.9: Results for the product of fiberloss and throughput for all spectrophotometric standards on PLATE 7027 and MJD 56448 for the blue (left) and red (right) cameras.

It was later found that data for the expected throughput of the eBOSS spectrograph, based on the engineering design, were available from the eBOSS collaboration (Smee et al. 2013), depicted as the black line in Fig. 5.13. Fresnel losses at the fiber faces, focal ratio degredation, and "slit" losses for 1" FWHM seeing conditions modeled with a double Gaussian PSF are not shown in this curve and had to be accounted for. These effects degraded the final throughput curve by an additional 15%, 4%, and 17%, respectively. A comparison of the throughput derived from standards stars on four different plates versus the expected throughput data provided by the eBOSS collaboration is shown in Fig. 5.14. The shapes of the dotted and solid curves are generally consistent, although there is a vertical offset in all four cases that is likely the result of different normalizations in the flux-calibration procedure. The Smee et al. (2013) throughput data was ultimately used for the `specsim` eBOSS configuration.

Fiber acceptance fraction

The fiber acceptance fraction, or fiberloss, is the fraction of light incident on the fiber that makes it through the fiber face. It is ultimately determined by the source profile, as a smaller fraction of the total light from extended sources is likely to pass through the fiber than the fraction of total light from sources with star-like profiles, such as quasars.

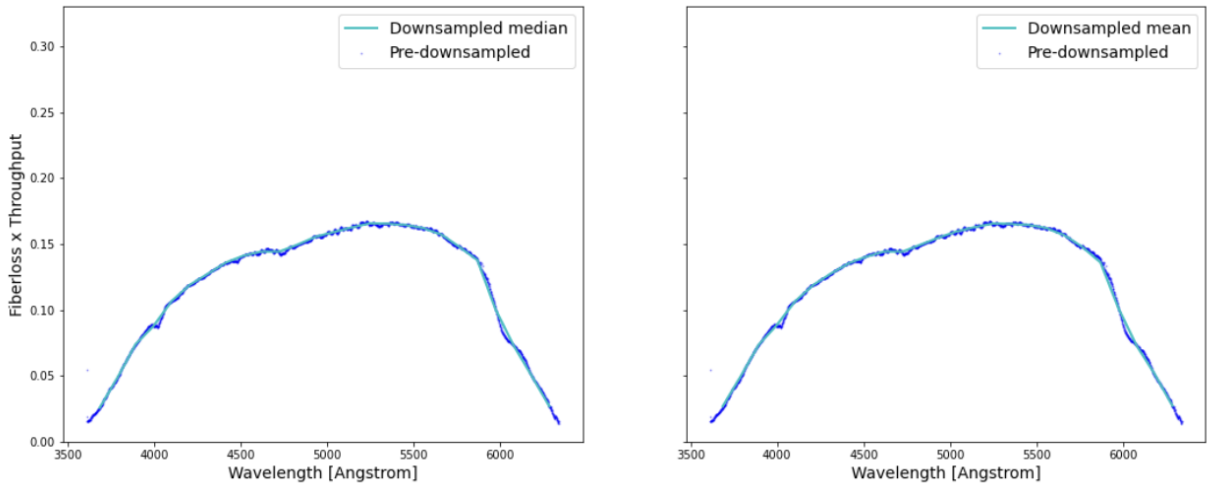


Figure 5.10: Results of downsampling the product of fiberloss and throughput in the blue camera by a factor of 100 using the median (left) and mean (right) for all standards on PLATE 7027 and MJD 56448 (left panel of Fig. 5.9).

There are three different methods to account for fiberloss in `specsim`. The `table` method uses a pre-computed table of fiberloss fractions based on the source type. The `galsim` method uses information about the transverse profile of the source on the sky to calculate fiberloss fractions via `GalSim`. Finally, `specsim` can interpolate fiber acceptance values pre-computed with `GalSim` with the `fastsim` method. This assumes a fixed axis ratio of 0.7, a fixed Moffat PSF model with $\beta = 3.5$ for the atmosphere, and a fixed fiber diameter of 107μm.

The `table` method for the eBOSS configuration required fiberloss values for each source type. This was done by adjusting the DESI fiberloss data to cover the range of the simulation grid by the beginning or end of each fiberloss array with the value at the smallest or largest wavelength, respectively. The final values for three different source types are shown in Fig. 5.15, which shows both DESI and eBOSS fiber acceptance fractions.

Additional configuration data

There were a few more inputs that needed to be specified for eBOSS, namely the extinction, the sky brightness, and a model quasar source spectrum. These data were generated in

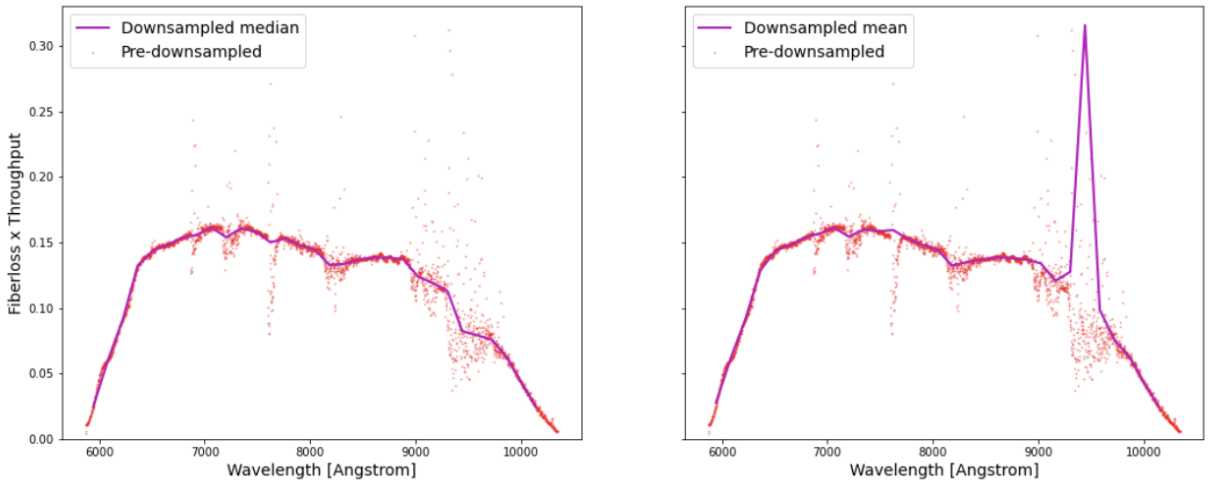


Figure 5.11: A smoothed version of the combination of the product of fiberloss and throughput in the red camera for all standards on PLATE 7027 and MJD 56448 (right panel of Fig. 5.9). The results of downsampling by a factor of 100 using the median and mean are shown in the left and right panels, respectively. The spike in the red camera, noticeable particularly in the mean of the downsampled curve, is introduced after dividing out the extinction.

a similar fashion to the fiberloss in that they were derived from the existing DESI values, however, there was an additional option to extend the array to cover the range over which to interpolate the data over the simulation grid. Rather than simply padding the ends of the array with a constant value, both ends of the extinction and the source spectrum were fit to a low-order polynomial to give more continuous coverage and better match the trends at low and high wavelengths, as seen in Fig. 5.16. The sky brightness was flat at each end of the spectrum, and so a constant value was used for padding. The DESI configuration allows the user to choose from one of three observing conditions (dark, bright, and grey, based on a combination of moon altitude and phase). The sky spectrum for eBOSS was based on the DESI dark sky spectrum, since quasars are typically observed during dark conditions where there is little scattered moonlight.

The left panel of Fig. 5.17 shows the final eBOSS instrument configuration data, where the blue and green curves correspond to the blue and red camera, respectively. A similar result is shown for the three channels in the default DESI configuration on the right panel.

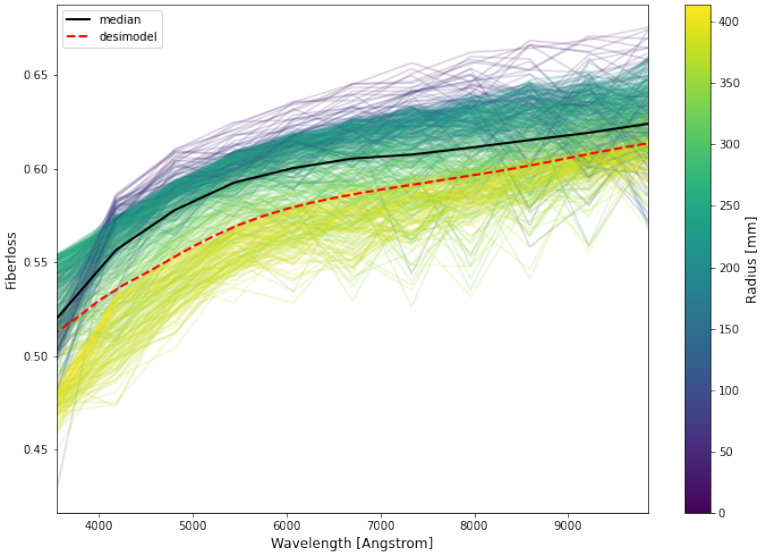


Figure 5.12: Simulated quasar fiberlosses using the `desimodel` parameters with the `galsim` package. Each curve shows an individual fiber, color-coded according to its focal plane radius.

5.3.4 Verification of the eBOSS configuration

Simulated spectra produced in both DESI and eBOSS configurations were compared as an initial test of whether this new mode of `specsim` was working properly. The same model source spectrum of a 22.4 magnitude quasar at z= 1.5 was used in both cases shown in Fig. 5.19. Unlike the default linear output wavelength grid for DESI, these spectra are plotted in the native binning scheme for eBOSS, which is logarithmic.

The eBOSS Lyα working group requested that `specsim` be able to accommodate this binning scheme for the eBOSS configuration, and comparisons between the two modes are given in terms of the same logarithmic binning for consistency. The package was modified to do a flux-conserving downsampling of the simulation grid. In effect, the flux density was multiplied by the bin width to convert to counts, and re-binned to the desired eBOSS resolution, which is linear in $\log(\lambda)$. Finally, the re-binned counts were multiplied by the new bin widths to recover the flux density. The simulated SEDs in Fig. 5.20 are the perfectly calibrated observed flux densities of the source spectrum for DESI and for eBOSS.

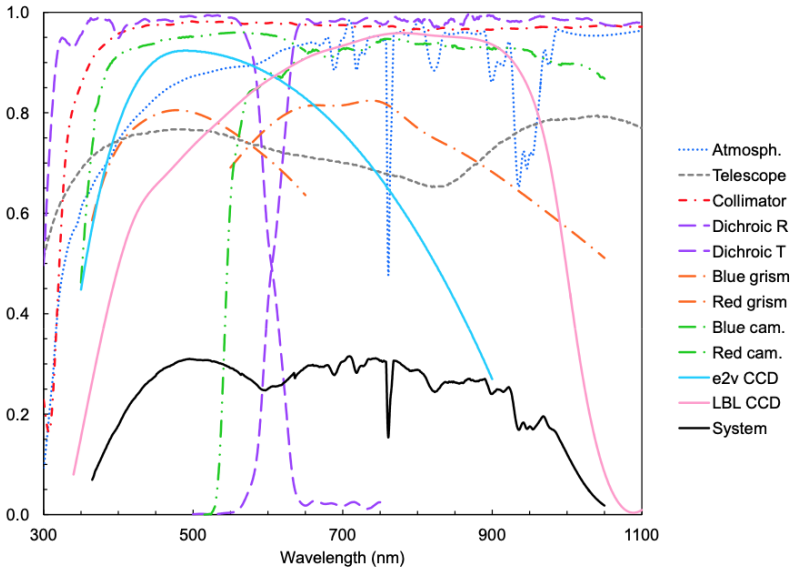


Figure 5.13: Expected overall system throughput for the SDSS spectrographs (black), along with all wavelength-dependent contributors to the throughput (labeled by color on the right-hand side). Figure is from Smee et al. (2013).

A proper study of the instrumental noise was also done by studying the inverse variance due to dark noise, read noise, and random noise in the camera. The simulated camera inverse variance for DESI and eBOSS setups is shown in Fig. 5.18. To evaluate the success of the eBOSS configuration in modeling the instrumental noise, a (noiseless) model eBOSS stellar spectrum used for calibration was compared with the observed flux. An example of one particular stellar model and sky background, both obtained from eBOSS `spec` files, that were used as inputs to the configuration file is given in Fig. 5.21. Fig. 5.22 shows the simulated flux density in blue, along with the observed flux density in orange for each camera. The mean flux for both simulated and observed spectra are fairly consistent, however there is slightly more scatter for the simulated output.

Comparing the inverse variance of the simulated and observed stellar spectrum in Fig. 5.23, we see again that the shape of the curves generally agree, however `specsim` tends to over-estimate the noise by around 25% in each camera. This study was done using other stellar spectra, shown in Fig. 5.24, and the trend persisted. The difference between the simulated

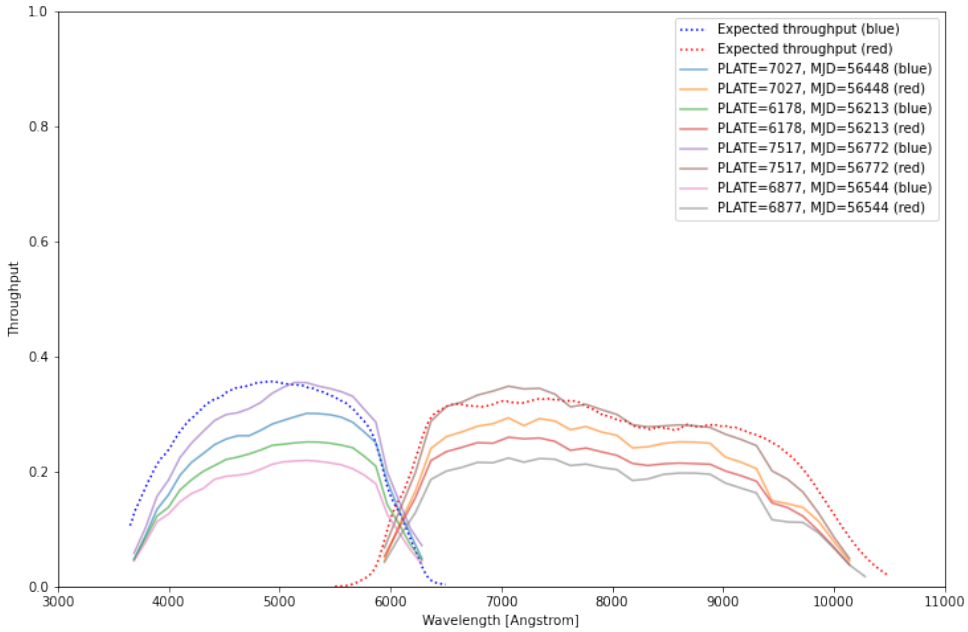


Figure 5.14: A comparison of throughput solutions derived by working backwards from standard star spectra via the procedure described above, to the expected throughput from Smee et al. (2013).

inverse variance and that obtained from the data was fairly consistent across both cameras for a given stellar spectrum, however the size of that difference varied from spectrum to spectrum. The likely cause of the difference between the `specsim` output and the data is that the eBOSS pipeline fit to a stellar spectrum model is artificially reducing the inverse variance in the data, and so the `specsim` are probably more accurate.

5.4 Verifying the `specsim` sky model

A second way of validating `specsim` was to go beyond comparing the real flux densities with synthetic eBOSS spectra to investigate the success of the lunar and sky model in capturing variations in sky brightness levels. This was done by comparing measured sky brightnesses

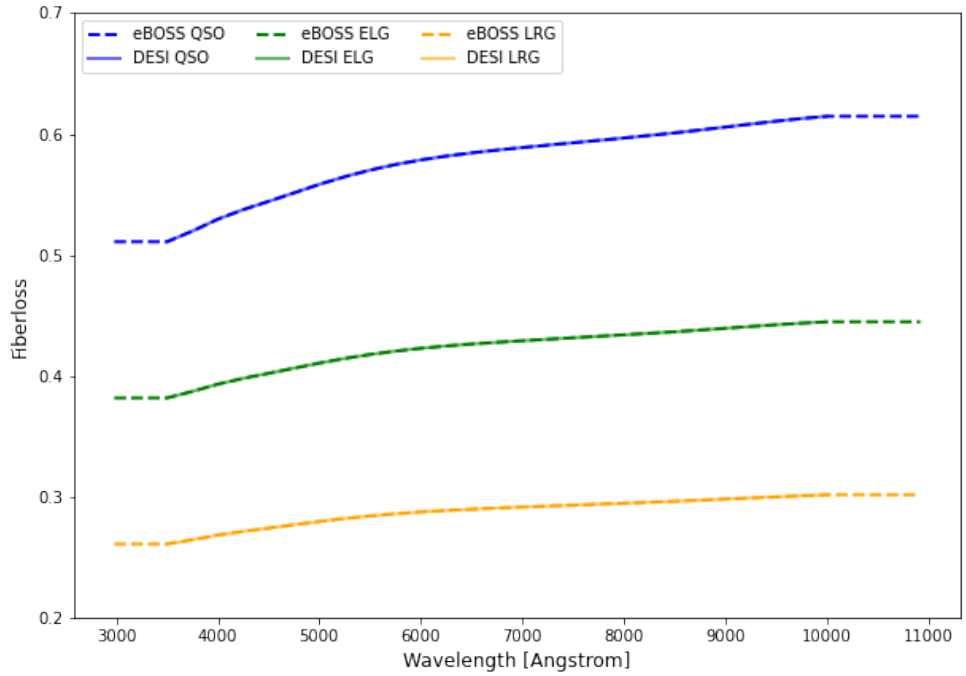


Figure 5.15: DESI fiberloss data (dashed curve), adjusted to cover the wavelength range of the eBOSS spectrograph (solid curve).

from the DECam Legacy Survey (DECaLS) with predicted sky levels using `specsim`. This study focused on cases where conditions were non-photometric (i.e. there was significant light contamination from the moon).

5.4.1 DECaLS sky data

The data used in this section came from Data Release 3 (DR3) of the DECam Legacy Survey, a photometric percursor survey to DESI (see Section 4.1.1). A plot of the unique observations in this dataset is shown in Fig. 5.26, where pointings are binned and colored according their separation angle, or opening angle between the observation and the moon in degrees.

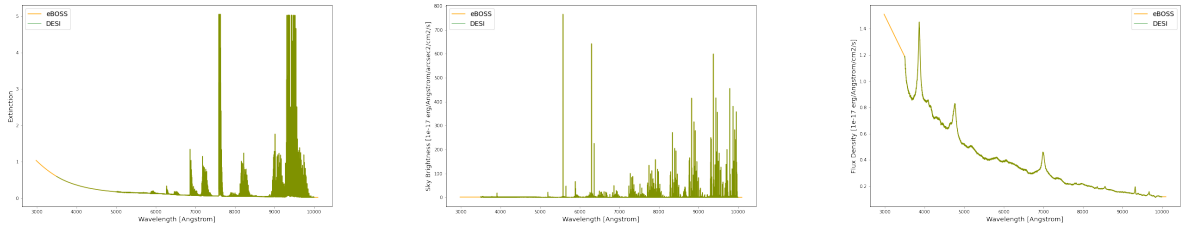


Figure 5.16: Additional wavelength-adjusted data for eBOSS atmospheric extinction (left), sky brightness (middle) and a model quasar source spectrum (right).

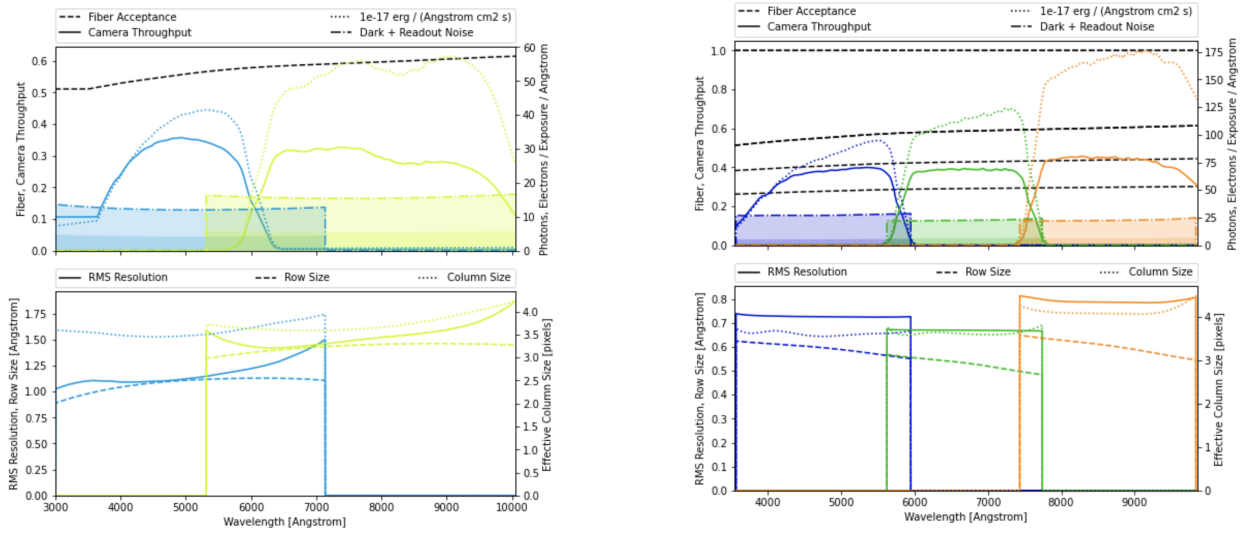


Figure 5.17: The final `specsim` instrument configurations for eBOSS (left) and DESI (right).

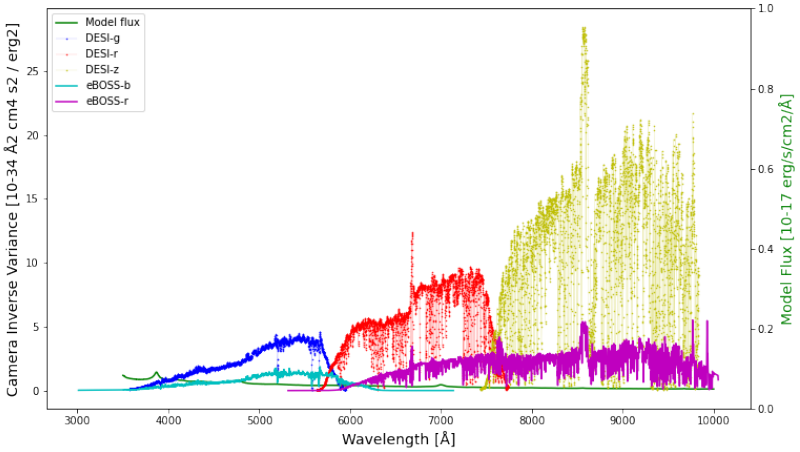


Figure 5.18: A comparison of the DESI and eBOSS inverse variances, color-coded by camera.

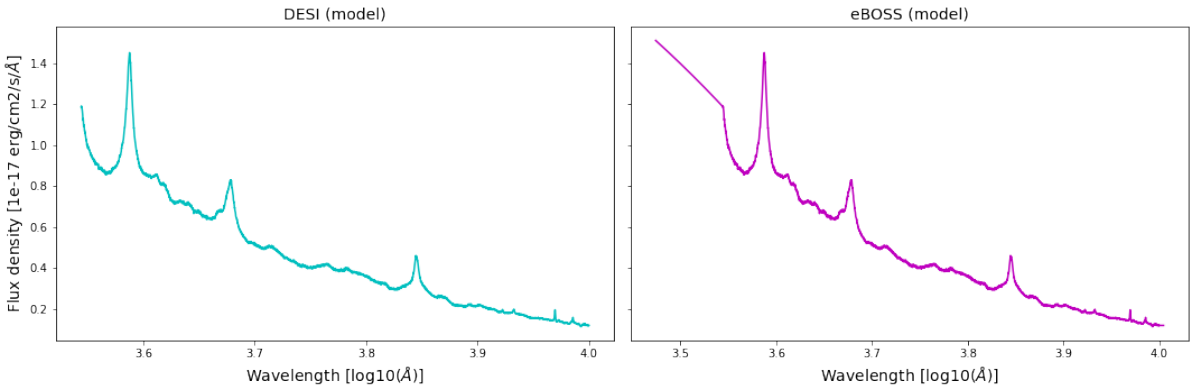


Figure 5.19: Final DESI (left) and eBOSS (right) model QSO spectra.

The effect of scattered moonlight for a single pointing can be better visualized on a polar plot. Fig. 5.27, generated using the `specsim` package, shows the amount of scattered moonlight in the V-band for two different positions of the moon. The result is based on a model developed by Krisciunas & Schaefer (1991) that predicts scattered moonlight as a function of the moon phase, the zenith angles of the moon and the telescope pointing, the separation angle between the moon and each pointing, and the local, V-band extinction coefficient. The position of each pointing was converted from equatorial coordinates into altitude-azimuth coordinates and plotted with respect to the moon zenith and azimuth angles. Moon zenith angles are shown as concentric circles of equal radius, and the azimuth angle increases from

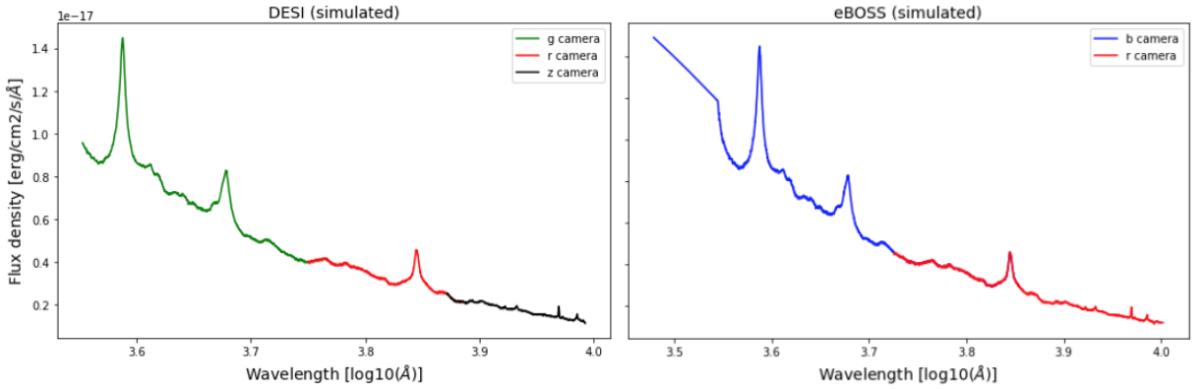


Figure 5.20: A comparison of a simulated quasar spectrum using the DESI (left) and eBOSS (right) `specsim` configurations.

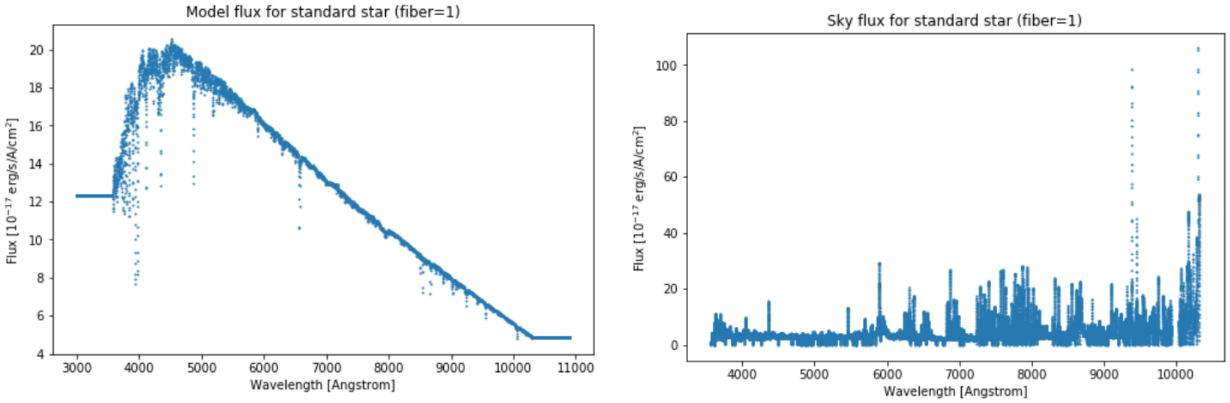


Figure 5.21: A stellar source spectrum and sky background spectrum used to study the simulated eBOSS inverse variances.

0° to 360° in a clockwise fashion. The phase of the moon, given by ϕ, is close to full in both instances, where 0 and 1 correspond to full and new moons, respectively. The V-band extinction coefficient is shown as k_V. What is clear from the plots, but perhaps not intuitively, is that there is a non-negligible amount of moon contamination along the edge of the horizon for all azimuth angles, even when the moon is far from the zenith and below the horizon, as seen in (b). This is the result of a geometrical effect due to the curvature of the Earth.

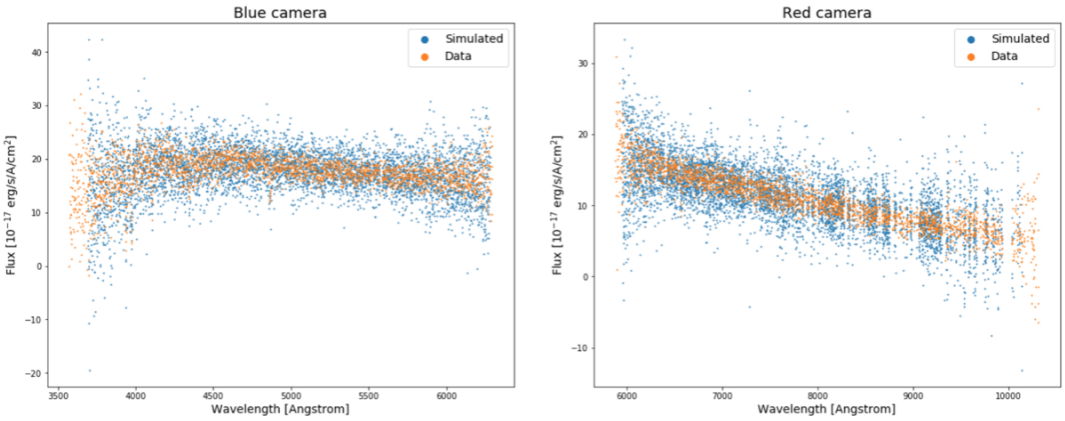


Figure 5.22: Simulated and true fluxes for a stellar source spectrum for the eBOSS blue (left) and red (right) spectrographs.

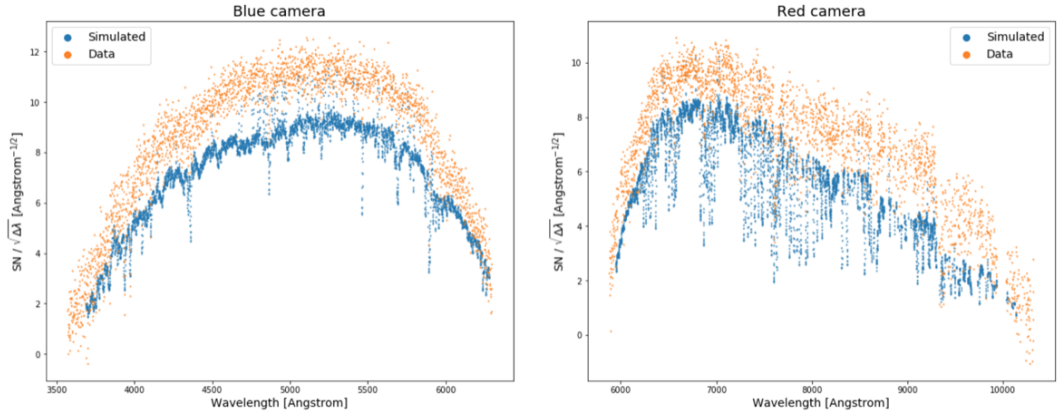


Figure 5.23: Simulated and true inverse variances for a stellar source spectrum for the eBOSS blue (left) and red (right) spectrographs.

5.4.2 Simulating the sky with `specsim`

Because I specifically focused on investigating sky levels under bright conditions, I chose to select observations taken when the moon was above the horizon. To this effect, a subset of the data was selected cuts on moon altitude, transparency and sun altitude (see Table 5.6). Sky brightnesses were generated in `specsim` for each unique DECaLS observation using metadata such as the airmass, moon phase, moon altitude, moon position, and pointing. They were then passed through a set of DECam filters using the `speclite`[7] package (Kirkby

[7] `https://speclite.readthedocs.io/`

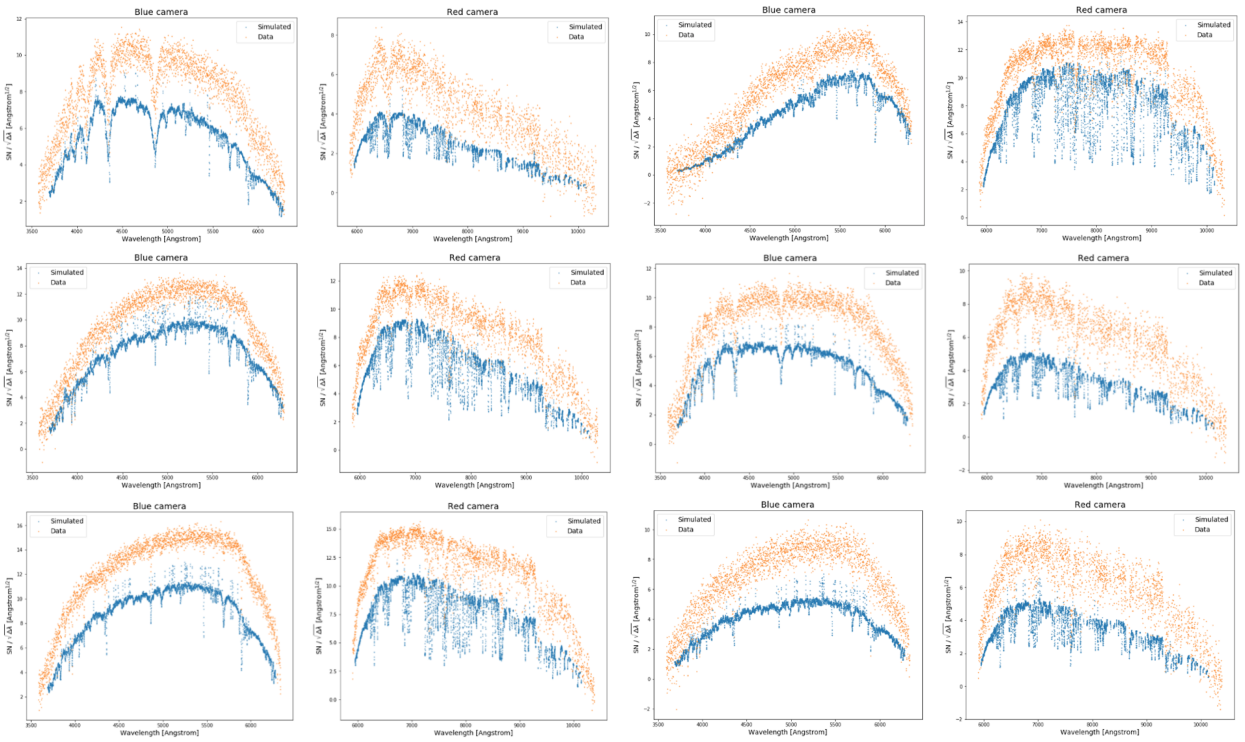


Figure 5.24: Simulated and true inverse variances of a range of stellar source inputs for the eBOSS `specsim` configuration. From left-to-right and top-to-bottom, they correspond to FIBERS 334, 395, 41, 694, 890, and 970 on PLATE 4055 observed on MJD 55359.

et al. 2021) to give the sky brightness in nanomaggies / sq arcsec. A maggie is a linear unit of flux density defined as:

$$1\,\text{maggie} = 10^{-0.4\,m_{AB}}, \tag{5.12}$$

where m_{AB} is the AB apparent magnitude. Finally, the output was converted into magnitudes using the standard conversion from flux in nanomaggies f to magnitude m:

$$m = 22.5 - 2.5\log_{10}(f \times 10^9). \tag{5.13}$$

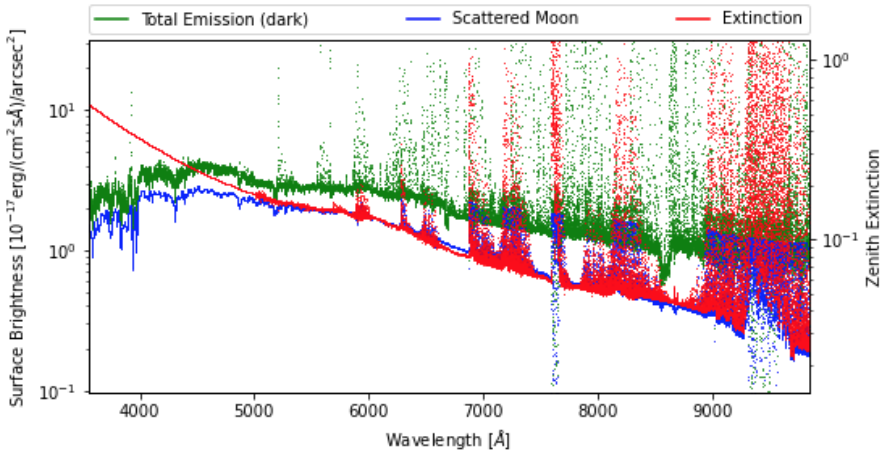


Figure 5.25: Contribution of different atmospheric components in `specsim`.

The simulated sky brightness for each photometric band is shown against empirical data from DECaLs in Fig. 5.28. There was considerable horizontal scatter across all filters, and so an additional cut was made on transparency in order to attempt to reduce it. Transparency, defined as the flux from a source relative to the expected flux under nominal conditions, is affected by elements in the atmosphere that absorb or scatter light, such as aerosols, water vapor, and clouds. Because the relationship between transparency and scattered moonlight is fairly complicated, a tight cut was made in order to isolate targets with high transparency. Fig. 5.29 shows the results after a transparency cut of $0.75 <$ transparency < 1.3.

This eliminated some of the streaks in the figures, however there were still observations where the measured brightness varied for the same value of predicted sky brightness. Upon closer inspection, some of these observations occurred on the same day, such as those in labeled in orange in Fig. 5.30. These pointings all occurred on 4 April, 2013, at times close to astronomical twilight. Because this scatter correlated with sun altitude, a third cut was made to only include observations where the sun altitude $< -15°$. This mostly eliminated the scatter in the g- and r-bands, with the remaining z-band scatter mostly due to the presence of OH lines and other gases along redder wavelengths, which is expected.

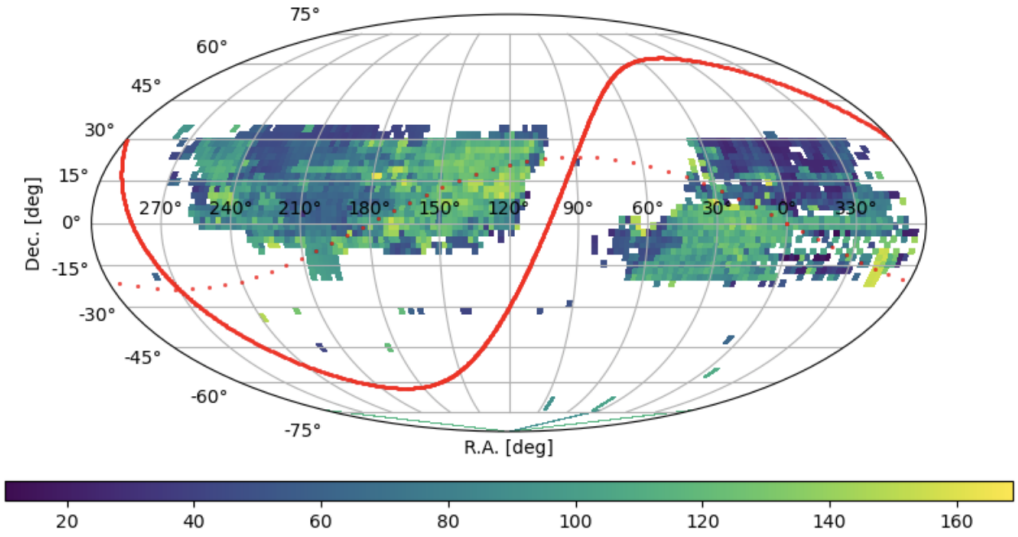


Figure 5.26: Moon separation angles for unique observations of sources in DR3.

Table 5.6: Cuts on DECaLS data for bright conditions.

moon altitude $> 15°$
$0.9 <$ transparency < 1.3
sun altitude $< -15°$

Addressing the g-band offset

The model is successful in predicting sky and moon contamination in the r- and z-bands, although there is a systematic offset in the g-band where the `specsim` model seems to underestimate the sky brightness. One possibility is that atmospheric extinction has been applied twice: the DECam filter responses already include a reference atmosphere with airmass 1.3, and `specsim` already applies atmospheric extinction to the spectra before they are convolved with the filters. The study was repeated using a set of DECam filters without atmospheric extinction. The results, shown in Fig. 5.32, show no significant change in predicted sky brightness. One of the primary contributors to the sky level during non-photometric nights is reflected light from the moon. The reflectivity of the moon, or moon albedo, has been found to decrease with wavelength, and is particularly low at near-UV wavelengths (Carver

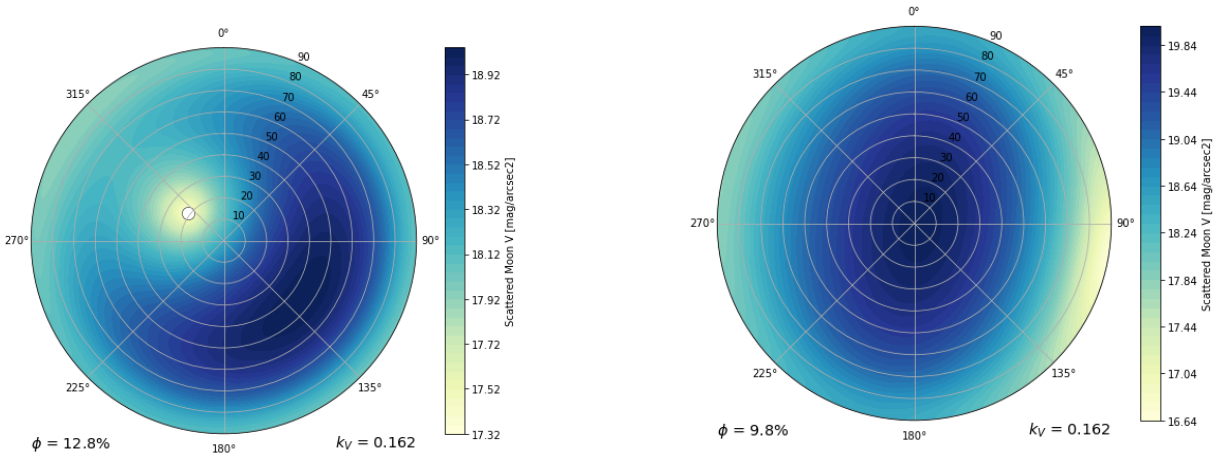


Figure 5.27: The amount of scattered moonlight for two different positions of the moon as a function of altitude and azimuth.

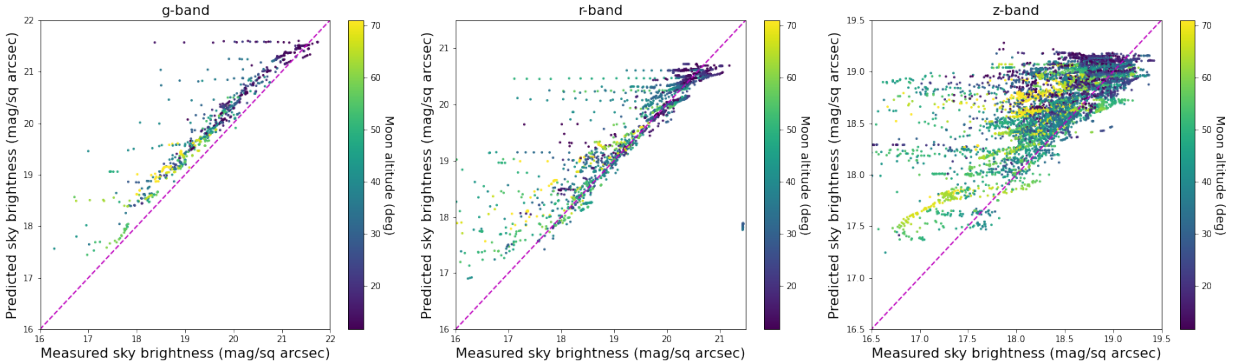


Figure 5.28: Predicted versus empirical sky brightness in three different bands after only imposing a cut on moon altitude.

et al. 1974). It may be possible that `specsim` is underestimating the moon albedo, leading to fainter sky levels.

Sky brightness residuals

Finally, histograms of the empirical versus predicted residuals were made and are shown in Fig. 5.33, along with the mean of the residuals for each passband. Overall, `specsim` accurately captures the measured variations in sky brightness levels in the r- and z-bands, where the offset in the z-band is expected due to the presence of sky emission lines in the

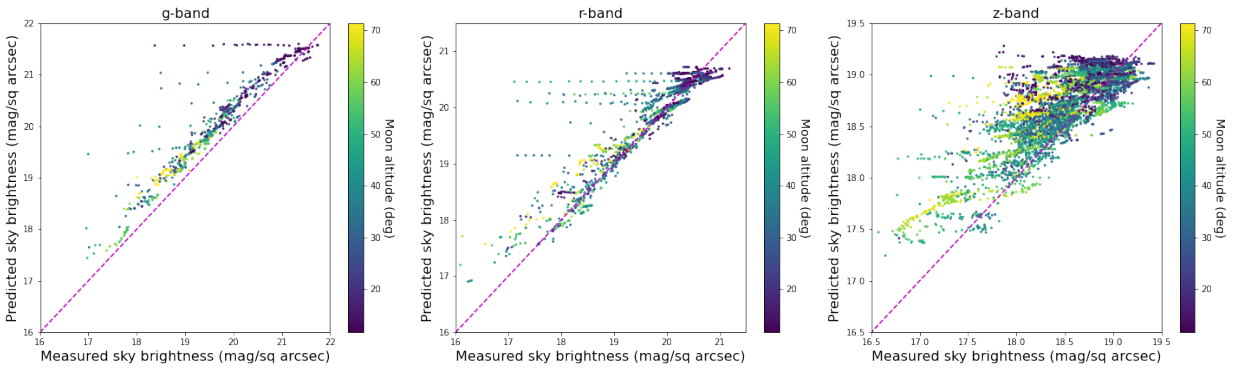


Figure 5.29: Predicted versus empirical sky brightness in three different bands after applying both moon altitude and transparency cuts.

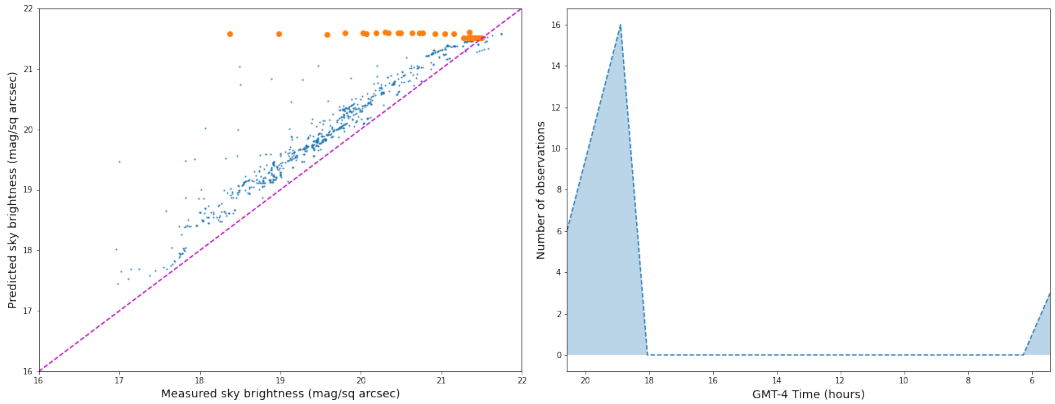


Figure 5.30: Predicted versus measured sky brightness in the g-band (left), showing the anomalous streak of observations taken during astronomical twilight in orange. The timestamps of these observations (right) show the peak at twilight.

near-infrared. The g-band offset is not as well understood, however further studies into moon albedo models could account for fainter predicted sky levels at shorter wavelengths.

5.5 Conclusion

The `specsim` software package was originally developed to simulate the response of the DESI telescope and instrument. I reconfigured this package to run in a separate mode where it would simulate eBOSS spectra for the eBOSS Lyα working group. This involved changing telescope and instrument parameters in the configuration file to their nominal eBOSS values,

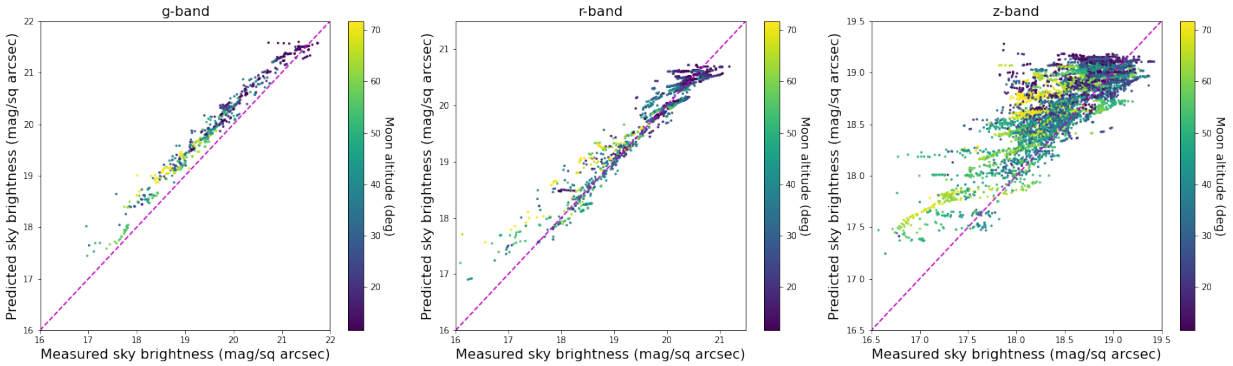


Figure 5.31: Predicted versus empirical sky brightness in three different bands after applying moon altitude, transparency, and sun altitude cuts.

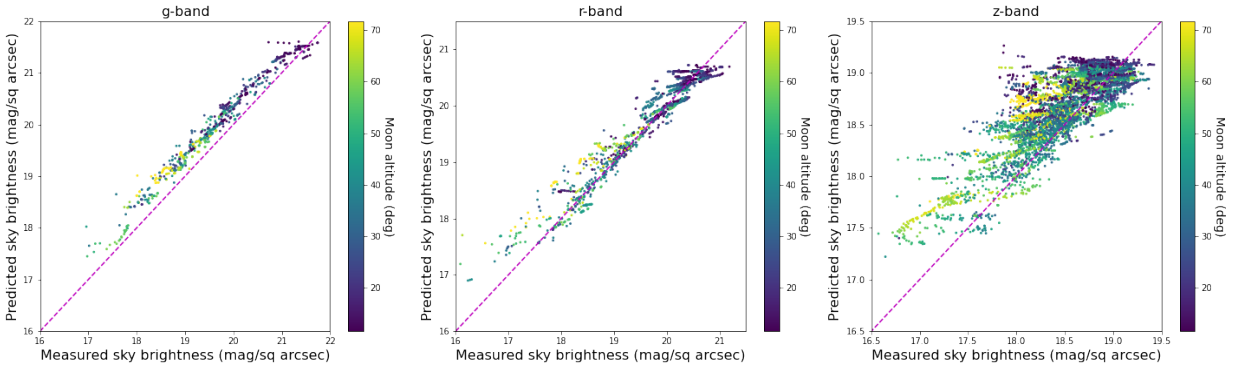


Figure 5.32: Predicted versus empirical sky brightness in three different bands after applying moon altitude, transparency, and sun altitude cuts, but without including atmospheric extinction.

as well as generating data for the eBOSS camera such as the wavelength binning, spectral resolution, throughput, and the fiber acceptance fraction.

A verification of `specsim` was performed by comparing the inverse variances of a model stellar spectrum with a real eBOSS stellar source. Another validation used the DESI configuration of `specsim` to compare sky brightness levels simulated via `specsim` with those from photometry from DECaLS, showing general agreement, other than an offset in the g-band.

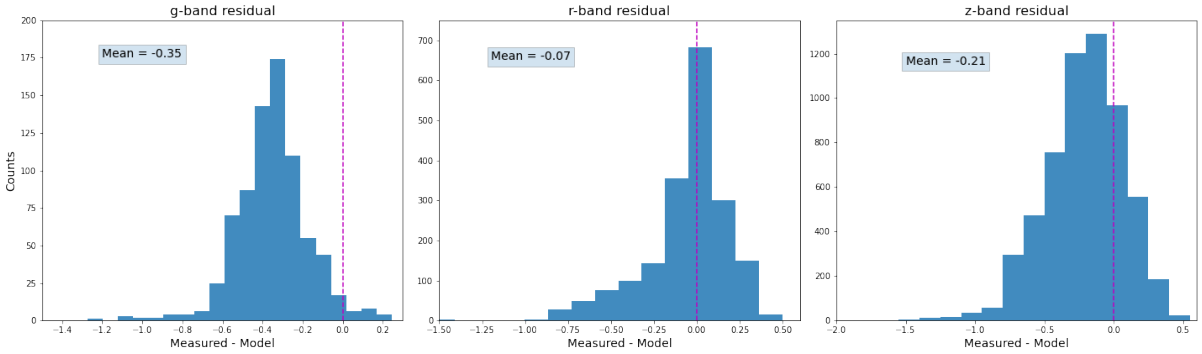


Figure 5.33: Sky brightness residuals (predicted - model) in all three photometric bands.

Bibliography

Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016, Phys. Rev. Lett., 116, 061102

Abolfathi, B. 2022, SOMVIZ, `https://github.com/belaa/SOMVIZ`

Abolfathi, B., Alonso, D., Armstrong, R., et al. 2021, The Astrophysical Journal Supplement Series, 253, 31

Aghanim, N., Akrami, Y., Ashdown, M., et al. 2020, Astronomy & Astrophysics, 641, A6

Alarcon, A., Sánchez, C., Bernstein, G. M., & Gaztañaga, E. 2020, Monthly Notices of the Royal Astronomical Society, 498, 2614–2631

Amon, A., Gruen, D., Troxel, M., et al. 2022, Physical Review D, 105

Arnouts, S. & Ilbert, O. 2011, Astrophysics Source Code Library, 08009

Asgari, M., Lin, C.-A., Joachimi, B., et al. 2021, Astronomy & Astrophysics, 645, A104

Bartelmann, M. & Schneider, P. 2001, Physics Reports, 340, 291–472

Benitez, N. 2000, The Astrophysical Journal, 536, 571–583

Bolzonella, M., Miralles, J. M., & Pello, R. 2000, , 363, 476

Bond, T., Roodman, A., Reil, K., et al. 2018, in Modeling, Systems Engineering, and Project Management for Astronomy VIII, ed. G. Z. Angeli & P. Dierickx, Vol. 10705, International Society for Optics and Photonics (SPIE), 150 – 169

Bonnett, C. 2015, Monthly Notices of the Royal Astronomical Society, 449, 1043

Bovy, J., Hogg, D. W., & Roweis, S. T. 2011, The Annals of Applied Statistics, 5

Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, The Astrophysical Journal, 686, 1503–1513

Buchs, R., Davis, C., Gruen, D., et al. 2019, Monthly Notices of the Royal Astronomical Society, 489, 820–841

Carrasco Kind, M. & Brunner, R. J. 2013, Monthly Notices of the Royal Astronomical Society, 432, 1483–1501

Carver, J. H., Horton, B. H., O'Brien, R. S., & O'Connor, G. G. 1974, The Moon, 9, 295

Chabanier, S., Etourneau, T., Le Goff, J.-M., et al. 2022, The Astrophysical Journal Supplement Series, 258, 18

Chabanier, S., Millea, M., & Palanque-Delabrouille, N. 2019, Monthly Notices of the Royal Astronomical Society, 489, 2247–2253

Cole, S., Percival, W. J., Peacock, J. A., et al. 2005, Monthly Notices of the Royal Astronomical Society, 362, 505–534

Collister, A. & Lahav, O. 2004, Publications of the Astronomical Society of the Pacific, 116, 345–351

Connolly, A. J., Angeli, G. Z., Chandrasekharan, S., et al. 2014, in Modeling, Systems Engineering, and Project Management for Astronomy VI, ed. G. Z. Angeli & P. Dierickx, Vol. 9150, International Society for Optics and Photonics (SPIE), 414 – 421

Connolly, A. J., Peterson, J., Jernigan, J. G., et al. 2010, in Modeling, Systems Engineering, and Project Management for Astronomy IV, ed. G. Z. Angeli & P. Dierickx, Vol. 7738, International Society for Optics and Photonics (SPIE), 612 – 621

Cox, D. R. 1955, Journal of the Royal Statistical Society. Series B (Methodological), 17, 129

Dark Energy Science Collaboration. 2020, LSSTDESC/imsim: May 2020 Release (Version v1.0.0), `https://github.com/LSSTDESC/imSim.git`

David Kirkby. 2020, dkirkby/zotbin, `https://github.com/dkirkby/zotbin.git`

DESI Collaboration. 2016, The DESI Experiment Part I: Science, Targeting, and Survey Design

Dey, A., Schlegel, D. J., Lang, D., et al. 2019, The Astronomical Journal, 157, 168

Dodelson, S. 2003, Modern Cosmology (San Diego: Elsevier)

du Mas des Bourboux, H., Rich, J., Font-Ribera, A., et al. 2020, The Astrophysical Journal, 901, 153

Dyson, F. W., Eddington, A. S., & Davidson, C. 1920, Philosophical Transactions of the Royal Society of London

Eisenstein, D. J., Seo, H., & White, M. 2007, The Astrophysical Journal, 664, 660–674

Eisenstein, D. J., Zehavi, I., Hogg, D. W., et al. 2005, The Astrophysical Journal, 633, 560–574

Farr, J., Font-Ribera, A., Bourboux, H. d. M. d., et al. 2020, Journal of Cosmology and Astroparticle Physics, 2020, 068–068

Flaugher, B., Diehl, H. T., Honscheid, K., et al. 2015, The Astronomical Journal, 150, 150

Gelman, A. & Rubin, D. B. 1992, Statistical Science, 7, 457

Gerdes, D. W., Sypniewski, A. J., McKay, T. A., et al. 2010, The Astrophysical Journal, 715, 823–832

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., et al. 2014, Generative Adversarial Networks

Grandell, J. 1976, Doubly Stochastic Poisson Processes (Heidelberg: Springer-Verlag Berlin)

Hamana, T., Shirasaki, M., Miyazaki, S., et al. 2020, Publications of the Astronomical Society of Japan, 72

Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, The Astrophysical Journal Supplement Series, 208, 19

Holoien, T. W., Marshall, P. J., & Wechsler, R. H. 2017, The Astronomical Journal, 153, 249

Hubble, E. 1929, Proceedings of the National Academy of Sciences, 15, 168

Huterer, D., Kirkby, D., Bean, R., et al. 2015, Astroparticle Physics, 63, 23–41

Huterer, D., Takada, M., Bernstein, G., & Jain, B. 2006, Monthly Notices of the Royal Astronomical Society, 366, 101

Ivezić, v., Kahn, S. M., Tyson, J. A., et al. 2019, The Astrophysical Journal, 873, 111

Ivezić, Ž., Connolly, A., Vanderplas, J., & Gray, A. 2014, Statistics, Data Mining and Machine Learning in Astronomy (Princeton University Press)

Jelinsky, P., Wechsler, R., Sharples, R., et al. 2018, Ground-based and Airborne Instrumentation for Astronomy VII

Jenkins, A., Frenk, C. S., Pearce, F. R., et al. 1998, The Astrophysical Journal, 499, 20–40

Kilbinger, M. 2015, Reports on Progress in Physics, 78, 086901

Kingma, D. P. & Welling, M. 2014, Auto-Encoding Variational Bayes

Kirkby, D. 2015, dkirkby/bossdata: September 2015 release (Version v0.2.7), `https://github.com/dkirkby/bossdata`

Kirkby, D. et al. 2020, desihub/specsim: August 2020 Release (Version v0.14), `https://github.com/desihub/specsim`

Kirkby, D. et al. 2021, desihub/speclite: January 2021 Release (Version v0.14), `https://github.com/desihub/speclite`

Kobyzev, I., Prince, S., & Brubaker, M. 2020, IEEE Transactions on Pattern Analysis and Machine Intelligence, PP, 1

Kohonen, T. 1982, Biological Cybernetics, 43, 59

Kolb, E. W. & Turner, M. S. 1990, The Early Universe (New York: Addison-Wesley Publishing Company)

Korytov, D., Hearin, A., Kovacs, E., et al. 2019, The Astrophysical Journal Supplement Series, 245, 26

Kramer, M., Stairs, I. H., Manchester, R. N., et al. 2021, Phys. Rev. X, 11, 041050

Krause, E. & Eifler, T. 2017, Monthly Notices of the Royal Astronomical Society, 470, 2100–2112

Krisciunas, K. & Schaefer, B. E. 1991, Publications of the Astronomical Society of the Pacific, 103, 1033

Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, Euclid Definition Study Report

Leistedt, B., Mortlock, D. J., & Peiris, H. V. 2016, Monthly Notices of the Royal Astronomical Society, 460, 4258–4267

Levi, M. E., Allen, L. E., Raichoor, A., et al. 2019, The Dark Energy Spectroscopic Instrument (DESI)

LSST Dark Energy Science Collaboration. 2012, Large Synoptic Survey Telescope: Dark Energy Science Collaboration

LSST Dark Energy Science Collaboration. 2021, DESC DC2 Data Release Note

LSST Project. 2019, lsst-camera-dh/eotest: February 2019 Release (Version v0.1.1)

LSST Science Collaboration and LSST Project. 2009, LSST Science Book (Version 2.0)

Mandelbaum, R. 2018, Annual Review of Astronomy and Astrophysics, 56, 393–433

Masters, D., Capak, P., Stern, D., et al. 2015, The Astrophysical Journal, 813, 53

Matthews, D. J. & Newman, J. A. 2010, The Astrophysical Journal, 721, 456–468

Morrison, C. B., Hildebrandt, H., Schmidt, S. J., et al. 2017, Monthly Notices of the Royal Astronomical Society, 467, 3576

Ménard, B., Scranton, R., Schmidt, S., et al. 2014, Clustering-based redshift estimation: method and application to data

Newman, J. A. 2008, The Astrophysical Journal, 684, 88–101

Newman, J. A., Abate, A., Abdalla, F. B., & Allam, S. 2015, Astroparticle Physics, 63, 81

Papamakarios, G., Nalisnick, E., Rezende, D. J., Mohamed, S., & Lakshminarayanan, B. 2021, Normalizing Flows for Probabilistic Modeling and Inference

Peacock, J. A. 1999, Cosmological Physics (New York: Cambridge University Press)

Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, The Astrophysical Journal, 517, 565–586

Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, The Astronomical Journal, 116, 1009–1038

Roodman, A., Bogart, J. R., Bond, T., Borgland, A. W., & Chiang, J. 2018, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 10705, Modeling, Systems Engineering, and Project Management for Astronomy VIII, ed. G. Z. Angeli & P. Dierickx, 107050D

Rowe, B., Jarvis, M., Mandelbaum, R., et al. 2015, GalSim: The modular galaxy image simulation toolkit

Ryden, B. 2003, Introduction to Cosmology (San Francisco: Pearson Education, Inc.)

Sanchez, J., Mendoza, I., Kirkby, D. P., & Burchat, P. R. 2021, Journal of Cosmology and Astroparticle Physics, 2021, 043

Schmidt, S. J., Malz, A. I., Soo, J. Y. H., et al. 2020, Monthly Notices of the Royal Astronomical Society

Smee, S. A., Gunn, J. E., Uomoto, A., et al. 2013, The Astronomical Journal, 146, 32

Spergel, D., Gehrels, N., Baltay, C., et al. 2015, Wide-Field InfrarRed Survey Telescope-Astrophysics Focused Telescope Assets WFIRST-AFTA 2015 Report

Sánchez, C. & Bernstein, G. M. 2018, Monthly Notices of the Royal Astronomical Society, 483, 2801–2813

The LSST Dark Energy Science Collaboration. 2021, The LSST Dark Energy Science Collaboration (DESC) Science Requirements Document

Vanderplas, J., Connolly, A., Ivezić, Ž., & Gray, A. 2012, in Conference on Intelligent Data Understanding (CIDU), 47 –54

Weinberg, D. H., Mortonson, M. J., Eisenstein, D. J., et al. 2013, Physics Reports, 530, 87–255

Wikipedia contributors. 2021, Conjugate prior — Wikipedia, The Free Encyclopedia, [Online; accessed 2-September-2021]

Zhao, C., Variu, A., He, M., et al. 2021, The completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: Cosmological implications from multi-tracer BAO analysis with galaxies and voids

Zou, H., Zhou, X., Fan, X., et al. 2019, The Astrophysical Journal Supplement Series, 245, 4

Zuntz, J., Lanusse, F., Malz, A. I., et al. 2021, The Open Journal of Astrophysics, 4

Željko Ivezić, and the LSST Science Collaboration. 2018, The LSST System Science Requirements Document