UNIVERSITY OF CALIFORNIA,
IRVINE

Probing the Dark Universe at the Pixel-Level in Large-Scale Surveys

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Physics

by

Alexander Broughton

Dissertation Committee:
Professor Simona Murgia, Chair
Professor David Kirkby
Professor Pierre Baldi

2024

Part I © 2023 American Physical Society
Part II © 2024 Alexander Broughton

DEDICATION

To my family and friends.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

This thesis would not have been possible without the kind and generous support of many friends, family, colleagues, and institutions.

First, I would like to thank **Simona Murgia** for opening her research group and giving me the freedom to explore any and all scientific questions as well as the resources and guidance to establish myself within the professional scientific community. Through her leadership, and thankfully, empathy and endless optimism, she has has shown me what it means to be a good scientist, and above all, an ethical one.

I would also like to thank my fellow collaborators in the Department of Physics and Astronomy, **Johanna Paine**, **Thomas Li**, **Chris Karwin**, and in the Department of Computer Science, **Pierre Baldi**, **Alex Shmakov**, and **Mohammadamin Tavakoli**, who in particular worked on the machine learning models in Part I of this thesis.

Of course I would also like to thank the fantastic research team at SLAC National Lab, where I spent a large fraction of my time and where much of the work for this thesis was performed. In particular, I would like to express my gratitude to **Aaron Roodman**, who was like a second advisor to me. He welcomed me into the LSST camera team under his kind leadership, both during my graduate school career and again in my new role as a postdoctoral researcher. Special thanks to **Yousuke Utsumi**, for his unwavering encouragement. His expertise was indispensable, and he showed me not only how to be a more precise researcher, but also a more effective science communicator and collaborator.

Thanks to **Andrés Plazas-Malagón, Louise Edwards, Tianqing Zhang, Mike Jarvis, Jim Chiang, Andy Rasmussen, Pierre Astier**, who were all the best collaborators to bounce ideas off. Likewise, thanks to **Seth Digel** who always gave me thoughtful feedback on papers and talks. Also a big thanks to (in no particular order) **Chris Waters**, **Eli Rykoff**, **Yusra alSayyad**, **Merlin Fisher-Levine**, **Robert Lupton**, **Jim Bosch**, **Arun Kannawadi**, **Andy Rasmussen**, **Tony Tyson**, **Agnès Ferté**, **Phil Marshall**, **Pierre Antilogus**, **Craig Lage**, **Greg Madjeski**, **Andrew Bradshaw**, **Shuang Liang**, **Adam Snyder**, **Stuart Marshall**, **Pat Burchat**, and **Josh Meyers** for their combined instrumentational, computational, and theoretical expertise which enabled much of this thesis.

Special thanks to **Phil Mansfield**, the best roommate I could have ever asked for. He and **Eric Charles** have given me the best career advice at the times that I most needed it, and which I hope to pass on to others someday.

The friends that I've made along the way have made this at all possible. **Bernardita Reid-Guachalla**, who–oddly enough–has made me feel more at home in a country that isn't her own; **Aaron Anderson**, who helped introduce me to film photography; **Johnny Esteves**, who taught me to bachata; **John-Franklin Crenshaw**, the first friend I met within the Dark Energy Science Collaboration (DESC); and of course **Sydney Erickson**, **Mahlet Shiferaw**, **Theo Schutt**, **Sid Mau**, **Jadyn Anczarski**, **Josh Tong**, and **Sarah Gaiser**, with whom I've shared good eating, wine tasting, dancing, beaching, hiking, and so much more. However, the advice, wisdom, ambition, and values that we've shared has truly made this journey worthwhile.

I also wish to thank my loving and supportive family, who have given me everything I ever needed or could have asked for throughout all my academic adventures. My parents

Katie and **Jeff**, encouraged me to pursue every curiosity from an early age. And I'm also lucky to have an older sister, **Elizabeth**, who continues to inspire me with her grace and relentless passion for education. I love you.

And of course, absolutely none of this would have taken place without the generous support from the institutions who have recognized the importance of this research. Thank you.

VITA

Alexander Broughton

EDUCATION

Doctor of Philosophy in Physics — **2024**
University of California, Irvine — *Irvine, CA*

Bachelor of Science in Physics — **2019**
University of California, Davis — *Davis, CA*

RESEARCH

Graduate Student Researcher — **2022–2023**
SLAC National Accelerator Laboratory — *Menlo Park, California*

Graduate Student Researcher — **2019–2024**
University of California, Irvine — *Irvine, California*

Undergraduate Researcher — **2017-2019**
University of California, Davis — *Davis, California*

Undergraduate Research Assistant — **2016**
Lawrence Berkeley National Laboratory (LBNL) — *Berkeley, California*

TEACHING

Teaching Assistant — **2019**
University of California, Irvine — *Irvine, California*

HONORS & AWARDS

Department of Energy Office of Science Graduate Student Research (SCGSR) Fellow — **2022**
Department of Energy (DOE)

Data Science Fellow (DSF) — **2021**
Physical Sciences and Machine Learning (PSML)

REFEREED JOURNAL PUBLICATIONS

Instrument Signature Removal and Calibration Products for the Rubin Legacy Survey of Space and Time **2024**
Journal of Astronomical Telescopes, Instruments, and Systems (JATIS)

Mitigation of the Brighter-Fatter Effect in LSSTCam **2024**
Publications of the Astronomical Society of the Pacific (PASP)

Photometry, Centroid and Point-Spread Function Measurements in the LSST Camera Focal Plane Using Artificial Stars **2023**
Publications of the Astronomical Society of the Pacific (PASP)

Deep Learning Models of the Discrete Component of the Galactic Interstellar γ-ray Emission **2023**
Physical Review D

Improved Modeling of the Discrete Component of the Galactic Interstellar γ-ray Emission and Implications for the Fermi–LAT Galactic Center Excess **2023**
Physical Review D

ABSTRACT OF THE DISSERTATION

Probing the Dark Universe at the Pixel-Level in Large-Scale Surveys

By

Alexander Broughton

Doctor of Philosophy in Physics

University of California, Irvine, 2024

Professor Simona Murgia, Chair

Future astrophysical surveys will provide unprecedented datasets in both scale and precision for testing our most concordant theories of dark matter and dark energy. For the first time, the confidence of our understanding of the components of the Universe will be limited by systematic uncertainties rather than the statistical precision we can achieve. Improved modeling of measurement biases in survey data is necessary to achieve the theory-constraining power of ambitious future astrophysical surveys. Recent in-lab, on-sky, and simulated experiments offer opportunities to investigate coherence between systematic and phenomenological observables and invent new techniques to differentiate them. This thesis details my contribution to measuring the Dark Universe, and it is structured in two parts.

The first part of this thesis improves template modeling of the foreground gamma-ray emission in the Galactic plane in 12 years of Fermi-LAT data. This has implications for the observed excess γ-ray emission near the Galactic center, which was previously associated with dark matter. We compare the likelihood of observing gamma-ray emission from cold hydrogen gas clouds using models derived from spectral line emission maps of abundant (^{12}CO) and rare (^{13}CO) tracer molecules. Using this approach, the rare tracer predicts previously unmodeled gamma-ray emission at 47σ significance, localized to the pixels around the densest regions, where emission is undercounted due to reabsorption along the line-of-sight. We also identify unassociated point-like sources in the Galactic plane which were

previously smoothed out by re-absorption consistent with peaked gas structure as well as point-like gas structure which could have been mistaken for millisecond pulsars by earlier works. We further utilize a convolutional neural network (CNN) to try to predict the ^{13}CO distribution in the same region, which could prove useful to predict the ^{13}CO distribution in data-poor regions such as the Galactic center.

The second part of this thesis develops a correction for the observed size-magnitude relation of stars, which are used calibrate observables for the Legacy Survey of Space and Time (LSST). The brighter-fatter effect results in distorted images of bright calibration stars, creating a flux-dependent point spread function (PSF) that if left unmitigated, could make up a large fraction of the error budget in Stage IV weak-lensing (WL) surveys such as the Legacy Survey of Space and Time (LSST). Using data from the LSST camera, we compare the BFE in flat field and star images and test the ability of scalar correction methods to mitigate it. We observe that the BFE evolves anisotropically as a function of flux due to higher-order BFEs which were previously assumed to be negligible by a commonly used correction technique. We also find that poor assumptions of the correction result in unphysical violations of flux conservation in the images at the O(0.01%) level. Based on these analyses, we then propose a new, unbiased correction methodology to account for these effects. With these improvements we can correct on average 95% of the total effect in the sensors we tested. We use these results to predict the impact of these previously unmodeled effects on sensitive observables of inferred cosmological outcomes and advise data processing strategies to limit the impact of the effects on LSST science pathways.

Part I

The Fermi Gamma-ray Space Telescope

Chapter 1

Introduction

Astrophysical observations over the past century have played a critical role in the discovery of dark matter (DM), and it continues to be a promising avenue to test predictions of different particle dark matter scenarios that potentially lie beyond the Standard Model (SM). In the 1930s, Fritz Zwicky conducted his famous "weighing" of galaxies experiment, which found that the mass of the Coma galaxy cluster was much larger than the mass of the observed luminous matter and coined the term "dark matter" to fill the gap in our understanding. Later, in the 1970s, observations by Vera C. Rubin confirmed that the stable rotations of the outer regions of galaxies were faster than could be explained by the gravitational effects of luminous matter alone. Around the same time, it became apparent that Big Bang nucleosynthesis could not have produced enough baryons to explain these observations throughout the Universe. Since the 1990s, a variety of numerical simulations and observations of the Cosmic Microwave Background (CMB), type Ia supernovae, and gravitational lensing of dynamical galaxy systems, show overwhelmingly that dark matter exists, and that galaxies are primarily made up of dark matter. However, the nature of dark matter is still elusive.

Current theories of particle dark matter span 90 orders of magnitude, ranging from ultra-

light axions to weakly-interacting massive particles (WIMPs) and even to planet/star mass objects (massive compact halo objects or MACHOs, such as primordial black holes). Searches for these particles on on Earth, including *collider* experiments (SM + SM $\rightarrow$ DM + DM, e.g. ATLAS Collaboration et al. 2022) and *direct detection* experiments (DM + SM $\rightarrow$ DM + SM, e.g. Aalbers et al. 2023), have found a few interesting events yet nothing statistically definitive. The ability of these probes to detect particle dark matter depends on the relative content of dark matter available, the interaction couplings involved, and the detector constraints. Direct detection experiments have very good sensitivity, but are limited in scale and need to contend with large backgrounds (Gaitskell et al. 1996). And while collider experiments have well-determined backgrounds, technical constraints (data volume, size, efficiency) have limited their testable parameter space.

The space of untested theoretical dark matter scenarios is still rich (for an overview see Tait 2021), and the astrophysical landscape offers a laboratory with regimes that we can not reproduce on Earth. The most promising candidate to search for in this space is the WIMP dark matter particle, which is predicted to self-annihilate into high-energy SM products (e.g. γ-ray photons, quark/anti-quark pairs, and even neutrinos). *Indirect detection* of dark matter (DM + DM $\rightarrow$ SM + SM) involves looking for the remnants of these exotic processes. Studies of the gravitational influences of dark matter from galaxy rotation curves show that dark matter is concentrated in the interior regions of the Milky Way Galaxy, making it the most likely location to potentially observe these rare and energetic phenomena. However, indirect detection probes for WIMPs can be very sensitive to background models of other sources of high mass-energy SM particles (Murgia 2020a). Systematic errors from background modeling are the largest hurdle to confirming or ruling out the indirect-detection of self-annhilating dark matter particles.

1.1 The Galactic Center γ-ray Excess

The Fermi Gamma-ray Space Telescope has been surveying the γ-ray sky since June 2008, and has observed potential signals from the products of self-annihilating dark matter particles. The primary on-board detector is the Large Area Telescope (Fermi-LAT), which is a particle tracker made of towers of layered high-Z foils coupled to a calorimeter, which allows it to reconstruct the energy and trajectory of incoming γ-rays (Atwood et al. 2009). Long-standing analyses of the Fermi-LAT data, first by Goodenough & Hooper (2009) and later by Hooper & Goodenough (2011); Hooper & Linden (2011); Abazajian & Kaplinghat (2012); Hooper & Slatyer (2013); Gordon & Macías (2013); Abazajian et al. (2014); Daylan et al. (2016); Calore et al. (2015); Ajello et al. (2016), report an unassociated excess of γ-ray emission, consistent with a spherical morphology that is brightest near the Galactic Center region (coincident within 0.05° the black hole Sgr A*). It extends approximately 15° (< 2 kpc) from the Galactic plane (Daylan et al. 2016; Calore et al. 2015). The Galactic Center Excess (GCE) makes up $5-10\%$ of the γ-ray emission above 1 GeV within this region. The GCE has been the fascination in the astrophysics community and the subject of increasingly detailed analyses for 15 years in part because of its tanatalizing agreement with the WIMP self-annhilation dark matter particle scenario in both morphology and spectrum, which has been reproduced by multiple independent analyses.

In general, the distribution of the γ-ray flux from self-annhilating dark matter across a solid angle element ($d\Omega$) depends on the density distribution of the dark matter particles:

$$\frac{dR(\theta)}{d\Omega} = N_T \frac{1}{2} \frac{\langle \sigma_{ann} v \rangle}{m_{DM}^2} \frac{1}{4\pi} \int_{\text{los}} ds \rho^2(\vec{r}(s, \vec{n})), \tag{1.1}$$

where m_{DM} is the mass of the hypothetical dark matter particle, $\langle \sigma_{ann} v \rangle$ is the velocity averaged self-annihilation cross-section, N_T is the number of nuclei in the detector, which

acts as a detection efficiency factor. The observed density distribution ($\rho(r)$) is consistent with a Navarro–Frenk–White (NFW) profile:

$$\rho(r) = \rho_0 \left(\frac{r}{R_s}\right)^{-\gamma} \left(1 + \frac{r}{R_s}\right)^{\gamma-3}. \tag{1.2}$$

with $\gamma = 1.1 - 1.3$, a characteristic scale radius $R_s = 20$ kpc, and a characteristic density consistent with the local dark matter density $\rho_0 = 0.3$–0.4 GeV cm^{-3}. Interestingly, an NFW profile is consistent with measurements of the Coma cluster's dark matter halo density distribution, and a distribution with $\gamma \sim 1$ is consistent with predictions from N-body simulations of cold, collisionless, virialized (equilibrated) dark matter density distributions after baryonic feedback effects are included (Gnedin et al. 2004). The GCE is found to be consistent with a predicted distribution of WIMP dark matter with a mass of approximately $30 - 60$ GeV, annihilating with a velocity-averaged cross-section of order $\sim 10^{-26}$ cm^3 s^{-1} with a dominate annihilation mode into a bottom quark hadron (Hooper & Linden 2011; Abazajian & Kaplinghat 2012; Gordon & Macías 2013; Daylan et al. 2016; Calore et al. 2015).

Similarly, the expected spectral density distribution of self-annhilating dark matter would be given by:

$$\frac{dN_\gamma}{dE} = \frac{\langle \sigma_{ann} v \rangle}{4\pi\eta m_{DM}} \frac{dN_\gamma}{dE} \times \int_{\Delta\Omega} d\Omega \int_{\text{los}} ds \rho^2(\vec{r}(s, \vec{n})) \tag{1.3}$$

Most analyses agree that the most of the spectrum across the GC region comes from a band between 1-3 GeV (Goodenough & Hooper 2009), however the soft (sub-1 GeV) and hard (super-10 GeV) power-law regimes of the GCE spectrum are still debated. Some analyses have observed a power-law tail that can extend beyond 10 GeV (Ajello et al. 2016). In addition, both Abazajian et al. (2014) and Horiuchi et al. (2016) argue that the GCE spectrum varies spatially and exhibits a higher energy spectrum in regions above and below

the Galactic plane. However it is generally agreed that the peak of the excess alone is consistent with theory predictions of WIMP annihilation into SM particles. In addition, the inferred annihilation cross-sections (σ_{ann}) from the measured spectra are consistent with the weak-scale annihilation cross-section needed to produce the relic abundance of dark matter measured by observational probes.

1.2 Interstellar Emission Models (IEMs)

The spectrum and spatial morphology of the GCE is crucial to identifying its source, and these properties depend on the removal of foreground and background γ-ray emission sources. An estimation of the uncertainties by Calore et al. (2015) outside the Galactic Center finds that the presence of the GCE above background is robustly above statistical uncertainty, but discrepancies between its morphology and spectrum can be traced back to different treatments of interstellar emission models. And interpretations of the GCE can change within the systematic uncertainties of these different emission models. None of the models considered in previous analyses has been able to reach Poisson noise-level agreement with the observational data, and properties of the GCE are strongly correlated with the properties of these other components. This can lead to poorly-constrained inferred properties and interpretations of the GCE. In this section we will cover the different types of γ-ray emission sources, their uncertainties, and how they can impact the characterization and interpretation of the GCE.

We will discuss two types of γ-ray background components in the interstellar region, a diffuse component and a discrete component. The diffuse component of emission originates from cosmic ray (CR) interactions with the interstellar medium: (i) CR protons collide with other free protons or protons within nuclei of interstellar gas and hadronize into neutral pions which quickly decay in to pairs of γ-rays ($pp \rightarrow X + \pi^0 \rightarrow X + \gamma\gamma$); (ii) high-energy CR

electrons upscatter starlight in the interstellar radiation field into γ-rays via inverse Compton (IC) scattering; (iii) high-energy cosmic ray electrons interact with gas in the interstellar medium and emit bremsstrahlung radiation at γ-ray energies (although this component is typically sub-dominant). Diffuse emission can also originate outside the interstellar medium, for example extragalactic active galactic nuclei (AGN) or blazars, or the Fermi bubbles of our own galaxy (two lobes of hot plasma extending $\pm 50,000$ly from the Galactic plane, which are thought to be remnants of earlier outflows from the Galactic Center and emit γ-rays via IC scattering. In addition, there are discrete sources of GeV level γ-rays from pulsars and X-ray binaries. Pulsars are rotating neutron star remnants (with periods ranging from milliseconds to seconds) from supernova explosions, and the subclass of millisecond pulsars (MSP) emit photons strongly at GeV (γ-ray) scales. In addition, X-ray binaries (particularly high mass X-ray binaries or HMXBs) are binary systems with a star (larger than about $10M_{\odot}$) and either a neutron star or a black hole that emit strongly in γ-rays via relativistic jets.

These models are derived as "templates" which are pixel-level counts maps of the sky that are binned by energy, and the total observed γ-ray sky is a linear combination of these templates (see Murgia 2020a). The templates can be derived in two ways, and both introduce their own systematic uncertainties. One method involves fitting a set of pre-selected physical parameters (e.g. γ-ray emissitvity, cosmic-ray spectra and radial dependence, etc.) to the γ-ray data, however this risks over-fitting the data and eating up the GCE. Another way to construct interstellar emission models is to simulate the propagation of known physics with empirically measured initial conditions to the resultant templates, and while these models typically agree with observations to the 10-20% level, they do not fully capture the realistic complexity of processes and regions in the Galaxy. All template models depend on the distribution of sources, particle propagation, the distribution of the interstellar medium, and the completeness of our understanding of the physical sub-processes.

1.3 Alternate Interpretations

One possible interpretation of the GCE is a distribution of unresolved point sources. The 12-year 4FGL-DR3 data release by Fermi-LAT (Ballet et al. 2023) identified 5,064 distinct point sources, however the resolution likely undercounts the number in the Galactic Center. The confirmed MSPs have spectra that peak between 1-3 GeV, which is coincident with expected peak from the dark matter annhilation scenario, and with the Galactic Center Excess (Abazajian & Kaplinghat 2012; Linden et al. 2016; Abazajian 2011; Petrović et al. 2015). To distinguish between these components, the low-energy and high-energy regimes need to be modeled both accurately and precisely. Below 1 GeV, the spectra of MSPs falls off gradually while the predicted dark matter spectrum falls off sharply. In addition, the dark matter annhilation scenario predicts a spectrum that could extend beyond 10 GeV, however MSPs are invisible above this level (Linden et al. 2016). However, the low- and high-energy regimes of the GCE spectrum vary across analyses with different assumptions, models, and analysis methods. For example, including the Fermi-bubble component, Ackermann et al. (2017a) finds that the GCE spectrum completely disappears above 10 GeV, although it is difficult to properly model this component, particularly near the Galactic Center, which can make it difficult to disentangle the Fermi bubbles from the other γ-ray components near the GC.

Composite likelihood analyses of subdivided galactocentric regions within the Galaxy by Horiuchi et al. (2016) find that the spectral energy distribution is not uniform across the GCE region. Depending on the background models used, the GCE emission spectrum can have higher-energy at higher latitudes, and in some analyses the GCE is not spherically symmetric and the GCE centroid could be pulled away from the Galactic Center by as much as 1° (Ajello et al. 2016; Linden et al. 2016; Karwin et al. 2017). If these findings are true, it would disfavor a single-component DM scenario and favor multiple origins associated

with background modeling.[1] Higher-energy at higher latitudes could be explained by a population of high-latitude MSPs sourcing an extra IC scattering component. In fact, it has been proposed that a population of MSPs could exist in this region as relics of ancient globular clusters (Brandt & Kocsis 2015). However, even this would still not be able to rule out DM annhilation scenario. If the hypothesized WIMP could self-annhilate into e^+e^- pairs, there could be a secondary IC and bremsstrahlung emission component that would also explain the softened the spectrum closer to the Galactic plane, thus making it even more difficult to distinguish between DM an MSP interpretations (Daylan et al. 2016; Cirelli et al. 2013, 2014).

Evidence for the MSP interpretation is driven by a non-Poissonian template fitting (NPTF) analysis of photon-counting statistics (Lee et al. 2015; Lee et al. 2016). NPTF is a Bayesian statistical tool that looks for deviations from photon-counting Poisson fluctuations in pixels across extended regions of the sky as a result of unresolved, underlying point sources. The method infers a posterior distribution of the possible source function in a given region (dN/dF for N sources contribution a flux between F and $F + dF$). The analysis by Lee et al. (2016) found the GC excess can be fully accounted for by a population of order 100 point sources accumulating at a flux just below the minimum Fermi–LAT sensitivity threshold.

Other known dark matter accumulation sites outside the Galactic Center seem to offer new areas to reproduce results with different systematics. An investigation of γ-ray emission from dwarf spheroidal galaxies by Ackermann et al. (2017b)) and another investigation of nearby galaxies such as Messier 31 & 33 by McDaniel et al. (2024) find that the presence of excesses in these regions is questionable and the overall uncertainty is still dominated by model uncertainties. The space of possible interpretations is so large, and the GCE is so

[1]Note that this conclusion is highly model-dependent, and other studies have concluded that the high-energy and low-energy components favor the same radial profile (Linden et al. 2016; Karwin et al. 2017).

sensitive to different model and analysis choices that it is difficult to rule out any scenario. The assumptions and analysis choices that generate the interstellar emission models are the largest source of systematic uncertainty in indirect search for WIMP self-annhilation modes. The origin of the GCE cannot be resolved without improvements to the assumptions and methodologies for analyzing the background an foreground models in these analyses.

Chapter 2

Improved Modeling of the Interstellar Discrete H_2 Component

The Galactic γ-ray interstellar emission (IE) traces interactions of cosmic rays (CRs) with the interstellar medium and constitutes most of the γ-ray emission observed by Fermi–LAT. Uncertainties in modeling the IE are large and difficult to constrain, and they impact the study of other γ-ray sources in the Fermi–LAT data, point-like as well as extended. We focus on modeling the small-scale structure of the IE which, which if not robustly captured by the model, confuses the determination of point sources, especially along the Galactic plane (Abdollahi *et al* 2020). It was shown in Ajello et al. (2016) that a large fraction of the point sources detected by Fermi–LAT in the Galactic Center (GC) region could be misidentified gas structure, and strongly dependent in number and spatial distribution on the IE model employed to analyze the data. This result indicates the presence of point-like emission in the data arising from unmodeled structure in the interstellar gas, and it underlines the importance of accurately modeling this component to reliably identify point sources in the γ-ray data. An investigation of the 4FGL-DR3 catalog by Abdollahi *et al* (2022) reported evidence for clustering of unassociated, spectrally soft sources along the Galactic plane that

appears to be tracing interstellar clouds. This emission could be originating from small-scale molecular gas structure.

More reliable modeling of the small scale gas structure could also impact the characterization of extended sources. A prominent example is the Fermi–LAT GC excess (see e.g. Goodenough & Hooper 2009; Hooper & Goodenough 2011; Hooper & Linden 2011; Abazajian 2011; Abazajian & Kaplinghat 2012; Hooper & Slatyer 2013; Gordon & Macías 2013; Abazajian et al. 2014; Daylan et al. 2016; Calore et al. 2015; Ajello et al. 2016). Striking features of this excess are its spatial morphology and spectrum which could be consistent with annihilating dark matter (DM). Alternative explanations have been proposed, with the leading one attributing the signal to a collection of discrete emitters such as an unresolved point source population of millisecond pulsars. Deviations from a spherically symmetric morphology for the excess have also been claimed which would imply the emission to be compatible with the stellar distribution in the Galactic bulge (Macias et al. 2018; Macias et al. 2019; Bartels et al. 2018; Pohl et al. 2022). The origin of the *Fermi*-LAT GC excess remains a debated topic (Murgia 2020b). This debate could be settled by determining whether the spatial morphology of the excess is consistent with a smooth distribution, as predicted for DM, or with the cumulative emission of a collection of point-like emitters. Several statistical techniques have been employed to accomplish this (see Bartels et al. 2016; Lee et al. 2016; Leane & Slatyer 2019; Zhong et al. 2020; Leane & Slatyer 2020; Leane 2020; Buschmann et al. 2020; Calore et al. 2021; List et al. 2021; Mishra-Sharma & Cranmer 2022), including the non-Poissonian template fit (NPTF) which can detect upward fluctuations in the photon statistics above Poisson noise which are associated with a collection of point sources. However, the results of the NPTF technique depend on the modeling of the IE (e.g. Leane & Slatyer 2019; Leane & Slatyer 2020; Leane 2020; Buschmann et al. 2020; List et al. 2021), and an uncontroversial resolution to the origin of the GC excess has not yet been reached. We posit that if the interstellar gas is more structured and point-like than current models

predict, the unresolved point source contribution in the data could be erroneously inflated by the fainter component (also below detection threshold) of the small scale gas features. In particular, statistical methods such as the NPTF might attribute this component to a smooth GC excess and conclude it is point-like. These uncertainties might therefore hinder the ability to distinguish the smooth versus point-like duplicity of the excess. We note that wavelet decomposition is another statistical technique that has been employed to resolve the GC excess (Bartels et al. 2016). While these studies are not as impacted by the IE model, the related results on the nature of the GC excess remain uncertain (Balaji et al. 2018; Zhong et al. 2020).

In this work, we present a novel approach to improve modeling of the small-scale structure in the interstellar gas. The IE is due primarily to CRs interacting with the interstellar hydrogen gas (and radiation field), in molecular (H_2), atomic (HI), and ionized (HII) forms. The H_2 component is the most highly structured, and this structure is traced by the related γ-ray emission. In this work, we focus on H_2 gas because of its high degree of structure and our methodology hinges on the availability of additional proxies better suited to trace its structure. The impact of the other components of the γ-ray IE (specifically HI and dark gas Grenier et al. 2005), and related uncertainties, are not considered in this study and will be the focus of later work. Since H_2 does not emit at a characteristic radio frequency, other molecules are used to trace its distribution. In particular, carbon monoxide (^{12}CO, or CO hereafter) is typically employed as a proxy. Radio surveys trace the distribution of CO across the sky and the H_2 column density is inferred by scaling the CO content by a conversion factor (X_{CO}) which gives the ratio of the integrated CO line emission to the H_2 column density. The bulk of the H_2 is traced following this methodology, and the survey of the CO J=1–0 transition line from Dame et al. (2001) has been widely employed. However, the CO line emission is typically optically thick in the denser cores of molecular clouds and it underestimates the total H_2 column density there. γ-ray IE models that employ CO to trace

H_2 gas may therefore underestimate its finer structure. This limitation can be addressed by exploiting CO isotopologues, ^{13}CO and C^{18}O, also found in H_2 clouds. The line emission from these rarer isotopologues remains optically thin to larger column densities and therefore is more reliable to probe dense cores.

2.1 Tracers of H_2

We employ the data from the Mopra Southern Galactic Plane CO Survey (data release 3, Braiding et al. 2018) to trace the denser H_2 regions. The survey covers a 50 square degree region, spanning Galactic longitudes $l = 300°$-$350°$ and latitudes $|b| \leq 0.5°$, and it targets the $J = 1$–0 transitions of CO, ^{13}CO, C^{18}O, and C^{17}O. Mopra is a 22 meter single dish radio telescope located ∼450 km from Sydney, Australia, targeting the $109 - 115$ GHz $J = 1 - 0$ transitions of CO, ^{13}CO, C^{18}O, and C^{17}O. It utilizes a "Fast-On-The-Fly" mapping in 1 square degree segments, with each segment being comprised of orthogonal scans in longitudinal and latitudinal directions. Each square degree has an exposure of at least ∼20 hours. The data reduction process by the Mopra team, which involves six main stages of processing to perform the band-pass correction and background subtraction, yields data cubes of the brightness temperature for each spectral line, as a function of Galactic coordinates and local standard of rest velocity. The spectral resolution is 0.1 km sec^{-1}. To efficiently work with the Mopra observations we create a mosaic of the data over the full 50 square degree region for both CO and ^{13}CO. These maps are shown in Figure 2.1, where the color scale gives the total line strength, $W(\mathrm{CO})$, integrated over all velocities (-150 to 50 km s^{-1}). The vertical black lines show the footprint of the mosaic, i.e. regions where pixels from adjacent fields overlap. As can be seen, the emission shows discontinuities in these footprint regions. This is likely due to the fields having different noise levels resulting from the varying weather conditions of the observations. To avoid any potential complications

with these regions they are masked in the analysis. Additionally, we slightly degrade the spatial resolution from 0.0083° to 0.03125°. This is done to speed up the computation time of the γ-ray intensity maps (see Section 2.3). We note that the CO maps customarily used to trace H_2 (Dame et al. 2001) have a spatial resolution of 0.25°, and thus the Mopra maps that we employ represent an 8× improvement. The emission shown in Figure 2.1 also displays vertical and horizontal strips. This is a systematic feature due to the scanning method used in the observations, and it does not have any significant impact on our analysis.

To ensure the highest quality data in the rare isotopologues, only pixels for which the brightness temperature exceeds the 1σ noise provided by Mopra are used. Noise removal is necessary since the noise could otherwise predict spurious γ-ray emission (see Section 2.2 for more details.) The noise threshold is chosen for consistency with the next stage of this work, where we will train machine learning (ML) models on these data to predicts ^{13}CO in regions of the sky where observations are not available. Because the optimization of ML models learns about the noise, our choice favors retaining more data for training purposes (see §3). In this analysis we utilize the CO and ^{13}CO data, as the C^{18}O emission is extremely sparse, and the C^{17}O emission was too faint to be detected in the survey. For the region it surveys, the Mopra data provide a sharper view of the CO emission compared to Dame et al. (2001), in addition to probing the rarer H_2 tracers.

2.2 Deriving H_2 column densities

We calculate the H_2 column density corresponding to Mopra's CO and ^{13}CO, referred to as $N(H_2)_{CO}$ and $N(H_2)_{CO13}$, respectively. Following the method of Ackermann et al. (2012), the gas is separated into 17 Galactocentric radial bins (or annuli) based on its velocity and therefore corresponding to different distances from the GC, assuming uniform circular motion

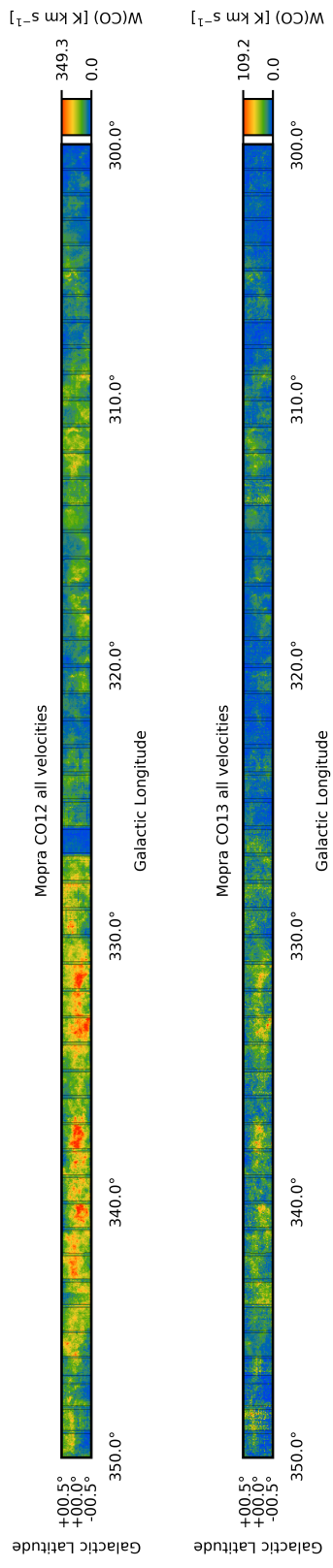


Figure 2.1: Total line strength integrated over all velocities (-150 to 50 km s^{-1}) for CO (top) and ^{13}CO (bottom) from Mopra. The color scale is set to 2/3 of the maximum value, for illustration, and no smoothing is applied. The vertical black lines delimit the footprint of the mosaic, and the narrower region they enclose is masked in the analysis (see text).

about the GC (see Appendix A for details). The $N(H_2)_{CO}$ is determined as:

$$N(H_2)_{CO} = W(CO) \times X_{CO}, \tag{2.1}$$

where $W(CO)$ is the line strength of the CO gas, and we adopt the radially-dependent X_{CO} from Ackermann et al. (2012) for the 17 radial bins. An alternative assumption of a constant X_{CO} could also be considered when our methodology is applied to the data. $N(H_2)_{CO13}$ is evaluated as:

$$N(H_2)_{CO13} = N(CO13) \times \Big[\frac{H_2}{CO13}\Big]. \tag{2.2}$$

where $N(CO13)$ is the ^{13}CO column density, and the abundance ratio $[\frac{H_2}{CO13}]$ is the conversion factor for the rarer species, ^{13}CO in this case. Following Wilson et al. (2009); Braiding et al. (2018), we derive $N(CO13)$ assuming that the gas is in local thermodynamic equilibrium at a fixed excitation temperature of 10 K (the procedure is summarized in Appendix B). Since the conversion to column densities involves multiplying the line strength observed by Mopra by large factors, of order $\sim 10^{20}$, noise in regions with no gas emission would yield significant column density and, in turn, spurious predictions for the related γ-ray emission. This is especially problematic for our ultimate goal of comparing $N(H_2)_{CO12}$ to $N(H_2)_{CO13}$. To mitigate this issue we use the noise maps which have been provided by the Mopra team, as anticipated in Section 2.1. For each spatial pixel, the noise is determined from the standard deviation of the continuum channels for line-of-sight velocities containing no apparent signal. The observations were carried out over a wide variety of weather conditions, and this is reflected in the noise maps, which are not uniform over the Mopra region, but rather vary from field to field. In calculating the column density we only use pixels that are above the 1σ noise level. Otherwise, pixels that are below the 1σ noise level are masked. The abundance ratio $[\frac{H_2}{CO13}]$ is not tightly constrained. We choose the upper bound of the range (2.7 - 7.5)

$\times 10^5$ provided in Pineda et al. (2008), in an attempt to assess the maximal impact on the γ-ray data, although it is not excluded that larger values are possible.

2.3 Excess γ-ray emission from the H_2 component

The CR propagation code GALPROP (v56) (Moskalenko & Strong 1998, 2000; Strong & Moskalenko 1998; Strong et al. 2000; Ptuskin et al. 2006; Strong et al. 2007; Vladimirov et al. 2011; Jóhannesson et al. 2016; Porter et al. 2017; Johannesson et al. 2018; Génolini et al. 2018) is used to evaluate the γ-ray intensity maps for the H_2-related emission traced by CO and ^{13}CO. GALPROP self-consistently calculates spectra and abundances of Galactic CR species and associated diffuse emissions (γ-rays, but also radio, X-rays) in 2D and 3D. We adopt the same GALPROP input parameters as described in Karwin et al. (2019). GALPROP returns radially dependent γ-ray intensity all-sky maps (in the 17 Galactocentric annuli) which allow us to determine, for the region observed by Mopra, the additional contribution in γ-rays from the H_2 traced by ^{13}CO, and assess its significance in simulated Fermi–LAT data.

The difference between $N(H_2)_{CO13}$ and $N(H_2)_{CO}$ gives an estimate of the H_2 that is missed in dense regions when only CO is used as a tracer. To calculate the corresponding intensity in γ-rays, we define the "Modified Map". This is calculated in the following way: for pixels with $N(H_2)_{CO13} > N(H_2)_{CO12}$, the value of $N(H_2)_{CO12}$ is replaced with the value of $N(H_2)_{CO13}$. We define the "Modified Excess Template" as the difference between the Modified Map and the baseline CO map, which accounts for the additional H_2-related γ-ray emission not included in current IE models, shown in Figure 2.2.

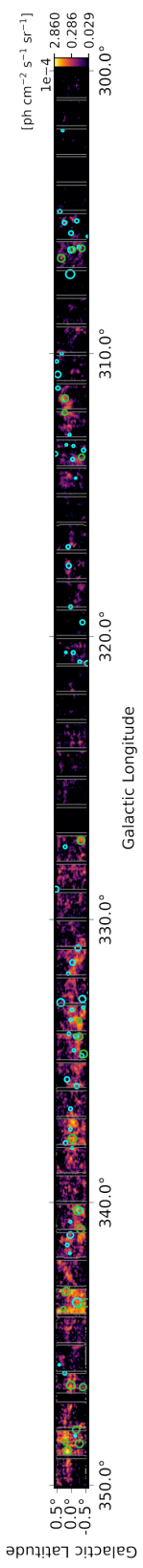


Figure 2.2: The Modified Excess Template, calculated as the difference between the Modified Map and the baseline CO map. The color scale shows the γ-ray intensity. The solid cyan circles show unassociated γ-ray point sources in version 4FGL-DR2 of the fourth Fermi–LAT catalog (Ballet et al. 2020), and the green circles show new sources that we find in this work (see text). The radius of each source corresponds to the 95% localization uncertainty.

2.4 Likelihood of excess γ-ray emission

We evaluate the significance of the Modified Excess Template in the Fermi–LAT data by simulating the data collected between 2008 August 04 to 2020 November 11 ($\sim$12 years). The simulated events have energies in the range $1 - 100$ GeV and are binned in 8 energy bins per decade, for event class P8R3_CLEAN (FRONT+BACK). The analysis is performed using Fermipy (v0.19.0), which utilizes the Fermitools (v1.2.23). In these simulations, we only focus on the H_2-related γ-ray emission, and exclude all other components of the γ-ray sky. The goal is to assess the significance of the Modified Excess Template, i.e. the contribution of the newly modeled H_2 fine structure, in the optimistic scenario where all other components are satisfactorily modeled. The simulated events trace the H_2-related γ-ray emission modeled with the Modified Map. The simulated data are then fit based on a binned maximum likelihood method to a model that includes two components, the γ-rays traced by the baseline CO map and the Modified Excess Template. The latter is assigned the spectrum determined by GALPROP (see Appendix C), and its normalization is free to vary in the fit. The normalization of the CO map is also free to vary and its spectrum constrained to that calculated by GALPROP. As mentioned above, the γ-ray flux is calculate in 17 radial bins, since the predicted H_2-related γ-ray emission depends on the CR density, which is a function of Galactocentric radius. In the simulations, the total emission is integrated along the line-of-sight. Moreover, the individual maps have a high level of degeneracy. We therefore combine the maps into 4 radial bins. Specifically, we combine bins 1-6, 7-10, 11-13, and 14-17, which we refer to as A1, A2, A3, and A4.

We simulate 1000 realizations of Fermi–LAT data, and calculate the Test Statistics (TS) for the nested models: $-2\log(L_0/L)$, where L_0 corresponds to value of the likelihood function for the null hypothesis (CO baseline) and L to the alternative hypothesis (CO baseline and Modified Excess Template). The statistical significance is approximated by $\sigma \approx \sqrt{\mathrm{TS}}$. The

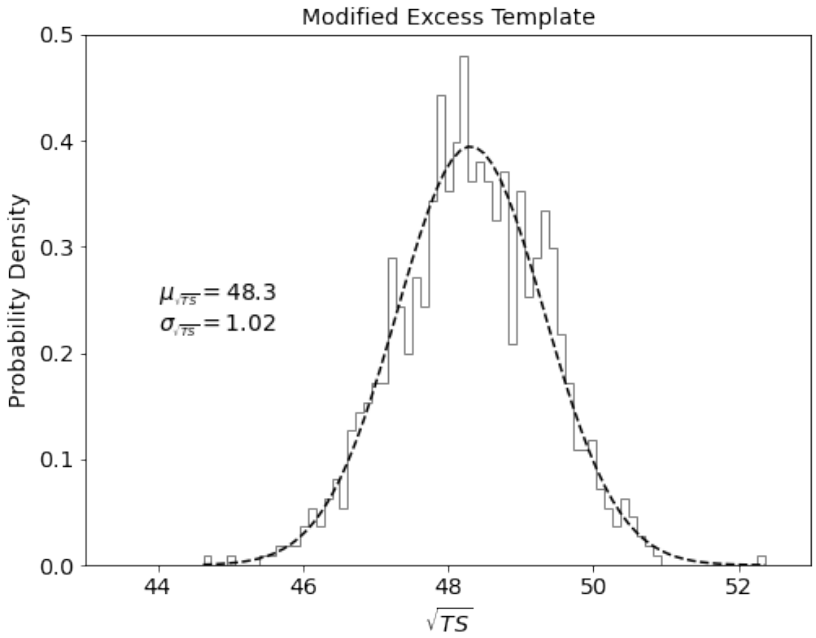


Figure 2.3: Distribution of the statistical significance ($\sigma \approx \sqrt{\mathrm{TS}}$) of the Modified Excess Template for the 50°×1° region covered by Mopra for 1000 realizations of ∼12 years of Fermi–LAT data. A fit with a Gaussian function is overlaid.

distribution of the $\sqrt{TS}$ for the 1000 simulations is shown in Figure 2.3. The mean of the distribution is 48.30±1.02, for the 50 squared degree Mopra region, and therefore very significant in a scenario where other components contributing to the Fermi–LAT data are perfectly modeled, and if the γ-ray emission traced by ^{13}CO is at the high end of the range we have considered (with the caveats discussed above.)

The fractional residuals as a function of energy for the 1000 simulations are shown in Figure 2.4. They are consistent with zero, as expected. Figure 2.5 shows the distributions of the flux of each model component, including the Modified Excess Template. In the Mopra region, the mean integrated flux for this component is $(8.3 \pm 0.2) \times 10^{-8}$ ph cm^{-2} s^{-1}.

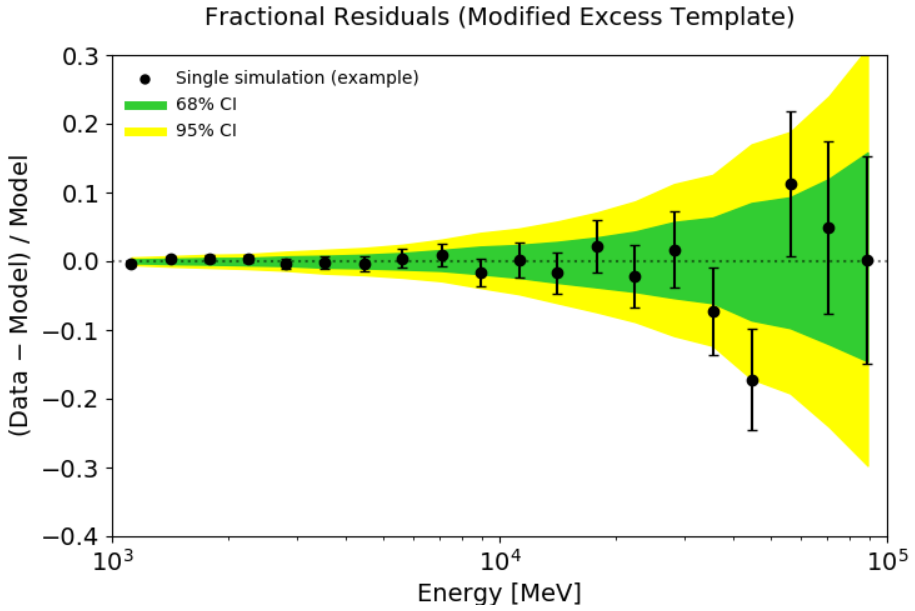


Figure 2.4: Fractional count residuals as a function of energy. The green and yellow bands show the 68% and 95% confidence regions from 1000 simulations, respectively. As an example, we also plot the results for a single simulation, shown with black data points.

2.5 Implications for the Galactic Center Excess

Overall, the Modified Excess Template accounts for a fair fraction (21%) of the total H_2-related γ-ray emission in the Mopra region. For comparison, the integrated GC excess flux in a $15^\circ \times 15^\circ$ region around the GC (the full region is considered), which is $4\times$ larger than the region observed by Mopra, is in the range $18.3 - 25.0 \times 10^{-8}$ ph cm^{-2} s^{-1} (from Ajello et al. 2016). In intensities, the GC excess corresponds to $2.67 - 3.65 \times 10^{-6}$ ph cm^{-2} s^{-1} sr^{-1} compared to 5.43×10^{-6} ph cm^{-2} s^{-1} sr^{-1} for the Modified Excess Template. Because of the different spatial morphology of the GC excess and H_2-related γ-ray emission traced by ^{13}CO, a direct comparison is unwarranted and we do not expect the latter to account for the majority of the GC excess. However, the H_2 emission extends beyond the latitudes considered in this study and, albeit dimmer at higher latitudes, the estimates provided here do not indicate its small scale component to be negligible. We estimate this contribution in

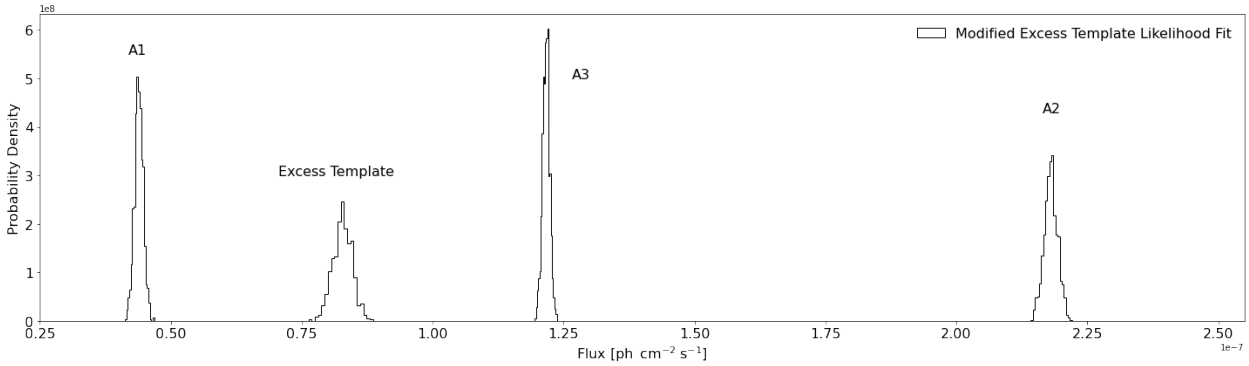


Figure 2.5: The distributions of fluxes for each component of the best-fit model using the Modified Excess Template over the full region for 1000 realizations of $\sim$ 12 years of Fermi–LAT data. Note that one of the annuli (A4) has negligible contribution and the normalization was fixed in the fitting process (corresponding to a flux of 4.98×10^{-9} ph cm^{-2} s^{-1}).

the GC region by considering, as above, a $15^\circ \times 15^\circ$ region around the GC. In this region, the contribution from H_2-related emission (π^0 decay only) traced by CO is $\sim 55 \times 10^{-8}$ ph cm^{-2} (Ajello et al. 2016). If we adopt the same fraction of unmodeled pointlike emission for H_2 as we find in the Mopra region, i.e. approximately 21%, this fraction translates into a very rough estimate of this point-like contribution in the same region of the GC excess, i.e. about 11.5×10^{-8} ph cm^{-2}, or 46-63% of the GC excess for the same region. This estimate shows that the point-like emission could correspond to a significant fraction of the GC excess.

This contribution could therefore confuse the GC excess morphology, depending on the exact H_2 emission outside of $\pm 0.5^\circ$ in latitude, because the techniques that track the point-like fluctuations could erroneously ascribe unaccounted point-like emission (originating from H_2 gas in this case) to the GC excess (Leane & Slatyer 2019). Considerations based on the spectrum of the unresolved source population could be powerful to settle this degeneracy, however a robust determination is required.

2.6 Implications for point source detection

We identify the structure traced by ^{13}CO that is bright enough to be detected as a Fermi–LAT point source and compare it to the unassociated sources in version 4FGL-DR2 of the fourth Fermi–LAT catalog of γ-ray sources (Ballet et al. 2020). A significant overlap would indicate that its brightest contribution has already been detected by Fermi–LAT and validate the methodology to trace this component. To this end, we perform a likelihood fit where, instead of including the Modified Excess Template in the model, we only consider the baseline CO map and find additional point sources using the fermipy *find_srcs* function. The method calculates TS maps, and identifies point sources based on peaks in the TS. We use a test point source modeled with a power law spectrum with spectral index=2, and a minimum TS threshold of TS $\geq$ 16. An index of 2.7 is compatible with CRs interacting with gas, and if used, it yields consistent results. In total we find 23 new point sources with TS$\geq$16, which are shown with green circles in Figure 2.2. For comparison, we also overlay 4FGL-DR2 unassociated sources. We find that 8/23 (35%) of the new sources are spatially consistent with an unassociated 4FGL-DR2 source (based on an overlap of the 95% localization errors), which accounts for 8/46 (17%) of the total unassociated sources in the region. To quantify the probability that the associations happen by chance, we perform 1000 simulations, where for each realization we randomly distribute 23 sources in the Mopra region. For each source the Galactic coordinates are drawn from a uniform distribution and we assign an error radius drawn from a Gaussian distribution, with the mean and standard deviation determined from the 23 detected sources. From the 1000 simulations we find the average number of random associations to be 3.16 $\pm$ 1.72, which corresponds to a p-value (i.e. N($\geq$ 8)/1000) of 0.012. It is therefore likely that the brightest of the point-like emission traced by Mopra ^{13}CO have already been detected in the Fermi–LAT data.

2.7 Discussion

The results presented here demonstrate that there likely exists significant structure in the H_2-related γ-ray emission detected by Fermi–LAT which is not currently included in the IE models. Since its spatial morphology has point-like features, it directly affects the detection of resolved and, potentially, unresolved point source populations in the γ-ray data. More specifically, we have discussed the impact on the 4FGL-DR2 catalog and on the interpretation of the GC excess. Although in this analysis we have focused on the region covered by Mopra, similar conclusions would hold elsewhere because of the same limitations in tracing H_2. The methodology described here however can only be readily applied to limited regions of the sky because of the paucity of observations of the rare H_2 tracers. Unfortunately, we cannot simply extrapolate the Modified Excess Template to other regions, even with a less ambitious goal of providing only an estimate of this emission. This is because the Mopra data on which it is based is tightly confined, especially in latitude. To address the limited available observations, we have resorted to ML to develop a methodology that predicts the distribution of the small scale H_2-related γ-ray emission for other regions of the sky based on the existing Mopra data. Because of its complexity, we describe the ML related work in §3, where we show that we can train a convolutional neural network to predict ^{13}CO based on CO Mopra observations, and reproduce the H_2-related small scale structure traced by ^{13}CO within the statistical uncertainty of the Fermi–LAT data in the Mopra region. We conclude that our results justify applying this methodology to other regions of the sky by extending it to employ the all-sky CO survey (Dame et al. 2001).

2.8 Conclusions

In this paper, we assess whether there is point-like emission in the γ-ray data associated with the interstellar gas and not currently included in models of the Galactic interstellar γ-ray emission. We argue that this emission could significantly impact the detection of γ-ray point sources along the Galactic plane, and hinder the characterization of extended γ-ray sources, such as the GC excess. We employ the data from the Mopra Southern Galactic Plane CO Survey, which includes tracers of the small scale structure of the H_2-related γ-rays, to improve models by more accurately describing the point-like features in the gas. We find that significant point-like emission originates from H_2 gas and the significance could be as high as $\sqrt{TS} \sim 48$ (depending on assumptions) in a $50^\circ \times 1^\circ$ region covered by Mopra, corresponding to $\sim 21\%$ of the modeled H_2-related γ-ray emission in the region. We also show that this newly found point-like component may account for some fraction ($\lesssim$17%) of γ-ray point sources detected by Fermi–LAT in the Galactic plane whose origin is so far unknown. That a significant number of unidentified 4FGL-DR2 sources along the Galactic plane originate from unmodelled structure in the gas is not unexpected, but here we develop a robust and reliable methodology to identify the contribution of both bright and dimmer components of the H_2 gas discrete emission. We emphasize that our results depend on the assumptions we have adopted, e.g. a less conservative ^{13}CO abundance ratio. Because of the significance of this emission under plausible assumptions, we conclude that its contribution in the GC region could introduce a significant systematic uncertainty in determining whether the GC excess is smooth or point-like, and therefore more consistent with a dark matter or pulsar interpretation, specifically when relying on statistical techniques such as the NPTF. Identifying this component in the GC region could be a crucial step to settle the origin of the GC excess.

Chapter 3

Improved Modeling of the Interstellar Discrete H_2 Component with Deep Learning

A significant point-like component from the small-scale (or discrete) structure in the H_2 interstellar gas might be present in the Fermi–LAT data, but modeling this emission relies on observations of rare gas tracers only available in limited regions of the sky. Identifying this contribution is important to discriminate γ-ray point sources from interstellar gas, and to better characterize extended γ-ray sources. We design and train convolutional neural networks to predict this emission where observations of these rare tracers do not exist, and discuss the impact of this component on the analysis of the Fermi–LAT data. In particular, we evaluate prospects to exploit this methodology in the characterization of the Fermi–LAT Galactic Center GeV Excess (GCE) through accurate modeling of point-like structures in the data to help distinguish between a point-like or smooth nature for the excess. We show that deep learning may be effectively employed to model the γ-ray emission traced by these rare H_2 proxies within statistical significance in data-rich regions, supporting prospects to

employ these methods in yet unobserved regions.

The Galactic γ-ray interstellar emission (IE) traces interactions of cosmic rays (CRs) with the interstellar gas and radiation field. In §2, we showed that interstellar H_2 gas is more structured and point-like than current IE models assume, and the related γ-ray emission might be a statistically significant component of the Fermi–LAT data. If this structure is not adequately captured by the IE model, it can impact the identification of resolved point-like sources as well as the characterization of extended components in the γ-ray sky. We demonstrated that unidentified sources in the the fourth Fermi–LAT catalog (Ballet et al. 2020) could indeed be originating from it. In addition, we have argued that this component could artificially inflate the unidentified and/or unresolved point source component in the data and, depending on its morphology, contribute to confounding the interpretation of the the Galactic Center (GC) excess observed by Fermi–LAT (Karwin et al. 2022; Goodenough & Hooper 2009; Hooper & Goodenough 2011; Hooper & Linden 2011; Abazajian 2011; Abazajian & Kaplinghat 2012; Abazajian et al. 2014; Gordon & Macías 2013; Daylan et al. 2016; Calore et al. 2015; Ajello et al. 2016) (see Murgia 2020c, for a thorough review). Improved modeling of small-scale gas related emission in the Fermi–LAT data is therefore necessary to robustly characterize γ-ray sources.

3.1 Tracers of H_2 predicted with deep learning

Interstellar H_2 gas is traced indirectly via the emission lines of other molecules that are found concurrently in gas clouds. Carbon monoxide (^{12}CO, or CO hereafter) is used as a proxy, however, its line emission is typically optically thick in the denser cores of molecular clouds, and therefore it underestimates the H_2 column density there. The line emission of rarer isotopologues of CO also present in the clouds, such as ^{13}CO, remains optically thin to larger column densities and are therefore more reliable to probe dense H_2 cloud cores, and

therefore the H_2 small-scale structure. We briefly summarize the methodology we developed in §2 to model this emission. We employ observations of the $J = 1$–0 transitions of CO and ^{13}CO from the Mopra Southern Galactic Plane CO Survey (Braiding et al. 2018), which cover a 50 square degree region, spanning Galactic longitudes l=300–350° and latitudes $|b| \leq 0.5°$. As described in §2, the data reduction process by the Mopra team yields data cubes of the brightness temperature for each spectral line, as a function of Galactic coordinates and local standard of rest velocity. As described in detail in the rest of this chapter, we employ these data to train machine learning (ML) models to predict ^{13}CO in regions of the sky where observations are not available. To ensure the highest quality data in the rare isotopologues, only pixels for which the brightness temperature exceeds the 1σ noise provided by Mopra are used. The performance of ML models may by biased by noise, and this conservative 1σ noise threshold allows us to retain the majority of the data for training purposes while still rejecting very low quality data.

We calculate H_2 column densities corresponding to Mopra's CO and ^{13}CO, referred to as $N(H_2)_{CO}$ and $N(H_2)_{CO13}$ respectively. We integrate the column density to produce 17 velocity bins, corresponding to increasing radial distances from the Galactic Center. The column density computation and integration procedure is described in §2. The final data is stored as measurements in $1° \times 1° \times 17$ tiles.

With this as input, we build a Modified Map by replacing $N(H_2)_{CO12}$ with $N(H_2)_{CO13}$ for pixels where $N(H_2)_{CO13} > N(H_2)_{CO12}$. The "Modified Excess Template", which corresponds to the small-scale H_2-related γ-ray emission traced by ^{13}CO, is determined from the difference between the Modified Map and the baseline CO map.

The central idea of this paper is to harness ML, in particular deep learning (Baldi 2021), to predict the distribution of ^{13}CO based on the CO observation, and therefore infer the H_2 small-scale structure in regions where ^{13}CO observations do not exist. Since a straightforward and robust analytical mapping between the two distributions is not available, ML

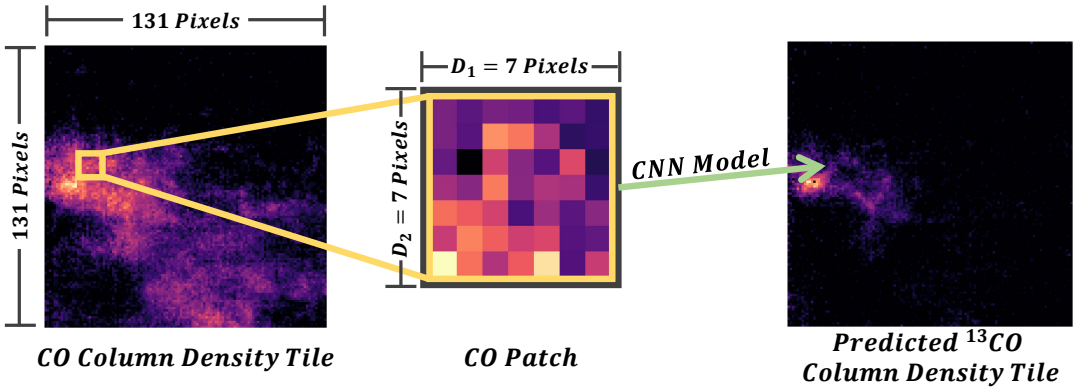


Figure 3.1: A Graphical overview of our CO $\rightarrow{}^{13}$CO modeling process for a single velocity bin. A CO column density tile of size $1^\circ \times 1^\circ = 131 \times 131$ pixels (*left*) is split into into overlapping smaller patches (*center*) of size $0.053^\circ = 7$ pixels. Patches are fed to a CNN model to produce the corresponding predicted ^{13}CO column density tile (*right*) by estimating the ^{13}CO column density at the center of each CO patch.

would estimate this mapping from data. We train a deep learning model to map between CO and ^{13}CO column densities in the Mopra region. To simplify this regression problem and enlarge our effective data set, we subdivide the tiled Mopra data into small $D_1 \times D_2$ patches of the sky to their respective center points (Figure 3.1). The validity of this simplification requires the gas column density to be locally correlated. That is, we assume it is unlikely that column densities outside of our chosen patch will significantly change the ^{13}CO estimate. This simplification greatly reduces the model size and provides us with an effective data set of over 50 million patches when $D_1 = D_2 = 0.0581$, or 7 pixels at the Mopra resolution. We find that larger patch sizes do not improve model accuracy, justifying this locality assumption (see the supplemental material from C. Karwin, A. Broughton, S. Murgia, A. Shmakov, M. Tavakoli, and P. Baldi 2022). We also apply smoothing techniques to eliminate noise from the data (also C. Karwin, A. Broughton, S. Murgia, A. Shmakov, M. Tavakoli, and P. Baldi 2022). The CO$\rightarrow{}^{13}$CO modeling problem may now be formally written as finding a parameterized function f_θ mapping a source patch $S \in \mathbb{R}^{D_1 \times D_2}$ to an estimate target column density at the center of that same patch $T \in \mathbb{R}$. The source S_i corresponds to Mopra CO

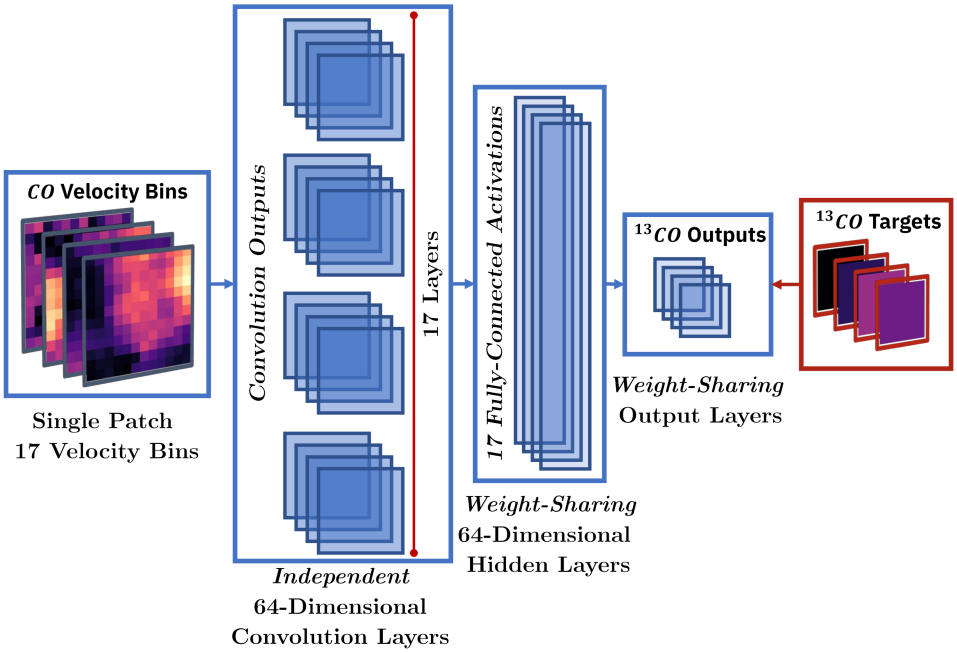


Figure 3.2: A diagram of the internal layers of the network. We maintain independent weights for the 17 convolution layers, one for each velocity bin. However, the hidden and output fully-connected layers share a single set of weights and the same layers are applied to every velocity bin.

column densities and the targets T_i to the corresponding ^{13}CO column densities at the center of the patch (Figure 3.1).

We evaluate our estimates by splitting the Mopra region into independent sub-regions which we designate for training and testing. We employ different splitting choices (see C. Karwin, A. Broughton, S. Murgia, A. Shmakov, M. Tavakoli, and P. Baldi 2022)), but we report on one, the *Alternate Tiled* split, throughout the remainder of this paper. This split consists of alternating $1^\circ \times 1^\circ$ tiles, training on longitudes $\{301-302, \ldots, 349-350\}$ and testing on longitudes $\{300-301, \ldots, 348-349\}$. This allows us to evaluate the model's performance on neighboring data, reducing the variance between training and testing data distributions.

We employ convolutional neural networks (CNNs) to model and predict the ^{13}CO col-

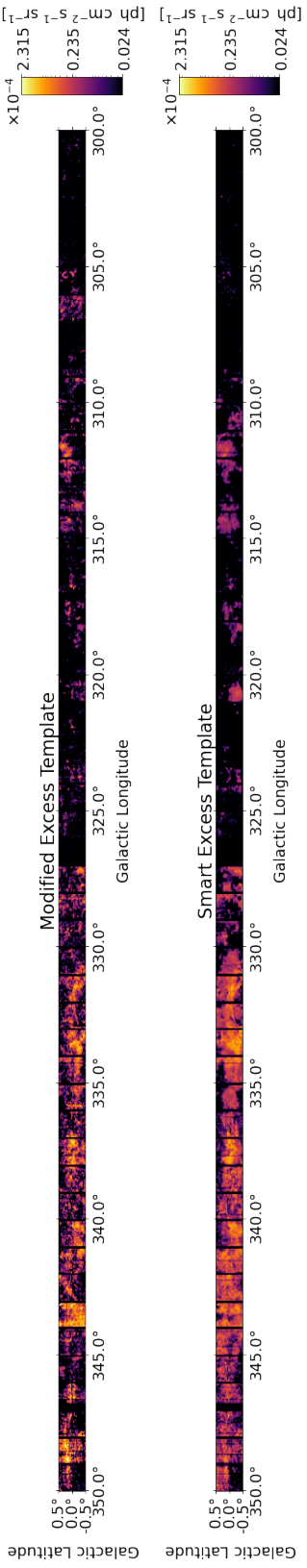


Figure 3.3: *Top*: the Modified Excess Template (from §2). *Bottom*: the Smart Excess Template. The color scale (logarithmic) indicates the γ-ray intensity per 0.03125^2 degree pixel. The first tile to the left (closest to the GC, and covering longitudes $349 - 350$) is testing data, and the neighboring one is training data for the CNN. The interleaving pattern of testing and training tiles is adopted for the entire Mopra region.

umn density in a translation-invariant manner. CNNs use neurons with shared connection parameters in order to implement convolution operations that provide the basis for building translation invariant architectures (Baldi 2021). Each convolution operation is associated with a *kernel*, or filter, corresponding to a set of connection weights that are shared by all the neurons in the corresponding layer. For each of the 17 velocity bins, we apply K learnable convolution filters of size $D_1 \times D_2$, operating on an entire patch and independently for each velocity bin, producing 17 K-dimensional latent vectors for each patch. Afterwards, we apply parametric rectified linear units ($PReLU$ Agostinelli et al. 2014; He et al. 2015; Tavakoli et al. 2021) with a learnable slope, a batch normalization layer (Ioffe & Szegedy 2015), and a random dropout layer (Srivastava et al. 2014; Baldi & Sadowski 2014). The latent vectors are then processed by several fully-connected layers, each with their own $PReLU$, batch-normalization, and dropout. Unlike the convolution layer, these hidden layers are shared between velocity bins, learning identical weights for every bin. This design allows spatial components (CNN) of the network to be specialized for each velocity bin while allowing latent higher level components to be shared between bins. The resulting latent vectors are fed through a final fully-connected layer which produces the ^{13}CO concentration point estimates for each patch and velocity bin. A diagram of the network architecture is presented in Figure 3.2.

The Mopra dataset contains widely varying column densities from bright to very dim. To effectively learn this high-spread distribution, we model the incidence of photons on the Mopra detector as a Poisson process and use a Poisson log-likelihood loss. We further re-weight this loss based on the ^{13}CO density to elevate the importance of bright pixels, prioritizing accuracy in hot-spots over a slight degradation in the background. This increases the importance of accurate measurement in the bright regions, encouraging the network to focus on the accuracy of these regions first. As the network trains, we anneal this weighting back to uniform (with all target values having equal weighting) in order to minimize any bias

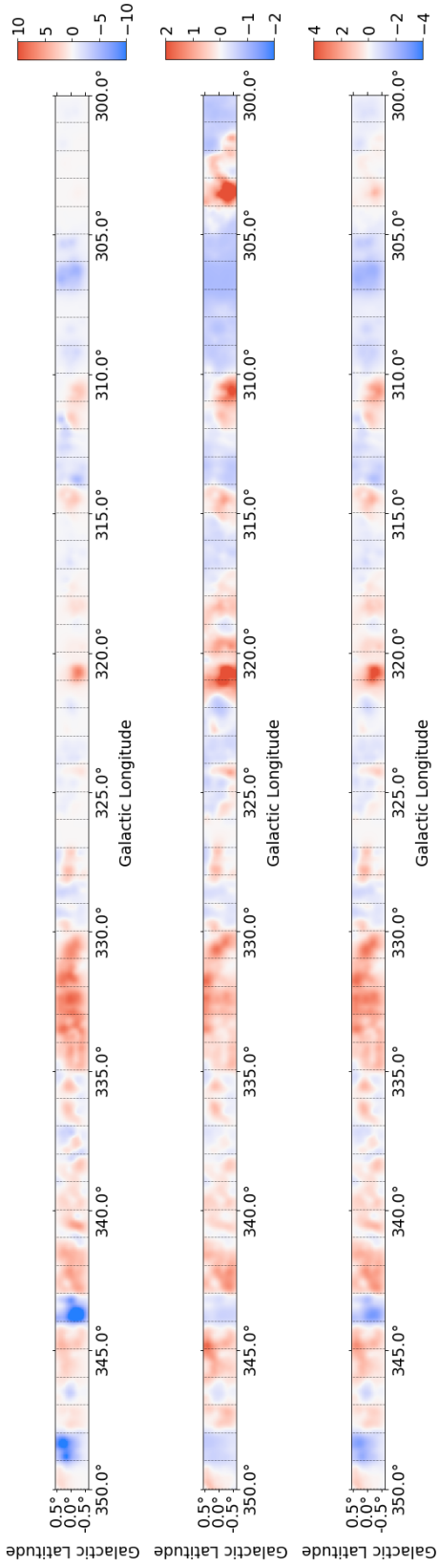


Figure 3.4: The pixel-wise residuals for the full region between the predicted counts from the Smart Excess Template and the predicted counts from the Modified Excess Template. *Top:* residual in counts (Smart-Modified), *Middle:* fractional residuals ((Smart-Modified)/Modified), *Bottom:* Residual in units of 1 standard deviation for the Modified Excess Template ((Smart-Modified)/$\sqrt{\text{Modified}}$). The color scale is per 0.08^2 degree pixel.

introduced by this loss. The design of this loss is guided by the overall goal of finding small angular scale features while limiting the amount of overestimation throughout the Mopra region. A formal description of the loss function is available in C. Karwin, A. Broughton, S. Murgia, A. Shmakov, M. Tavakoli, and P. Baldi (2022).

We tune the CNN's hyperparameters using the Sherpa hyperparameter optimization library (Hertel et al. 2018, 2020), testing 2000 network variations, scoring each parameterization with Poisson likelihood, and using Gaussian Process optimization to suggest parameters. Our final parameterization is presented in C. Karwin, A. Broughton, S. Murgia, A. Shmakov, M. Tavakoli, and P. Baldi (2022), which is evaluated after training the network for 200 iterations. We generate the predicted ^{13}CO density by stitching the network output for every patch in a source image. When performed on a single Nvidia Titan XP GPU, and batching 8192 patches at once, this inference process requires approximately 11 seconds to cover the entire Mopra region.

3.2 Deriving H_2 column densities

Following the same procedure as in §2, we employ the CR propagation code GALPROP (v56)(Moskalenko & Strong 1998, 2000; Strong & Moskalenko 1998; Strong et al. 2000; Ptuskin et al. 2006; Strong et al. 2007; Vladimirov et al. 2011; Jóhannesson et al. 2016; Porter et al. 2017; Johannesson et al. 2018; Génolini et al. 2018) to calculate γ-ray sky maps in 17 Galactocentric radial bins for the H_2-related emission, where the latter is traced by CO and ^{13}CO. The CO and ^{13}CO related H_2 column densities, $N(H_2)_{CO13}$ and $N(H_2)_{CO}$ respectively, are used to determine the H_2 that is missed in dense regions when only CO is used as a tracer. We follow the procedure from §2 to determine the Modified Map and "Modified Excess Template" (Figure 3.3, top panel) to construct their ML analogs. The "Smart Map" is determined similarly to the Modified Map, but instead of using the true

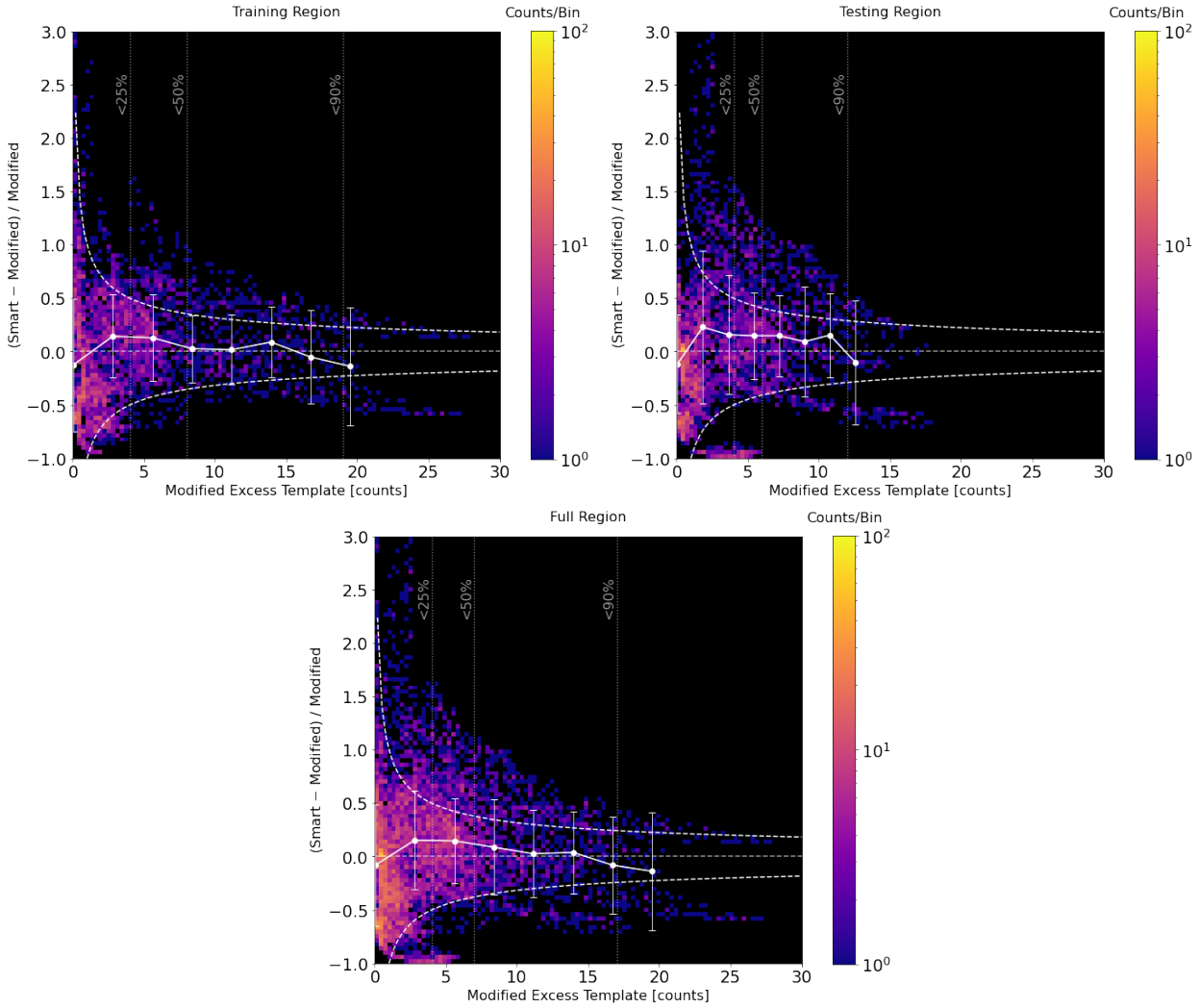


Figure 3.5: The fractional residuals between the nominal Smart Excess Template (*Smart*) and Modified Excess Template (*Modified*) predicted counts per pixel as a function of Modified Excess Template pixel counts for training tiles (*left*), testing tiles (*right*), and all tiles (*bottom*). The points and error bars represent the median and median absolute deviation of the fractional difference for the pixels in 8 domain intervals (we note that the scatter of the points within each interval is not simply characterized by these quantities and does not always indicate a most probable outcome.) We also plot the boundaries of the $\pm 1\sigma$ of the Poisson noise fluctuations of the Modified Excess Template. We overlay the 25%, 50%, 95% flux percentiles to indicate the flux fraction encompassed by pixels below that threshold.

Mopra ^{13}CO data, we use the ^{13}CO estimates from the CNN. We use this to determine the "Smart Excess Template", defined as the difference between the Smart map and the baseline CO map. The Smart Excess Template, also integrated over all annuli and energies, is shown in Figure 3.3, bottom panel. The map covers the full Mopra region, which includes alternating training and testing tiles for the CNN.

To assess how closely the CNN predicts the γ-ray emission inferred by Mopra observations, we compare the excess templates with three different metrics. We quantify the differences with the following three metrics which are determined by using the predicted photon counts per pixel for each of the templates and scaled to $\sim$12 years of Fermi–LAT data (we assume the same Fermi–LAT observation parameters and event selection as for the simulations described in §2 and in this work): the absolute difference (Smart-Modified), the fractional difference ((Smart-Modified)/Modified), and the difference in units of the standard deviation, σ, of the Modified counts ((Smart-Modified)/$\sqrt{\text{Modified}}$). The latter metric allows us to compare the magnitude of the differences due to the CNN performance in light of the statistical power of the data. All results are shown in Figure 3.4 as a function of longitude and latitude for the Mopra region (see C. Karwin, A. Broughton, S. Murgia, A. Shmakov, M. Tavakoli, and P. Baldi 2022, for distributions of residuals integrated over the region). We observe that for the vast majority of the region (83.6% of the pixels), the predicted counts for the Smart Excess Template are within $\pm 1\sigma$ of the Modified Excess Template counts, and therefore the difference is generally within the statistical uncertainty of the data. This result indicates that the CNN performance is adequate for modeling the small-scale structure in H_2-related γ-ray emission traced by Mopra, for the statistics achieved by Fermi–LAT.

The CNN performance varies with respect to longitude. Most tiles display either an overall underprediction by the CNN across the entire tile or an overprediction, rather than a comparable mixture of under/over predictions across the same tile. This may be explained by the design of the loss function, placing increased importance towards brighter pixels while

also biasing the network to prefer underprediction via Poisson regression. Prominent examples include the training tile $343^\circ < l < 344^\circ$ and testing tile $348^\circ < l < 349^\circ$, the brightest in each dataset, where the CNN underpredicts the gas column density, as shown in Figure 3.4. The dependence of the residuals as a function of pixel brightness (in counts/pixel), is shown in Figure 3.5 for the full Mopra region, and for the training and testing sub-regions separately. Generally, better agreement between Mopra (Modified Excess Template) and the CNN prediction (Smart Excess Template) is found for the brighter pixels. There is a broad spread in the fractional residuals, but it is confined to be between $\pm 50\%$ for the vast majority of the pixels. We overlay contours to indicate the $\pm 1\sigma$ statistical fluctuations in the Modified Excess Template, which encompasses 83.6% of the pixels. We also overlay the median and median absolute deviation of the fractional difference in bins of the Modified Excess Template counts to guide the eye, but emphasize that the scatter of the points within each interval cannot be simply characterized by these quantities and does not always indicate a most probable outcome. As expected, we find that the CNN performs (somewhat) better in the training tiles. Approximately 58% of pixels where the CNN prediction is beyond the $\pm 1\sigma$ level are in the testing sub-region. The CNN is more likely to overpredict the emission for the dimmer pixels (60.8% of pixels across the full region are below the 25% flux percentile), consistently with the design. The CNN overprediction in the dim pixels essentially spreads out the brightness of hot-spots over a larger area. We also note that small statistics causes the fractional difference to increase dramatically for the faintest pixels. Above the 90% flux percentile, the distribution of the fractional residuals bifurcates into two separate distribution. The underpredicted pixels are in the $343^\circ < l < 344^\circ$ tile, and the overpredicted pixels are in $333^\circ < l < 334^\circ$ tile, both in the training sub-region. This is likely caused by our choice to combine independent convolution layers for each velocity bin with a shared hidden layer. Since certain bins have higher overall brightness than others, the shared hidden layer will attempt to average the error between the two extremes.

3.3 Likelihood of excess γ-ray emission

We determine the significance of the Smart Excess Template in the Fermi–LAT data, similarly to §2 for the Modified Excess Template. The simulations cover the same observations and event selection (∼12 years, $1 - 100$ GeV, P8R3 CLEAN FRONT+BACK). We only simulate the H_2-related γ-ray emission in the Mopra region, excluding all other components, since our goal is to establish the performance of the Smart Excess Template in the optimistic scenario where all other components are known. The simulated events trace the H_2-related γ-ray emission modeled with the Modified Map from Mopra. The simulated data are fit based on a binned maximum likelihood method to a model that includes the baseline CO map (as observed by Mopra) and the Smart Excess Template. The normalization of the Smart Excess Template is free to vary in the fit. The energy spectrum, calculated by GALPROP, is assumed, and held fixed during the likelihood fit. The normalization and spectral index of the CO baseline contribution is also free to vary. The 17 radial bins from GALPROP are combined into 4 radial annuli, which we refer to as A1, A2, A3, and A4, with the same partitioning as in §2. Just as in §2, the normalization of the dim A4 bin is held constant to a normalization of 1.0, contributing a partial flux of 4.98×10^{-9} ph cm^{-2} s^{-1}.

We simulate 1000 realizations and calculate the Test Statistics (TS) for the nested models: $-2\log(L_0/L)$, where L_0 is the null hypothesis (CO baseline) and L is the alternative hypothesis (CO baseline and Smart Excess Template). The statistical significance is approximated by $\sigma \approx \sqrt{\mathrm{TS}}$. The distribution of the $\sqrt{TS}$ for the simulations is shown in Figure 3.6 for the full Mopra region (solid line), and for the testing sub-region only (dashed line). The Smart Excess Template corresponds to $\sqrt{TS} = 27.5 \pm 1.1$ (mean and standard deviation) in the full Mopra region and 12.9 ± 1.0 in the testing sub-region. The significance in the testing sub-region is lower, not only because of the smaller statistics, but also because the full region contains the training tiles where the CNN more closely matches the

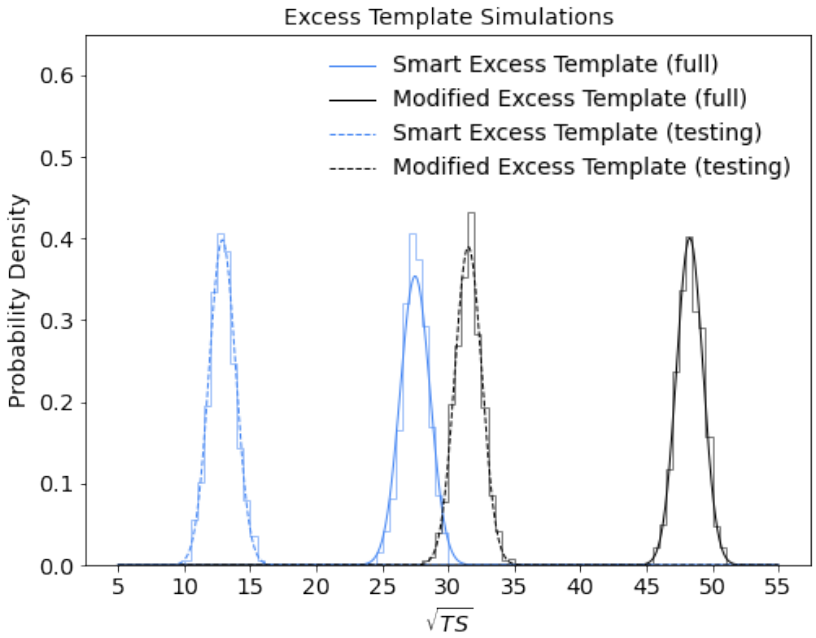


Figure 3.6: Statistical significance of the Smart Excess Template for the full Mopra region (solid lines), and for the testing sub-region (dashed lines). Each contains 1000 realizations of 12 years of Fermi–LAT data. A fit with a Gaussian distribution is overlaid to each distribution. These distributions are compared to the results for the Modified Excess Template from §2.

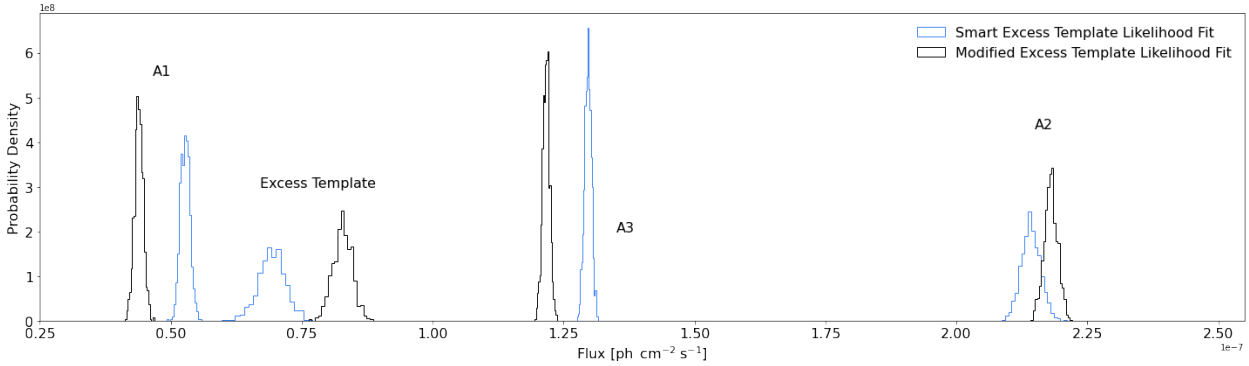


Figure 3.7: Distribution of fluxes for each component of the best-fit model using the Smart Excess Template maps over the full region for 1000 simulations of Fermi–LAT data. The results for the Modified Excess Template maps from §2 are overlaid. Note that one of the annuli (A4), not shown, has negligible contribution and the normalization was fixed in the fitting process.

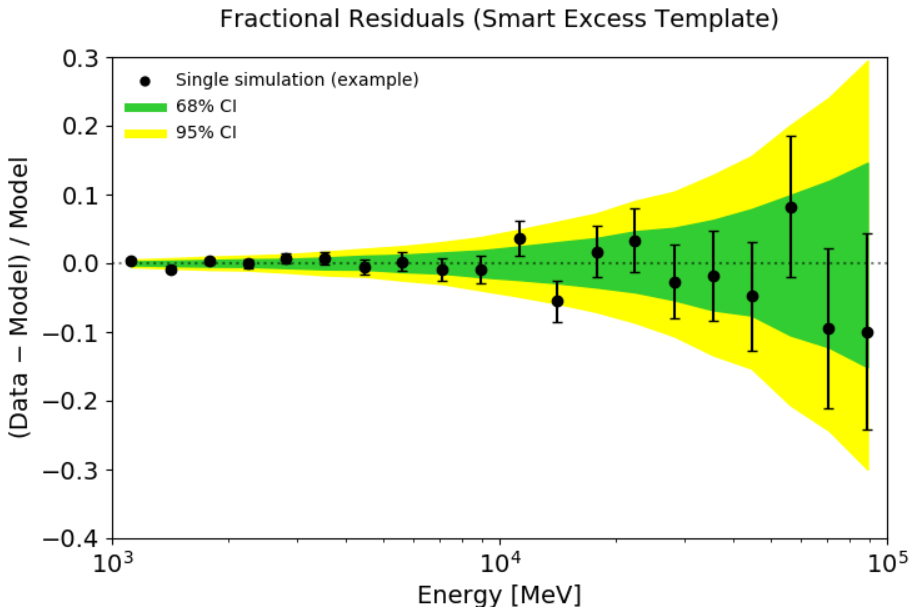


Figure 3.8: Fractional count residuals for the fit with the Smart Excess Template. The green and yellow bands show the 68% and 95% confidence regions from 1000 simulations, respectively. We also plot the results for a single simulation as an example in each case, which is shown with black data points.

Mopra data. For comparison, the Modified Excess Template has a $\sqrt{TS} = 48.3 \pm 1.0$ and $\sqrt{TS} = 31.5 \pm 1.0$ for the full region and testing sub-region, respectively. In Figure 3.7, we show the distributions for the best-fit flux of the Smart Excess Template overlaid with those for the Modified Excess Template, as well as the flux for the CO baseline emission, separated into annuli. The integrated flux for the Smart Excess Template is approximately 83.7% of the flux of the Modified Excess Template, indicating an overall underprediction by the CNN. The fractional count residuals ((Data-Model)/Model) as a function of energy for the best fit model are shown in Figure 3.8 for the fits that include Smart Excess Template, in addition to the other CO components. They are consistent with zero. The best fit spectra of the CO components agree with the GALPROP prediction. Finally, Figure 3.9 shows the fractional count residuals in latitude, longitude. The residuals, which incorporate differences between the (best fit) Smart Excess Template and the (simulated) Modified Excess Template, are

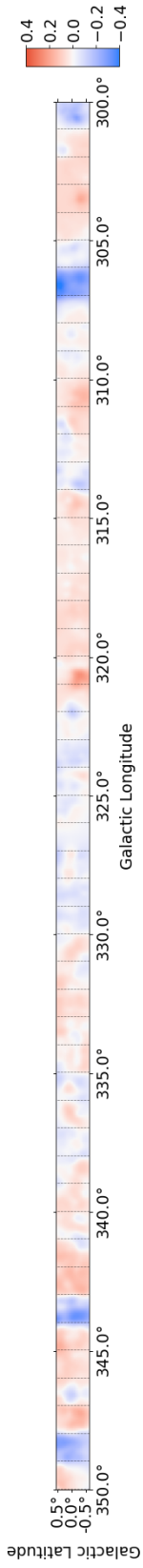


Figure 3.9: Fractional count residuals, (Data-Model)/Model, in latitude and longitude for the fits with the Smart Excess Template, in addition to the other CO components. The residuals are calculated using the mean over 1000 simulations.

smaller than those shown in Figure 3.4, (middle panel) indicating the other CO model components partially compensate for the discrepancies between the Excess Templates. We have performed this analysis also using Gaussian Processes and find that the performance is worse compared to the CNN (C. Karwin, A. Broughton, S. Murgia, A. Shmakov, M. Tavakoli, and P. Baldi 2022).

3.4 Discussion

The ML methodology presented here must be refined to be extended to other regions of the sky. The available all sky CO J=1–0 observations have significantly poorer spatial resolution ($8'$ from Dame et al. (2001)) compared to Mopra ($0.6'$). The CNN model will therefore require additional transfer learning to operate on lower resolution maps. The resolution of the Fermi–LAT data is worse (for most energies and event types), suggesting that the impact of the poorer CO resolution on the CNN predictions for the γ-rays might ultimately be less pronounced. Another consideration is that the available ^{13}CO training data for the CNN from Mopra and other observations (Urquhart et al. 2018; Eden et al. 2020; Kalberla et al. 2020; Blackwell et al. 2017; Di Teodoro et al. 2018), are confined to the Galactic plane. The H_2 scale height is low, but its small-scale structure contribution at higher latitudes might not be negligible because of the contribution of more local H_2. However, because of the lack of adequate training data the CNN prediction for this component could be more uncertain and this is relevant for the characterization of the GC excess, which extends to high latitudes. Finally, this analysis inherits the limitations and uncertainties in modeling the H_2 component with traditional methods, including the loss of kinematic resolution for the gas in the direction of the GC (Ackermann et al. 2012). And it adds another: the H_2 column densities from CO and ^{13}CO have been treated independently in this work. Their estimates can be related analytically, but rely on a assumptions pertaining to the optical

depth, beam filling factor, spatial variation, etc. (e.g. Burton et al. 2013; Cormier et al. 2018). More ambitiously, ML may be used to constrain some of these uncertainties by using the γ-ray data. Finally, small-scale structure in the γ-ray data could also arise from other components of the IE, e.g. related to HI, and shall be included in a more comprehensive study.

3.5 Conclusions

We have presented a methodology that harnesses ML to predict the small-scale component of the interstellar H_2 gas and its related γ-ray emission for the first time. Our work is motivated by the the fact that while CO is widely employed as a tracer of H_2, its emission alone does not reliably capture the more highly structured component of H_2. Incorporating this small-scale component in γ-ray IE models is crucial for determining γ-ray point sources and could affect the characterization of extended γ-ray sources, e.g. the Fermi–LAT GC excess. Here, we have improved H_2 modeling by employing observations from the Mopra survey which includes isotopologues of CO such as ^{13}CO that do not suffer from the same limitations and can, therefore, more reliably trace the H_2 small-scale structure (the survey spans Galactic longitudes l=300–350° and latitudes $|b| \leq 0.5°$.) In particular, we designed and trained a CNN to predict the concentration of ^{13}CO in the Mopra region.

We have tested the performance of this methodology in predicting the contribution of the H_2 related γ-ray small-scale structure to the Fermi–LAT data. We find that the accuracy of the predicted γ-ray emission, which incorporates the CNN-predicted contribution, is within the statistical uncertainty of the Fermi–LAT data for the vast majority of the region. Moreover, the predicted γ-ray contribution can have a large statistical significance, $\sim \sqrt{TS} = 28$ if we assume other components of the γ-ray IE are known. We designed a dynamically-weighted loss function for the CNN's training to prioritize emission accuracy

from brighter spots in the Mopra data while avoiding over-predicting the ^{13}CO concentration in these regions. This is a conservative choice aimed at minimizing the injection of spurious γ-ray sources in the IE model. As a consequence of this design, over-prediction is more likely for the dimmer spots, which we show to be less important. Based on these results, we conclude that deep learning can improve the modeling of γ-ray emissions in data-rich regions, supporting prospects to extend this methodology to other regions of the sky. In this work, we restrict our analysis to the Mopra observations of CO and ^{13}CO which are confined to the Galactic plane. Extending this methodology to other regions will require the CNN model to describe ^{13}CO from the poorer resolution CO maps available for the majority of the sky, and will require extrapolating predictions to higher latitudes.

Part II

The Legacy Survey of Space and Time

Chapter 4

Introduction

Consider the path of light, from the event it was created until the event it arrives at our telescope at our single vantage point in spacetime. This light, generated by some physical process, will propagate along a path defined by the geometry of the Universe and its interactions along the way. It's path may deviate from a straight line by the gravitational effects of nearby bodies. It also may be attenuated, boosted, polarized, or absorbed and re-emitted all by different materials it encounters along its worldline. When it arrives at Earth, we can only measure it's brightness, energy, directionality, and depending on one's detector, it's polarization. It's properties encode the sub-processes, contents, and evolution of the Universe since it was emitted. Because light travels at a large but finite velocity, the further away the light begins its journey, the earlier into history it can record. Our best cosmological theory has has developed from this idea by inferring from observational evidence obtained using advanced instrumentation and analytical techniques.

Three key observations made in the past century led to our current understanding of the Universe: (i) galaxies are all moving away from each other and accelerating despite the gravitational forces between them (Hubble 1929; Perlmutter et al. 1999); (ii) the current relative abundances of the lightest elements (e.g. [H/He]) cannot be explained by the nuclear

processes within stars, but can be explained by also considering nuclear reactions in an early hot universe; (iii) detection of emission from a nearly perfect blackbody spectrum at 2.7K with only small fluctuations ($\delta T < 10^{-5}$), which is predicted to be a colder relic from the moment the universe became cold enough to form neutral atoms. These observations suggest that the Universe was once very dense, very hot, and very homogeneous (same everywhere) and isotropic (same in all directions). To explain the expansion of the Universe, we have needed to propose the existence of a source of gravitational repulsion called dark energy. And in order to counteract the expansion enough to form large structures within this accelerating Universe, we have been forced to also propose the existence of an extra source of gravitational pull called dark matter. However, the true nature of these components unknown. Different versions of dark energy and dark matter fit into different inflationary models of the Universe. And differences between the best-fit inflationary models from early- and late- universe probes are already hinting at a potential paradigm shift in our understanding.

However, cosmological surveys have not yet achieved the necessary sub-1% level precision needed to distinguish between different inflationary models. One of reasons for this is inadequate correction of instrument-induced distortions. The model constraining power of future cosmological surveys will not greatly improve unless potential systematic errors in the instrumentation can be adequately treated. This thesis details my efforts to meet these targets in future cosmological surveys.

4.1 Modern Cosmology

Consider again the path of light as it travels through the Universe. Einstein's theory of general relativity describes how a distribution of matter and energy can distort the coordinate system of the universe in which our photon travels. In a universe that is homogeneous and isotropic, then the distance between any two events separated by space and time is governed

by Friedmann-Lemaître-Robertson-Walker (FLRW) metric:

$$ds^2 = -dt^2 + a^2(t)d\Sigma^2 \tag{4.1}$$

where in spherical coordinates $d\Sigma = \left[\frac{dr^2}{1-\kappa r^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2)\right]$ describes the geometry of the Universe, and we are following the convention of natural units ($\hbar = c = k_b = 1$).

This equation is a solution of Einstein's field equations and represents a generalized topological framework that maps a small step in coordinates in space and time (the path of our light) to a small step in a reference frame independent of the observer (us). The parameter κ is derived from Einstein's field equations, and describes the curvature of the coordinate system by the distribution of matter and energy inside of it. For a universe that contains energy distributed in a homogeneous and isotropic way, κ is allowed to be "flat" ($\kappa = 0$, parallel lines will never meet), positively curved ($\kappa = +1$, parallel lines will converge), or negatively curved ($\kappa = -1$, parallel lines will diverge).

The evolution of the geometry of the Universe is captured by the scale parameter $a(t)$, which is defined so that the size of the Universe today is 1. In the thermodynamic view of the Universe, the Universe expands adiabatically (i.e. no energy enters or leaves the Universe), so its evolution is just determined by the equation of state $w_X = P_X/\rho_X$ (the ratio of the partial pressure to the energy density) of each of the contents within it. This is given by the Friedmann equation with the set of cosmological parameters $\Theta = \{H_0, \Omega_M, \Omega_R, \Omega_{DE}, w\}$:

$$H(t) = \frac{\dot{a}}{a} = H_0 \left(\Omega_M a^{-3} + \Omega_R a^{-4} + \Omega_{DE} a^{-3(1+w)} + \Omega_k a^2\right)^{1/2}. \tag{4.2}$$

The ratio of the expansion rate of the Universe to its current size is defined as the Hubble parameter H (where H_0 refers to the value of the constant today). This parameter depends on the contents of the Universe, $\Omega_X = \rho_X/\rho_c$, which is the ratio of the density of component

X relative to the critical density $\rho_c = 3H_0^2/8\pi G$ (the critical density of matter and energy that would make a homogeneous and isotropic universe flat, $\Omega_{\rm TOT} = \sum_X \Omega_X = 1$). Here we consider just the contributions from:

1. Matter (Ω_M): This includes contributions from all matter and could include cold (non-relativistic) dark matter (e.g WIMP scenario) since both have the same equation of state. It's energy density scales inversely-proportional to the volume of the universe ($\propto a^3$).

2. Radiation (Ω_R): This includes contributions from massless particles like light (Ω_γ) as well as light-like (relativistic) particles such as neutrinos (Ω_ν), which have very low mass, or potentially hot (relativistic) dark matter. It's energy density scales inversely-proportional to volume of the universe and its own wavelength ($\propto a^4$).

3. Dark energy (Ω_{DE}): This is a hypothetical component with a generic equation of state w that would cause the universe to accelerate if $w < -1/3$.

4. The curvature of the universe (Ω_k): This includes the contribution from the energy density of the curvature of the universe itself. It is defined by the deviation from critical density from the combination of all other components $\Omega_k = 1 - \Omega_{\rm TOT}$.

Although, numerous versions of this theory exist (see the review by Huterer 2023), and one could easily remove components from the theory or extend it to include contributions from any theoretically abstract form of matter and energy as long as an equation of state can be defined.

The ΛCDM ("Lambda Cold Dark Matter") universe is a special case of this theoretical framework. This universe is described with contributions from matter and cold dark matter, radiation, dark energy, and curvature. The dark energy in this special model has an equation of state of $w = -1$, making dark energy a "cosmological constant" defined by $\Lambda = 3H_0^2\Omega_{DE}$.

Another special case, w_0w_aCDM, is similar except dark energy is allowed to evolve over time by allowing the dark energy equation of state to depend on the scale parameter so that $w = w_0 + w_a(1 - a)$.

A consequence of the expansion of the Universe is that the light is stretched to longer wavelengths (redshifted) with the expanding spatial coordinates. The ratio between the wavelength of the light when it was emitted (λ_e) and the wavelength of the light when it was observed (λ_o) is directly related to the ratio of the scales of the Universe at those times such that $z = \lambda_o/\lambda_e = a(t_o)/a(t_e) - 1$. The amount by which light is spectroscopically shifted depends on the distance it has traveled in a changing universe. Redshift is often a useful proxy for the distance that the light has traveled since it is a directly observable property of light. Certain processes within different elements and molecules emit light at known wavelengths, and we can measure the redshift directly by comparing the observed wavelengths to the known emitted wavelengths. It is therefore often useful to write out the theory with this parameter in place of a distance parameter.

Note that when we refer to the distance that a photon travels, there are actually three different ways to measure distances within this description of the Universe: the comoving distance, the luminosity distance, and the angular-diameter distance. The comoving distance evolves with the expanding universe (the so-called "Hubble flow") giving a distance measurement that does not change with time. The comoving distance (r) at a given moment in time is related to the physical distance we would measure if we could pause time and measure the distance between two places with ruler (the proper distance d_p) by the scale factor of the universe:

$$r(t) = d_p(t)/a(t) \tag{4.3}$$

Given our formalism for an evolving universe, the comoving distance over which a photon travels from the moment it was emitted until we observe it in the present epoch is related

to the content of the universe:

$$r(z) = \int \frac{dt}{a(t)} = -\int_z^0 \frac{dz'}{H(z')} = \frac{H_0^{-1}}{\sqrt{\Omega_k}} \sin\left[\sqrt{\Omega_k}\int_0^z \frac{dz'}{E(z')}\right] \tag{4.4}$$

where

$$E(z') = \left[\Omega_M(1+z')^3 + \Omega_R(1+z')^4 + \Omega_{DE}(1+z')^{3(1+w)} + \Omega_k(1+z')^2\right]^{1/2}. \tag{4.5}$$

The comoving distance is invariant with a changing scale factor, which is why we never directly measure this distance. Note that hereafter, we will specifically denote the comoving distance for a flat universe with $\chi(z)$.

More realistically, we might infer the luminosity distance (d_L) based on a comparison of the apparent flux of light of an object with a known intrinsic luminosity. If we have two objects of the same luminosity and one is further away from us, the one that is further away will look dimmer. We know how the flux of light reaching us decreases with distance:

$$f = \frac{L}{4\pi d_L^2} \tag{4.6}$$

where L is the known intrinsic luminosity of the object which gets spread out over a sphere of area $4\pi d_L^2$ resulting in us observing a flux f. The radius of the sphere over which the light is spread out is the comoving distance. The light will be spectroscopically shifted by an expanding universe by a factor of $(1+z)$, and it will also be received by the observer at a rate scaled by $(1+z)$. Therefore, the luminosity distance is related to the comoving distance by:

$$d_L(z) = (1+z)r(z) \tag{4.7}$$

Another way to measure distance is by inferring the angular diameter distance based on

a comparison of an object's angular size on the sky ($\Delta\theta$) with its known proper length (Δl). The arc length is related to the proper length by $\Delta l = a(t) r \Delta\theta$. This ratio is related to the comoving distance between the observer and the object:

$$d_A(z) = \frac{1}{1+z} r(z) \tag{4.8}$$

In this theoretical framework, we have only made two assumptions: (i) Einstein's theory of general relativity and (ii) the homogeneity and isotropy of the Universe. From this we have shown that the distribution and relative content of different constituent forms of matter and energy on the scales of the Universe deterministically constrain the the evolution of the Universe. It is apparent also that an evolving universe affects the brightness, directionality, and color of light observed from distant objects. Perhaps by some lucky accident these properties also happen to be observables we can measure from our own single vantage point on Earth.

4.2 Measurements of the Late Universe

The theory described in the previous section describes how the Universe evolves based on the distribution and relative content of matter and energy. Fundamental to the theory of general relativity, from which this framework is developed, is the idea that "matter tells spacetime how to curve, and curved spacetime tells matter how to move" as the eminent physicist John Archibald Wheeler once said. If we map all the matter we can find in the Universe, we observe that the Universe is generally homogeneous and isotropic on scales above 100 Mpc, but below this level there are structures formed by this matter (e.g. stars, globular clusters, galaxies, and galaxy clusters) which are inhomogeneities which formed from their interactions through gravity and other short-range physical processes (LSST Science Collaboration et al. 2009;

Planck Collaboration et al. 2020a; Huterer 2023). These structures form and move within the evolving Universe and trace its evolution through time. In this thesis, we primarily focus on just a couple of these cosmological probes that utilize this principle.

4.2.1 Type Ia supernovae

The most direct measurement of the evolution of the Universe through time would be to relate the distance to galaxies with redshift as they sit within the Hubble flow. Galaxies will of course be Doppler shifted by their kinematic motion under gravity, but they will also be redshifted as they move apart by simply sitting within the expanding Universe. Observing many galaxies will average out the dispersion of redshifts due to their gravitational kinematic motion, and leave us with just the redshifts resulting from the expansion of the Universe. And if we knew the exact distance to these galaxies, we would know exactly how long the light has been traveling to get to us. If could know both the distance and redshift to distant galaxies, we could construct a picture of the expansion of the Universe as a function of lookback time.

One way to do this is with a standard candle. To illustrate this, we can imagine two candles with the same luminosity. If we were to move one candle further away, we could easily determine the relative distance between them by how much the apparent flux has changed. If we had good knowledge of the distance to the first candle, we could confidently measure the distance to the second candle. Classical Cepheid variable stars are known standard candles; the periodicity of their brightness is correlated with their peak brightness, and they are typically close enough that we can use parallax angles to independently measure their distance from us (see the discovery of Cepheid variables in the Small Megellanic Cloud by Leavitt & Pickering 1912).

Type Ia supernovae (SN Ia) are a homogeneous class of standard candles that can be seen

over much greater distances (Branch & Tammann 1992). They are extremely bright explosions that occur when a white dwarf star, accreting matter from an unfortunate companion star, hits the Chandrasekhar mass limit. These explosions are very bright and short-lived. The few examples we have found (approximately 7,000 to date Brout et al. 2022; Rubin et al. 2023) are commonly bright enough to outshine their host galaxies and can be seen over many gigaparsecs away. They typically fade after about a month, making them tricky to catch not observing the right place at the right time. However, the SN Ia sub-class has a few properties that make it very useful standard candle. For one, SN Ia have approximately the same maximum total luminosity.[1] Secondly, the SN Ia sub-class is characterized by an easily recognized emission line from silicon (Si II) around 2eV. And thirdly, some SN Ia occur in galaxies with observable Cepheid variable stars, which allows us to thoroughly calibrate and test our empirical relations with other independent measurements of distance. Combining observations of parallax, Cepheids, and SN Ia, we can construct a *cosmic distance ladder* to map redshift as a function of distance.

The observed brightness is measured by the apparent magnitude m, which a logarithmic measurement of flux compared to a well-calibrated reference source, and the the true brightness is measured by the absolute magnitude M, which is a similarly reference-calibrated logarithmic measurement of its intrinsic luminosity. The difference between these magnitudes is the *distance modulus*, which is simply related to the luminosity distance:

$$\mu \equiv m - M = 2.5 \log_{10}(\frac{d_L}{10\text{pc}}) - K_{\text{corr}}, \tag{4.9}$$

[1] In general SN Ia are not intrinsically standard candles; their peak fluxes can vary depending on extinction from the environment (e.g. the interstellar medium or ISM) within their host galaxy, for example. However, the peak flux of a SN Ia are correlated with other properties of their light curve and the properties of their host galaxies (e.g. how long it takes to reach its peak or fall back down and its color) (Guy et al. 2007; Kenworthy et al. 2021). The diversity of different supernovae and the covariance of different SN Ia properties has been intensely studied, and many treatments exist. Many studies will include additional terms to distance modulus (Equation 4.9) that act as corrections to the absolute magnitude (also refer to Kenworthy et al. 2021).

where $K_{\rm corr}$ is a correction to account for the fact that we measure the apparent magnitude in a single physical filter while the supernova emits a broader spectral energy distribution (SED) in its own rest frame. These observables are clearly sensitive to the underlying cosmological model of the Universe. The distance to the supernova tells us how long the light from the explosion has been traveling, and the redshift of the light measures the relative size of the Universe from the exact moment of the supernova explosion.

In practice, we try to find the set of cosmological parameters that minimizes the difference (called the "Hubble residual"), between the observed distance modulus with the distance modulus predicted by our model of the Universe, which for a single galaxy (index i) is given by:

$$HR_i = \mu_{obs,i(z_i)} - \mu_{\rm MODEL}(z_i, \Theta). \tag{4.10}$$

Compiling searches over the past couple of decades have confidently turned up thousands of SN Ia in the late Universe, around $z = 0.001 - 2.26$ (Brout et al. 2022; Rubin et al. 2023), mainly split between low redshift samples (e.g. Riess et al. 1998; Kowalski et al. 2008) and high redshift samples (Perlmutter et al. 1999; DES Collaboration et al. 2024, e.g.). The largest and most recent self-consistent analysis of SN Ia comes from the Dark Energy Survey (DES) Supernova Program (DES Collaboration et al. 2024). Their observations were consistent with a ΛCDM Universe that is flat with $\Omega_k = 0.16 \pm 0.16$ and dominated by matter and dark energy $(\Omega_M, \Omega_\Lambda) = (0.291^{+0.063}_{-0.065}, 0.55 \pm 0.17)$ with a Hubble constant of $H_0 = 73$ km s^{-1} Mpc^{-1}, although these data are also consistent, within statistical uncertainties, with dynamical models ($w_0 w_a CDM$) that allow dark energy equation of state to change with time. Even the most recent and complete SN Ia analyses continue to be dominated by statistical uncertainties rather than systematic uncertainties (e.g. DES Collaboration et al. 2024), and more data will be needed in the future to increase our confidence to distinguish in these models.

4.2.2 Weak gravitational lensing

Another way to measure the the scale of the Universe through time is by tracing the size and distribution of the different structures in the Universe at different redshifts. The expansion of the Universe under the influence of dark energy directly opposes the growth of structure under the influence of gravity. The statistical distribution of different structures can therefore depend on the relative contents of the Universe and can constrain our cosmological models.

The simplest way to measure these properties (besides straightforward counting of galaxies/clusters of galaxies as a function of apparent magnitude per unit area, $dN \log N(m)/dm$, of course) is with the two-point correlation function (2PCF). If there is an object at position $\vec{x}$, then the excess probability of finding another object at position $\vec{x}' = \vec{x} + \vec{r}$ is the correlation function $\xi(r)$. Note that in an isotropic Universe, we can simply replace $\vec{r}$ with the scalar distance r. For any matter density fluctuation $\delta(\vec{\theta}) = (\rho(\vec{x}) - \bar{\rho})/\bar{\rho}$, where $\rho(\vec{x})$ is the density at $\vec{x}$ and $\bar{\rho}$ is the mean density, the two-point correlation function is:

$$\xi(r) = \frac{\langle [\rho(\vec{x}) - \bar{\rho}][\rho(\vec{x}') - \bar{\rho}] \rangle}{\bar{\rho}^2} \equiv \langle \delta(\vec{x}) \delta(\vec{x}') \rangle \tag{4.11}$$

where $\langle .. \rangle$ refers to an average over all positions $\vec{x}$. Perhaps more usefully, it is common to use the Fourier transform of the two-point correlation function, which is called the matter power spectrum (P):

$$\tilde{\xi}(k) = \langle \tilde{\delta}(\vec{k}) \tilde{\delta}^*(\vec{k}') \rangle = (2\pi)^3 \delta_D^{(3)}(\vec{k} - \vec{k}') P(\vec{k}) \tag{4.12}$$

where $\tilde{\delta}_m(\vec{k})$ is the Fourier transform of the matter density fluctuations $\delta_m(\vec{\theta})$ and $\delta_D^{(3)}(\vec{k} - \vec{k}')$ is the 3D Dirac delta function. Note that in an isotropic Universe, we can simply replace $\vec{k}$ with the scalar wavenumber k.

The structure of the Universe was seeded by small primordial density fluctuations that

grew under the influence of gravity as the Universe evolved. Structure formation requires complicated multi-dimensional models of all the physical sub-processes involved (Krause & Hirata 2010), however the problem can be simplified by only considering density fluctuations within the cosmic horizon which can causally interact to form structures and late times (so that the speed of sound in the medium is negligible), in which case the fluctuations grow linearly under the influence of gravity in the expanding Universe:

$$\ddot{\delta} + 2H\dot{\delta} - 4\pi G\rho_m(t)\delta = 0. \tag{4.13}$$

This equation is generally true for small density fluctuations ($|\delta| << 1$) in the early Universe, which is largely consistent with observations of relic radiation from the surface of last scattering (highly Gaussian fluctuations of rms $\sim 10^{-5}$ with a blackbody spectrum cold enough form neutral hydrogen at redshift $z \sim 1100$) (Planck Collaboration et al. 2020b). These densities grow independently of wavenumber, so we only write the density fluctuations as a function of time here. We call the large scales on which this equation accurately approximates the growth of structure as the "linear" scales and correspondingly refer to the "linear part" of the matter power spectrum. We also refer to small scales as the "non-linear part" of the matter power spectrum, where other physical processes complicate structure formation and evolution and require more specialized simulations/modeling to predict from theory (Krause & Hirata 2010). The largest (sub-horizon) objects are galaxy clusters, and they make up the so-called *large-scale structure* (LSS) of the Universe, although the form of these structures are still affected by non-gravitational processes and are actually part of the non-linear regime.

The over-densities of matter in the LSS distort the topology of spacetime around them, and light follows along a path defined by the distorted space. Image distortions of galaxies are therefore sensitive to the statistical properties of this structure through different cosmological

lookback times, and can be used to measure the evolution of the Universe as a whole. Gravitational lensing of galaxies by the LSS produce only a small distortion or "shear" to the shape of a galaxy (typically 0.01–1%). We call this weak gravitational lensing (WL).

For sufficiently small distortions and where the size of the galaxy is much smaller than the angular size of the lensing structure, we can approximate the effect of WL as a linear mapping from an unlensed "source" plane to a lensed "observed" plane. We will then describe the observed light from a galaxy as a distribution of light coming from a direction $\vec{\theta}$, which was distorted by a linear transformation from a true direction $\vec{\beta}$. We can write this formally as:

$$f^{obs}(\vec{\theta}) = f^{s}(\mathcal{A}\vec{\theta}), \tag{4.14}$$

where the approximate distortion can be described by a distortion matrix $\mathcal{A}$, which is the Jacobian of the transformation from the unlensed coordinate system to the lensed coordinate system. In a simple Born approximation, the change in the observed direction of the light is proportional to the integral of the 3D gravitational potential along the line-of-sight. We can write down the 2D lensing potential in the lensed coordinate system as the integral of the 3D gravitational potential along the line-of-sight to the source at a comoving distance χ_s:

$$\psi(\vec{\theta}) = \frac{2}{c^2}\int_0^{\chi_s} d\chi \frac{\chi_s - \chi}{\chi_s \chi}\Phi(\chi\vec{\theta}, \chi). \tag{4.15}$$

Where the gravitational potential is given by the Poisson equation around a small density contrast, which we can write in comoving coordinates:

$$\nabla^2\Phi(\chi\vec{\theta}, \chi) = \frac{3}{2}\Omega_M H_0^2 a^{-1}(\chi)\delta_m(\chi\vec{\theta}, \chi). \tag{4.16}$$

Since there is a physical symmetry here, we can decompose the distortion matrix into a diagonal term involving the *convergence* (κ) and a symmetric, traceless term involving the

shear (γ):

$$\mathcal{A}_{ij} = \frac{\partial \beta_i}{\partial \theta_j} = (\delta_{ij} - \frac{\partial^2 \psi(\vec{\theta})}{\partial \theta_i \partial \theta_j}) = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa-\gamma_1. \end{pmatrix}. \tag{4.17}$$

The shear component distorts the shape of the galaxy and a combination of the shear and convergence changes the size and brightness ("magnification") of the galaxy. These components are derived from the gravitational potential:

$$\begin{aligned} \kappa &= \frac{1}{2}(\partial_1\partial_1 + \partial_2\partial_2)\psi(\vec{\theta}) \\ &= \frac{1}{2}\nabla^2\psi(\vec{\theta}) = \frac{1}{c^2}\int_0^{\chi_s} d\chi W(\chi)\delta_m(\chi\vec{\theta},\chi) \\ \gamma_1 &= \frac{1}{2}(\partial_1\partial_1 - \partial_2\partial_2)\psi(\vec{\theta}) \\ \gamma_2 &= \partial_1\partial_2\psi(\vec{\theta}). \end{aligned} \tag{4.18}$$

The function $W(\chi)$ is called the *lensing efficiency* function, and it defines the distribution of lensing matter along the line-of-sight:

$$W(\chi) = \frac{3}{2}\Omega_M H_0^2 a^{-1}(\chi)\chi \int_\chi^{\chi_{lim}} d\chi' n_s(\chi')\frac{\chi' - \chi}{\chi'} \tag{4.19}$$

for a distribution of source objects $n_s(\chi)$, which for a single galaxy is $n_s(\chi) = \delta(\chi - \chi_s)$.

When we measure the shape of a galaxy, we measure some flux distribution over an image in pixels. One of the simplest ways to measure the shape using the radially weighted quadrupole moments in the image coordinate system:

$$M_{ij} = \frac{\int d\theta_1 d\theta_2 \theta_i \theta_j w(\theta) I(\theta)}{\int d\theta_1 d\theta_2 w(\theta) I(\theta)}, \tag{4.20}$$

where $I(\theta)$ is the value of the pixel in the image in direction θ, and $w(\theta)$ is a weight function to minimize the contribution of the low signal-to-noise (S/N) regions. The corresponding

size (T) and complex-valued ellipticity of the galaxy is then given by:

$$\begin{aligned} T &= M_{11} + M_{22} \\ e_1 &= \frac{M_{11} - M_{22}}{M_{11} + M_{22}} \\ e_2 &= \frac{2M_{12}}{M_{11} + M_{22}}. \end{aligned} \tag{4.21}$$

A galaxy's true ellipticity ($e^{\rm int}$) has a well-defined response to an applied shear. We can approximate the observed ellipticity ($e^{\rm obs}$) as:

$$e^{\rm obs} \approx e^{\rm int} + \frac{2\gamma}{1-\kappa}. \tag{4.22}$$

We do not observe the shear and convergence individually, we only observe the *reduced shear*, which is the combined effect of both types of distortions $g = \gamma/(1-\kappa)$. In addition, we do not know the true shape and size and orientation of a galaxy, so we can only ever estimate the reduced shear, which we will denote $\hat{g}$ (a detailed discussion of different shear estimators is in §4.3.1).

The physical symmetry of the shear components means that if we rotate the system by 180°, the rotated shear would be identical to the unrotated shear. This means that we can represent the shear as a spin-2 ("spinor") pseudovector quantity and write shear as a complex-valued number $\gamma = \gamma_1 + i\gamma_2 = |\gamma|e^{2i\phi}$, where ϕ is a polar angle orientation of the galaxy.

We have not yet defined a precise coordinate system to measure our shears. A convenient coordinate system then might decompose the shears into a $+/\times$ basis, where the $+$ part is aligned parallel or perpendicular to the line connecting the pair of galaxies, and the $\times$ part is defined along a 45° angle relative to that line. Therefore, we decompose our estimated

reduced shear as:

$$\begin{aligned} \hat{g}_+ &= -\Re[\hat{g}e^{-2i\phi}] \\ \hat{g}_\times &= -\Im[\hat{g}e^{-2i\phi}]. \end{aligned} \tag{4.23}$$

In the same way, it is common to decompose the correlation functions of spin-2 quantities like shear into powers of divergence (E-mode) and curl (B-mode) components (LSST Science Collaboration et al. 2009). The shear correlation caused by a scalar gravitational potential would produce tangential shear and no curl shear, and therefore E-mode would be nonzero and the B-mode would vanish. Non-zero B-modes could be a sign of non-Gaussian distribution of primordial density fluctuations sourced by early gravitational waves (Meerburg et al. 2016). The observed ratio of E/B amplitudes is less than 1/500 over a large range of angular scales (Hilbert et al. 2009; Krause & Hirata 2010), which is likely outside the realm of measurement for current WL surveys. Therefore, non-zero B-modes are more likely a sign of systematic errors like poor shear estimation, atmospheric effects, astigmatism in the optics, or a non-flat focal plane, and they have been useful metric for finding systematic errors in cosmic shear measurement and analysis (for examples of old and new analyses of B-modes in cosmic shear surveys see Schneider et al. 2010; Becker 2013; Hamana et al. 2022).

In a realistic analysis, we have a catalog of $e^{\rm obs}$ for many galaxies. The galaxies in this ensemble have intrinsically varied shapes and orientations ("shape noise"). If galaxies have a underlying random dispersion of shapes and intrinsic alignments, then these stochastic variations will average out if we consider an ensemble of many galaxies ($\langle e^{\rm int}\rangle = 0$). However, the shear caused by WL occurs coherently around features in the LSS of the Universe, and this component will not average away. We would expected there to be some amount of coherent shear ("called cosmic shear") along the axis of the tidal gravitational field on different angular scales. Under this assumption, we can infer the reduced shear from the

observed ellipticities of an ensemble of many galaxies in a field ($g^{\rm inf} = \langle e^{\rm obs} \rangle$).

In practice, it is more useful to use the two-point correlation function of shears at different angular scales. For a pair of galaxies, one at position $\vec{\theta_1}$ and the other at $\vec{\theta_2}$, the angular distance between them is $\theta = |\vec{\theta_1} - \vec{\theta_2}|$. A 2PCF can be written for the shears of all pairs of galaxies separated by an angular separation θ. Typically, we will weight the averages by the inverse of statistical shape noise (based on the number of galaxies at a given separation) and estimated systematic error. We will again denote this operation by the standard standard angle brackets $\langle .. \rangle$ notation. The correlation function is then given by $\hat{\xi}_{\hat{g}\hat{g}}(\theta) = \langle \hat{g} \cdot \hat{g}^* \rangle = \hat{\xi}_+(\theta) + \hat{\xi}_-(\theta)$, where

$$\begin{aligned} \hat{\xi}_+(\theta) &= \langle \hat{g}_+ \hat{g}_+ \rangle(\theta) + \langle \hat{g}_\times \hat{g}_\times \rangle(\theta) \\ \hat{\xi}_-(\theta) &= \langle \hat{g}_+ \hat{g}_+ \rangle(\theta) - \langle \hat{g}_\times \hat{g}_\times \rangle(\theta). \end{aligned} \tag{4.24}$$

Note that the cross-correlation terms conveniently vanish due to parity symmetry in the Universe ($\langle \hat{g}_+ \hat{g}_\times \rangle = \langle \hat{g}_\times \hat{g}_+ \rangle = 0$).

The utility of the shear-shear autocorrelation function is extended with galaxy samples that span redshift bins, which allows us to take a tomographic view of the LSS through lookback time. We can calculate the correlation functions between all pairs of galaxies between redshift bin i and j that are separated by an angular separation θ. The Hankel transform of this gives us the power of the coherence as a function of angular wavenumber, and it turns out that this is also directly related to the matter power spectrum. We write this out as:

$$\xi_\pm^{ij}(\theta) = \frac{1}{2\pi} \int dl l C_l^{ij} J_{2\mp 2}(l\theta). \tag{4.25}$$

where J_n are just n-th order Bessel functions of the first kind. The coefficients are given by:

$$C_l^{ij} = \int_0^\infty dz \frac{W_i(z) W_j(z)}{\chi(z)^2 H(z))} P_\delta \left(\frac{l}{\chi(z)}, z \right). \tag{4.26}$$

Notice that the shear-shear correlation function is related to the integral of the matter power spectrum over the line-of-sight, where we have a wavenumber (l/χ), and $W_i(z)$ is the lensing efficiency function for tomographic redshift bin i with a known number density of galaxies as a function of redshift ($n^i(\chi'(z))$). Our cosmological theories most directly predict matter power spectrum, which can be derived analytically in the linear regime from Equation 4.13, or predicted from simulations or modeling (Krause & Hirata 2010) in the non-linear regime, which we can then compare to the matter power spectrum based on the observed shear power spectrum.

Shear-shear correlations are typically less than 1% over a given region of the sky, and the correlation drops off rapidly with distance (e.g. Miyatake et al. 2023). The precision with which we can measure the anomalous correlations depends on the effective number of galaxies we have in the sample. Some galaxies have very low signal-to-noise ratio, which means they will not effectively contribute to averaging out shape noise in the galaxy sample (Mandelbaum et al. 2018). In order to infer a correlation function with significance, one needs to utilize samples of many galaxies, typically on the order of $10^6 - 10^{10}$ galaxies (Almoubayyed et al. 2020).

The shear-shear power spectrum's sensitivity to both the background geometry and the large-scale structure of the Universe makes this a particularly powerful probe of cosmology. The amplitude of the shear power spectrum at a given scale depends on the distance-redshift relation, the value of the matter power spectrum at that scale, and to the curvature of the Universe. In particular, the shear power spectrum is expected to scale proportionally with $\Omega_M^2\sigma_8^2$ in the linear regime, where σ_8 is the amplitude of the mass fluctuations on the scale of galaxy clusters ($8h^{-1}$ Mpc) and is essentially the normalization of the matter power spectrum. The shape of the power spectrum is therefore degenerate between these two parameters. Occasionally, these parameters are combined into $S_8 = \sigma_8(\Omega_M/0.3)^{0.5}$, which is orthogonal to this degeneracy, and this is an important parameter to compare between

late-time cosmic shear and early-time Cosmic Microwave Background (CMB) probes (Planck Collaboration et al. 2020a; Di Valentino et al. 2021).

Today, WL surveys have catalogs of $O(10^6 - 10^8)$ shears across $10^2 - 10^3$ sq. deg. of the sky, and the strongest constraints on σ_8 come from cosmic shear measurements (Asgari et al. 2021; de Jong et al. 2017; Abbott et al. 2022; Miyatake et al. 2023). In 2023, the Hyper-Suprime Cam (HSC) survey released a shear power spectra analysis from a compilation of 3 years of data (Miyatake et al. 2023), correlating $O(10^6)$ shears across 416 deg^2 and reaching a point source 5σ r-band depth of 26 mag (approximately $z \sim 1.5$). They found $\Omega_M = 0.256^{+0.056}_{-0.044}$, $\sigma_8 = 0.818^{+0.089}_{-0.091}$, and $H_0 = 67.9^{+6.60}_{-4.30}$, which is consistent with ΛCDM theory from relic weak lensing probes and other model-independent late Universe probes (e.g. local distance ladder measurements with SN Ia). However, these inferred parameters are in disagreement with early Universe probes (e.g. 2σ disagreement on σ_8 with predictions from the cosmic microwave background measurements of the early Universe at $z \sim 1100$ from Planck Collaboration et al. 2020a). In general, there is a $2-3\sigma$ disagreement between early- and late-time probes on S_8, which could hint at potentially new physics (for an overview of different tensions for different parameters and their breakdown into statstical/systematic components, see Di Valentino et al. 2021).

4.3 Measurements of Uncertainty

A major issue in contemporary observational cosmology is to measure observables precisely and accurately enough to confidently distinguish between different cosmological models and reliably compare datasets between different experiments (Mandelbaum 2015). Driving this issue is the exponentially increasing data volume from cosmological surveys (separated into stages by Albrecht et al. 2006). For example, the earliest cosmic shear analyses Stage II surveys in the early 2000s used only about $10^4 - 10^5$ galaxies. In the 2010s, Stage III surveys

used about $10^6 - 10^7$ galaxies. Ongoing and future Stage IV surveys are expecting $10^8 - 10^{10}$ galaxies. In addition, we have only observed approximately 7,000 Type Ia supernovae confidently to date (Brout et al. 2022). Future Stage IV surveys are expected to find over 200,000 new events (Ivezić et al. 2019). As the number of observed objects increases, the statistical uncertainty of our measurements decreases roughly with $1/\sqrt{N}$. Future surveys will soon no longer be limited by statistical uncertainties from limited numbers of objects, and the need to suppress the variety of different systematic uncertainties in analyses of future survey experiments will become ever more important.

Following the convention in report of the Dark Energy Task Force (DETF, also see Albrecht et al. 2006), surveys report the uncertainty on different cosmological parameters based on the 95% confidence interval width of the posterior probability distribution after marginalizing over a prior range of nuisance parameters corresponding to different sources of systematic uncertainties in the data. Statistical uncertainties refer to the widths of posterior probability density distributions with fixed systematic uncertainties. Systematic uncertainties refer to the lack of knowledge that results from the presence–or imperfect treatment–of systematic biases (or "errors") in cosmological analyses. Surveys will define an overall error budget (for either specific or combinations of cosmological probes) and then place a limit on the systematic uncertainties as a fraction of the statistical uncertainties. Some systematic biases (esp. observational or measurement effects) are "calibratable" meaning they can be explicitly estimated and marginalized in advance of cosmological inference using external data or simulations. Other systematic biases (esp. astrophysical effects like intrinsic alignments, selection effects, models of physical processes like baryonic feedback and galaxy formation, etc.) are only "self-calibratable" meaning they cannot be estimated in advance of the cosmological inference analysis. Therefore, most surveys make sure the models of self-calibrated systematics are sufficiently complex to capture all sources of errors and only lay out requirements on uncertainties from calibratable systematics (e.g. The LSST Dark

Energy Science Collaboration et al. 2021). This thesis will focus on sources of calibratable systematics since they can be measured prior to any real survey data or phenomenological inference analyses.

Systematic biases can occur during many stages in the pipeline from the sky to the data catalogs of the relevant cosmological information (e.g. lists of galaxies and their positions, fluxes, colors, redshifts, etc.) for a given cosmological probe. Figure 4.1 illustrates an example of the different stages where systematic biases can develop in an image analysis of a single a galaxy:

1. Observation: the optical system itself and the act of sampling the sky on a finite pixel grid degrades the overall image quality. This includes atmospheric effects like turbulence, atmospheric glow, sky background, differential chromatic refraction, etc. and optical system effects like aberration, vignetting, pixelization, shot noise, etc. These systematics are modeled per-image from on-sky calibration sources. More information will be given in §4.3.1.

2. Detector response: the instrument can add noise and imprint signature artifacts on the image. This includes electronic noise, and imperfect correction of quantum efficiency variations, charge-transfer inefficiency, pixel size variations, sensor defects, brighter-fatter effect, etc.). These systematics need to be characterized and calibrated individually and removed from raw data in a process called Instrument Signature Removal (ISR). More information will be given in §4.4

3. Measurement: there are different methodologies for determining the shape and size of different sources when generating source catalogs from images, and biased algorithms can impart their own systematic uncertainties. This firstly includes methods for stacking of exposures into coadded images. It also includes sky background subtraction, object detection. For a shear analysis we need to be able to deconvolve the PSF from

an image and estimate the shear from the observed shape. This depends on the choice of shear estimation algorithm and how it handles PSF deconvolution, noise bias, model bias, selection bias, etc.) It is also known that object detection is sensitive to the object shape Sheldon et al. (2023). More information will be given in §4.3.1.

4. Analysis: limitations of models or methods that go into compiling catalog data to constrain cosmological analyses. There are uncertainties associated with limitations of models of astrophysical physical processes in the non-linear regime like baryonic feedback, extinction effects, etc. and uncertainties associated with treatments of intrinsic unknowns like coherence of unlensed galaxy shapes or intrinsic alignments, etc. For example, the scale of the power spectrum probed with weak lensing by galaxy clusters is within the non-linear regime, and it requires the specialized simulations/modeling to predict its morphology from theory (Krause & Hirata 2010). The accuracy of these models are a source of systematic uncertainty, and while they are certainly complex, they are well-studied and unlikely to seriously affect modern ambitious weak lensing probes (Krause & Hirata 2010).

Understanding and mitigating systematic biases is essential to achieve the precision necessary to distinguish between cosmological models and reliably compare measurements between different experiments, particularly given the current tensions between early- and late-universe observations. In this section, I will cover each of these stages and how different sources of systematic uncertainty are parameterized, measured, and folded into cosmological data analyses.

4.3.1 The Point Spread Function (PSF)

One metric of scientific quality in an astrophysical measurement, an image, and even a survey is the quality of the *seeing*, which is #2 in the list of most important design drivers

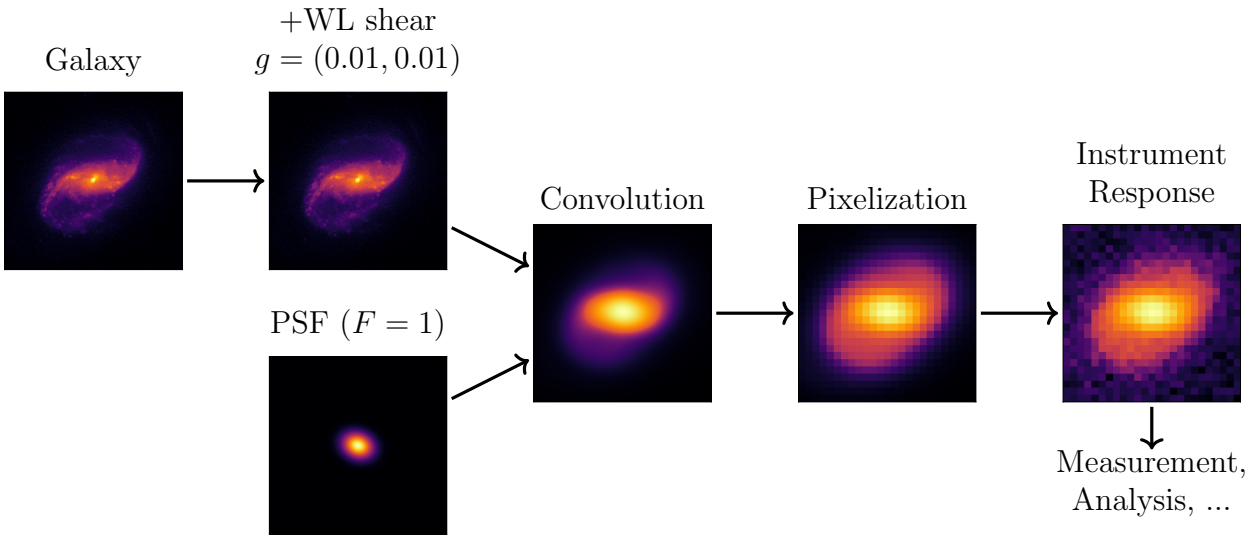


Figure 4.1: Illustration of the physical and observational distortions to objects for a general ground-based observatory. First, the object is distorted by gravitational fields along the line-of-sight (weak gravitational lensing), then it is distorted by the flux-normalized point spread function (atmospheric turbulence, optical system distortions, pixelization, charge diffusion in sensors, etc.), then the light profile is sampled by the imaging sensor which imparts it's own unique signature.

for Stage IV ground-based WL surveys (see Ivezić et al. 2019). All astrophysical analyses are fundamentally reduced to measuring light intensities over some region in the sky. The smallest unit of sky is defined by the size of the field-of-view of a single pixel on the detector array. In a perfect scenario, we would ensure that a photon coming from a specific region of sky gets recorded in the correct pixel, and that all photons that come from the same region of the sky always get recorded in that same pixel in the same image. However, when a photon makes it to Earth, it will first encounters the atmosphere and the wavefront of the light will be refracted and diffracted through turbulent layers with different densities, temperature gradients, and dusty/icy particulates. The amount this distorts the light of course depends on the spectral energy distribution of the light and can vary rapidly in time over the duration of an exposure. Once it finally makes it to the telescope, it then has to encounter the optical system which may further blur the light due to optical aberration, telescope vibration, and

telescope tracking errors, etc. Then it reaches the imaging sensor, the photon is converted into electrical charge, and this the light is further blurred through thermal diffusion of the charge before it is collected and confined in potential wells at the bottoms of pixels in the detector. Therefore, when we record a photon in an image, we are not measuring the exact point where that photon came from. Instead, we measure a sample from a probability distribution of where that photon actually came from. The transfer function that maps a single point on the image to a region of the sky is called the point spread function (PSF), and it is effectively the impulse response of the imaging apparatus.

The measured image ($I_{meas}(\theta_x, \theta_y)$) can be thought of a convolution between a true image ($I_{true}(\theta_x, \theta_y)$) and the PSF:

$$I_{meas}(\theta_x, \theta_y) = I_{true}(\theta_x, \theta_y) \otimes PSF(\theta_x, \theta_y) \tag{4.27}$$

The PSF is analytically a continuous distribution function. It is typically described by a flux-normalized 2D Gaussian, although other profiles with additional complexities and parameters are also sometimes used to model it (e.g. Airy, Moffat, Kolmogorov, and von Karman profiles are some common ones). The best case scenario is when there is no distortion at all, and the PSF is just a Dirac delta function. In reality, this never happens, even in rare, so-called fast "lucky" images, where we just happen to get a relatively clear shot into space (for an interesting example of a real analysis like this, see Law et al. 2006). We also have to consider that the PSF is sampled on a pixel grid. While PSFs with a large size make the image more blurry and take away useful information, PSFs that are too small could be undersampled by the pixel grid.[2]

[2]A space telescope is more likely to have an undersampled PSF than a ground telescope given the lack of atmosphere (Lauer 1999). E.g. the median seeing in optical wavelengths at Rubin Observatory's site on Cerro Panchón is 0".65 arcsec with a 0".2/pixel platescale (Ivezić et al. 2019), and the median seeing for Euclid (VIS instrument) is 0".16 arcsec with a 0".101/pixel platescale (Euclid Collaboration et al. 2024). However, being on the ground enables larger entendue (effective field-of-view × mirror size) as well as more complex optical systems, such as active or adaptive optical systems (Neill et al. 2014; Crenshaw et al. 2024).

The PSF is something that we can only estimate from our data. Conveniently stars are all effectively infinitely small point sources and can be used to calibrate the PSF. The profile of a star in a single exposure is an estimate of the time-averaged PSF at that location on the sensor and in the field-of-view. To estimate the local PSF, one can either measure the unweighted, radial second moments from flux values in pixels or fit an analytical profile to a cutout image around the centroid of the star (Equation 4.20). For an analytical Gaussian for instance, the PSF size can be estimated by assuming $T = 2\sigma^2 = I_{xx} + I_{yy}$.

The PSF can vary spatially across the sensor and the field-of-view. Several PSF estimation algorithms have been used (see PSFEx and PIFF, Bertin 2011; Jarvis et al. 2021, respectively), however the general process proceeds as follows:

1. Identify/select stars in an exposure.
2. Measure the PSF at the location of each star (the "observed PSF").
3. Interpolate the local PSFs across the image (a common choice is to use 2D polynomials or Gaussian Process interpolation).
4. Reconstruct the PSF at the location of each object in the image (the "model PSF").

Different methods can vary in terms of their treatments of stellar identification, local PSF estimation, and interpolation method. Perhaps not surprisingly, it is advantageous to use as many stars as possible in an image to estimate the PSF. In particular, high S/N ratio stars constrain their local PSF better than dim stars since the relative uncertainty for a Poisson-noise dominated measurement goes with the photon count as $\sqrt{N}/N$ (Paulin-Henriksson et al. 2008). Note that the number of stars in an image can vary with the telescope's pointing. Obviously, images taken in the direction of the Galactic plane will typically contain more stars than images taken above or below the Galactic plane.

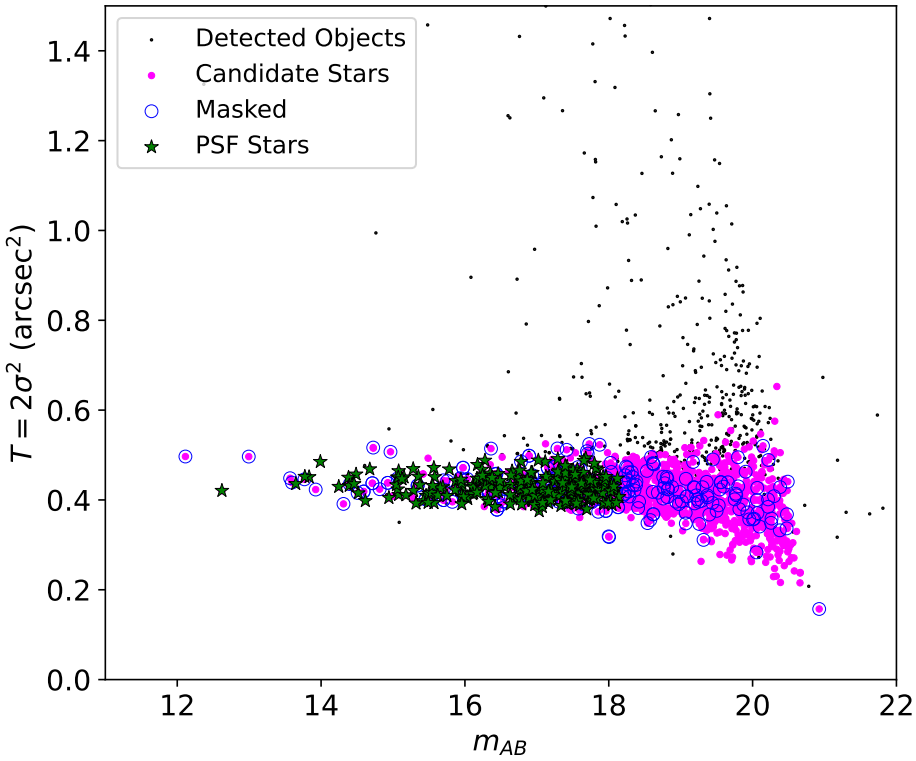


Figure 4.2: An example size-magnitude diagram for all detected sources from a single 30s exposure targeting a high stellar density field from a ground-based telescope that has a 6.7×6.7 arcmin field-of-view and a primary mirror with an effective diameter of 1.2m. We calculate the size T based on the best-fitting Gaussian to each source. The black points are all the detected sources in an image (determined via S/N thresholding), which may not be real objects (stars, galaxies, etc.). The magenta points are sources that are likely to be stars based on their extendedness (difference between the de Vaucouleurs flux and the Gaussian flux) and other parameters (ref. to Bertin 2011). The blue circles show which of these sources were masked due to their positions on an edge of the sensor or a known sensor defect or were blended with another source. The green points show which sources were ultimately selected as stars used to estimate the PSF. For details about the exact source detection algorithm and method of star selection, see Bertin (2011); Jurić et al. (2017); Bosch et al. (2019).

The average PSF of all stars across the focal plane is typically used to define image quality. Since all stars are intrinsically point sources, all stars in the same exposure should have the roughly the same PSF, regardless of their differences in brightness. In fact, this property is often used to identify stars to use for PSF estimation. Figure 4.2 shows an example of the size-magnitude relation of detected sources in a single exposure. The stars are

found in the so-called "stellar locus" which roughly follows a flat size vs. magnitude relation. The mean size of the stellar locus defines the average distortion during the exposure. The average full width at half maximum, which for a Gaussian is FWHM $= 2\sqrt{2\ln 2}\sigma$, is referred to the "seeing" and is the typical a metric of image quality.

4.3.2 PSF modeling error

The PSF has a significant effect on the apparent size and shape of objects. The measurements of all celestial objects need to be calibrated against the inferred PSF at their location in an image. For a systematic error in a PSF ($\Delta T_{\rm PSF} = T_{\rm PSF}^{\rm (true)} - T_{\rm PSF}^{\rm (model)}$), this will result in a systematic error on the shape of the imaged source.

For example, we use the "PSF flux" to measure the brightness of SN Ia. Since SN Ia are relatively faint, their profiles are typically dominated by noise, making it difficult to measure their flux over some circular footprint by straight-forward pixel-by-pixel integration ("aperture flux"). However, SN Ia are effectively star-sized objects from our point of view, so we know that they should be shaped like the PSF. We infer the flux of SN Ia by fitting the PSF to the SN Ia image. The total flux is then just the normalization of the PSF that results in the best fit to the image. However, if the model PSF is biased from the true PSF, this results in a flux error.

Similarly, PSFs are required to estimate shear in weak lensing analyses. Shear estimation essentially involves deconvolving the image with the PSF to get an accurate estimate of the shear field at the location of a galaxy. We typically describe modeling biases in shear estimation as either multiplicative or additive. For a given ensemble of galaxies, with some true shear (g_{true}), the estimated shear over an ensemble of galaxies ($\langle \hat{e} \rangle$) is parameterized as:

$$\langle \hat{e} \rangle = (1 + m) g_{true} + \alpha e_{\rm PSF} + c, \tag{4.28}$$

where m is the multiplicative bias and the combination of $\alpha e_{\rm PSF} + c$ is the additive bias. The additive bias is sometimes given as a single constant (Heymans et al. 2006), although in this case we follow the formalism by Jarvis et al. (2016), which separates the additive bias into *PSF leakage*, which is dependent on the PSF (α) and a separate additive offset (c).[3] To understand how errors in shear estimation propagate to errors in the shear-shear correlation function, we can simply plug Equation 4.28 into the correlation function:

$$\delta\xi_{\pm}(\theta) \simeq 2m\xi_{\pm}(\theta) + \alpha^2\xi_{\pm}^{PP}(\theta) + \xi_{\pm}^{cc}(\theta), \tag{4.29}$$

where $\xi_{\pm}^{PP}$ refers to the autocorrelation function of the PSF and $\xi_{\pm}^{cc}$ is the autocorrelation function of the additive error offset. This also assumes that the different types of biases are uncorrelated.

Previously, shear estimators took a moments-based approach (Kaiser et al. 1995; Rhodes et al. 2000; Melchior et al. 2011), measuring the image moments of both the galaxy and the PSF from the image, or a modeling approach (Massey & Refregier 2005; Refregier 2003; Refregier & Bacon 2003; Nakajima & Bernstein 2007; Miller et al. 2013), which which performs a maximum likelihood fit with a PSF-convolved galaxy profile to the image. Typically, these shear estimators are biased for the low S/N galaxies (Bernstein & Jarvis 2002). Even the best shear estimators will always introduce some "modeling bias" since there is always a limit to how well we can model all of the realistic complexity of the structure of a galaxy; galaxies do not usually have elliptical isophotes and are typically not well captured by second-order image moments alone (cf. Equation 4.20, ref. Massey et al. 2007; Bernstein 2010; Zhang & Komatsu 2011); even if the estimated PSF is perfectly known, the shear estimator can still cause some of the PSF size and shape to "leak" into the galaxy size and shape. In a blind

[3]Note that the multiplicative and additive biases could theoretically have different values along each direction of the decomposition, however Jarvis et al. (2016) measured these separately for DES Y1 and found that this anisotropy m is negligible.

challenge of different shear estimators on real galaxy shapes (GREAT3), Mandelbaum et al. (2015) found that the model bias was found to be approximately 1% for bulge+disc model across many shear estimators. Modern shear estimators have made large improvements to these earlier approaches. One approach to calibrate the noise bias with image simulations of representative ensembles of galaxies (Mandelbaum et al. 2018; MacCrann et al. 2022). Most recently, cosmic shear surveys have proposed to use perturbative methods (Huff & Mandelbaum 2017; Sheldon & Huff 2017; Sheldon et al. 2020, 2023; Li et al. 2018; Li et al. 2022; Li et al. 2023), which approximate the responsivity of the shape of the galaxy to a shear, and these methods produce sub-10^{-3} biases for isolated galaxies where the PSF is known perfectly.

In the case where we have a perfect shear estimator, PSF modeling errors can affect the shear-shear correlation function (from Paulin-Henriksson et al. 2008):

$$\delta e_{sys} = (e - e_{\rm PSF}) \left(\frac{T_{\rm PSF}}{T_{\rm gal}}\right) \frac{\delta T_{\rm PSF}}{T_{\rm PSF}} - \left(\frac{T_{\rm PSF}}{T_{\rm gal}}\right) \delta e_{\rm PSF}, \tag{4.30}$$

where $T_{\rm gal}$ is the intrinsic size of the galaxy and $T_{\rm PSF}$ is the size of the PSF. These errors enter into the tangential shear correlation function (ξ_+) as:

$$\begin{aligned} \delta\xi_+(\theta) = 2 \left\langle \frac{T_{\rm PSF}}{T_{\rm gal}} \frac{\delta T_{\rm PSF}}{T_{\rm PSF}} \right\rangle \xi_+(\theta) + \left\langle \frac{T_{\rm PSF}}{T_{\rm gal}} \right\rangle^2 \rho_1(\theta) - \alpha \left\langle \frac{T_{\rm PSF}}{T_{\rm gal}} \right\rangle \rho_2(\theta) \\ + \left\langle \frac{T_{\rm PSF}}{T_{\rm gal}} \right\rangle^2 \rho_3(\theta) + \left\langle \frac{T_{\rm PSF}}{T_{\rm gal}} \right\rangle^2 \rho_4(\theta) \\ - \alpha \left\langle \frac{T_{\rm PSF}}{T_{\rm gal}} \right\rangle \rho_5(\theta), \end{aligned} \tag{4.31}$$

where ρ_{1-5} are called the "ρ-statistics" and are given by:

$$\rho_1(\theta) \equiv \langle \delta e^*_{\rm PSF}(\vec{x}) \delta e_{\rm PSF}(\vec{x} + \vec{\theta}) \rangle, \tag{4.32}$$

$$\rho_2(\theta) \equiv \langle e^*_{\rm PSF}(\vec{x})\delta e_{\rm PSF}(\vec{x}+\vec{\theta})\rangle, \tag{4.33}$$

$$\rho_3(\theta) \equiv \left\langle \left(e^*_{\rm PSF}\frac{\delta T_{\rm PSF}}{T_{\rm PSF}}\right)(\vec{x}) \left(e_{\rm PSF}\frac{\delta T_{\rm PSF}}{T_{\rm PSF}}\right)(\vec{x}+\vec{\theta})\right\rangle, \tag{4.34}$$

$$\rho_4(\theta) \equiv \left\langle \delta e^*_{\rm PSF}(\vec{x}) \left(e_{\rm PSF}\frac{\delta T_{\rm PSF}}{T_{\rm PSF}}\right)(\vec{x}+\vec{\theta})\right\rangle, \tag{4.35}$$

$$\rho_5(\theta) \equiv \left\langle e^*_{\rm PSF}(\vec{x}) \left(e_{\rm PSF}\frac{\delta T_{\rm PSF}}{T_{\rm PSF}}\right)(\vec{x}+\vec{\theta})\right\rangle, \tag{4.36}$$

The first term is a multiplicative bias term and all others are additive biases. We see that errors in modeling the PSF can result in multiplicative biases, PSF leakage biases, and additive offset biases to shear estimation, depending on the size of the PSF to the intrinsic size of the galaxy. Note that the amplitude of the impact of a PSF modeling error on a galaxy image depends on the relative size of the PSF and the galaxy. Note that if the size of the PSF is zero, there are no biases to the shear-shear correlation function (again assuming an already perfect shear estimator). Note that we can also write a similar set of equations for the $\xi_\times$ component.

PSF estimation is one of the primary sources of systematic uncertainty in cosmological analyses (Paulin-Henriksson et al. 2008; Massey et al. 2007; Zhang et al. 2022; Zhang et al. 2023). The LSST Dark Energy Science Collaboration et al. (2021) defines several limitations on these biases for its SN Ia and weak lensing science goals for the Legacy Survey of Space and Time, an upcoming Stage IV WL survey. For SN Ia photometry, the calibrated flux must be accurate to within 1% of the true flux. For weak lensing science, the document defines the maximum PSF modeling bias as $\delta T_{\rm PSF}/T_{\rm PSF} < 10^{-3}$. It also places requirements on the residual PSF ellipticity via the correlation function of the biases on different components of ellipticity. For $E_1(\theta) = \langle \delta e_1 \delta e_1^* \rangle$, $E_2(\theta) = \langle \delta e_2 \delta e_2^* \rangle$, and $E_X(\theta) = \langle \delta e_1 \delta e_2^* \rangle$, averaged over all stars in the full (10 year) survey, $E_{1/2/X}$ must be less that 3×10^{-5} on $\theta < 1$ arcminute scales (around the scale of galaxy clusters). In particular, biases on E_X produce so-called B-mode

shear, which is expected to be approximately zero in the absence of systematic errors in the data catalog. It is common to fit these errors as nuisance parameters. In the past, cosmic shear analyses have only accounted for these PSF systematics due to the second moments of the PSF. However, Zhang et al. (2021) showed that the fourth moment terms were significant in the HSC Y3 PSF, which by itself could have lead to $\sim 0.3\sigma$ biases in cosmological parameters when not accounted for in the PSF model. Although, no requirements have yet been laid out for Stage IV surveys to constrain the higher-order statistics.

4.4 Measurements with CCDs

Our current cosmological model (ΛCDM) rose to prominence thanks to technological advancements in photon-detection technology since the 1970s that have enabled some of the most rigorous constraints of cosmological parameters to date. The development of solid-state devices for imaging such as the charge-coupled device (or "CCD") by Willard S. Boyle and George E. Smith at Bell Laboratories appeared to promise advantages in size, weight, power, noise, sensitivity, linearity with signal, dynamic range, photometric accuracy and repeatability, spectral response, durability, and importantly, scalability over the photoelectric ancestors used in the early days of astronomy (namely photographic film plates and tube-type detectors). These devices work by collecting charges produced by incoming light in arrays of pixels defined by potential wells. This format alone was attractive to early astronomers because the data could be digitized and transmitted, enabling telescopes to be put in space, and the devices were responsive to light in the ultraviolet (UV) to near-infrared (NIR) parts of the spectrum (4.1eV to 1.26eV), making them useful for a wide variety of astrophysical experiments. The 61-inch telescope on Mt. Lemmon was the first instrument to show the potential of CCDs for astrophysical experimentation (Smith 1976). And very early on, the CCD was proposed for the Wide Field/Planetary camera on the Hubble Space

Telescope, which was then called the Large Space Telescope (McLean 2004).

However, since the early days of the CCD development, they have presented few challenges. For example, data transfer efficiency in the earliest detectors was poor, which limited the size of the sensors to only 100 pixels (Leach et al. 1980). In addition, these detectors needed to be cooled below 200K to eliminate the production of infrared photons *inside* the detector itself (so-called "dark currents" or "thermal noise"). Another concern was manufacturing yield, as the sensors were notoriously difficult to produce (Janesick 2001). Detectors also experienced quantum efficiency variations that made it difficult to produce reliable relative flux measurements between sensors. Today, we also know that different physical effects inside CCDs can imprint artifacts on the images (e.g. electronic noise, charge-transfer inefficiency, pixel size variations, tree rings, annealing patterns, sensor defects, brighter-fatter effect, support structure, etc.). These can result in one or some combination of photometric, astrometric, size, and shape errors in images of celestial objects, and in many ways mimic the impact of new physics on these different astrophysical observables (Esteves et al. 2023). Despite many recent advancements to the design of CCDs, recent analyses have found it difficult to achieve ground-based flux measurements with fractional uncertainties in accuracy of less than 3% using them (Stubbs & Tonry 2006; Stubbs 2013). Improperly treating these instrument systematics can impart additional uncertainty in downstream scientific pathways os cosmological surveys.

Stage III cosmological surveys (defined by Albrecht et al. 2006) such as the Dark Energy Survey (DES; Dark Energy Survey Collaboration et al. 2016), the Kilo-Degree Survey (KiDS; de Jong et al. 2017), the Hyper Suprime-Cam survey (HSC; Aihara et al. 2018) utilized CCDs to efficiently map large volumes of universe since they could be easily scaled to cover a large field of view by multiplexing several sensors together. These surveys have constrained the dark energy equation of state ($\omega_0 \sim -1$) to around 3% precision using 100 million galaxies and $5,000$ type Ia supernovae (SNe) (Abbott et al. 2022). However, these uncertainties are

still dominated by statistical uncertainty from the limited quantity of data. Future Stage IV cosmological surveys will also use CCDs, including the Nancy Grace Roman Space Telescope (Spergel et al. 2015) and Euclid (Laureijs et al. 2011) and the Legacy Survey of Space and Time (LSST) (LSST; Ivezić et al. 2019; LSST Science Collaboration et al. 2009). They will drastically increase our view of the Universe and will be able to significantly tighten the statistical uncertainties between cosmological models, but that will make them more sensitive to systematic errors like those produced by the instruments themselves. These surveys therefore place strict requirements on systematic uncertainties and have extremely tight error budgets.

Chapter 5

Mitigation of the Brighter-Fatter Effect (BFE) in LSSTCam

The largest of the Stage IV surveys is LSST, which will be conducted with the Simonyi Survey Telescope at the Vera C. Rubin Observatory, which is under construction on the summit of Cerro Pachón in Chile (see Figure 5.1). The instrument for this is a 3.2 gigapixel camera (LSSTCam), which contains 201 individual CCDs, with 189 designated specifically for science imaging. The sensors are fully depleted high-resistivity, bulk silicon CCDs developed by two separate vendors. One type is made by Imaging Technology Laboratories (ITL), and the other is made by Teledyne e2v (E2V) to similar general architectural specifications. These sensors are arranged by type into 3×3 groups called raft-tower modules (RTMs) or rafts, which can each operate as an independent camera. Each CCD sensor is $4\,\mathrm{cm} \times 4\,\mathrm{cm}$ and made up of sixteen, 1 megapixel channels each read out by its own amplifier and readout electronics. Each pixel on the LSST Camera focal plane is $10\,\mu\mathrm{m} \times 10\,\mu\mathrm{m}$ and has a depth of $100\,\mu\mathrm{m}$ (see Figure 5.2) (Lopez et al. 2018).

The combination of a large focal plane of CCDs and a large field of view will enable LSST to generate the largest map of the Universe to date. The survey plans to image 18,000

Figure 5.1: A view of the Vera C. Rubin Observatory on the Cerro Pachón ridge in northern Chile, credit: Rubin Observatory/NOIRLab/NSF/AURA/B. Quint.

deg^2 of the sky with $O(1000)$, 30 second exposures over 10 years in six filters (*ugrizy*) in the optical and near-infrared (NIR) parts of the spectrum, generating approximately 20 terabytes of data every night. By the end of the survey it is expected to catalog over 20 billion galaxies and 200,000 type Ia SNe, and it aims to measure cosmological parameters to sub-1% precision (The LSST Dark Energy Science Collaboration et al. 2021; Ivezić et al. 2019).

Much of the design and current development of instrumentation for LSST focuses on characterizing and reducing the impact of systematic sensor artifacts with the goal to produce sub-percent level precision measurements of cosmological parameters and test currently prevailing thermodynamic models of the universe and theories of dark matter (Albrecht et al. 2006). It is also no surprise that all of LSST's scientific analyses rely on the assumption that observable measurements come from pixels in an image that are all the same size, that are all independent from one another, and all have well-defined response functions. Detector effects in the instrument can impart systematic errors that violate these assumptions, and the only way to handle them is to individually isolate, diagnose, and treat each of these potential sources of calibratable systematic uncertainties and test their corrections. To meet

these requirements for LSST (The LSST Dark Energy Science Collaboration et al. 2021), we need to understand the systematic effects in our sensors at the far limits of the capabilities of our instrumentation (Ivezić et al. 2019).

Many of the problematic sensor behaviors are considered generally static and can be simply calibrated from well-known sources or modeled as intrinsic constants on a sensor and easily removed from raw images (Plazas Malagón et al. 2024). However this method would not work with locally variable or otherwise signal-dependent effects, and such effects are known to exist (Stubbs 2013; Antilogus et al. 2014; Astier et al. 2019).

One such dynamic effect is the brighter-fatter effect (BFE), which makes brighter sources appear larger in a recorded image. This effect occurs during the integration of an exposure, when captured photocharges produce transverse electric fields on incoming photocharges, causing the incoming charges to deflect into neighboring pixels (Downing et al. 2006; Holland et al. 2014; Lage et al. 2017). The effect can be modeled as a reaction to quasistatic changes in effective pixel area as charges collect in the potential wells of the pixels. This results in correlations between neighboring pixels that arise at high signal levels. The magnitude of the BFE depends on the surface brightness profile of the source itself and thus cannot be modeled solely by its location on a sensor. The consequence is that the BFE breaks the critical assumption of all experimental image analysis that the pixels are independent light collectors that perfectly obey Poisson statistics.

Reduction of Poisson noise by "smearing" due to the BFE in exposures with taken with uniform illumination across all pixels ("flat field" images) has been observed in LSSTCam sensors by Antilogus et al. (2014) and Lage et al. (2017), and in detectors used in other astronomical cameras such as Hyper Suprime-Cam (HSC) by Coulton et al. (2018), the Dark Energy Camera used by the Dark Energy Survey (DES) by Gruen et al. (2015), the Wide Field Camera 3 H1RG detector of the Hubble Space Telescope by Plazas et al. (2017), MegaCam by Guyonnet et al. (2015), the Mid-Infrared Instrument (MIRI) on board the

James Webb Space Telescope (JWST) by Argyriou et al. (2023), and in the near-infrared (NIR) detectors of the Wide Field Imager of NASA's Nancy Grace Roman Space Telescope (Plazas et al. 2018; Hirata & Choi 2020; Choi & Hirata 2020; Freudenburg et al. 2020; Plazas Malagón et al. 2023). This effect was originally observed as a growth in the stellar size-magnitude relation. Stars are particularly impacted by the BFE since they typically have the highest signal-to-noise ratios of all detected sources in single exposures. This presents a potentially major issue for downstream scientific efforts since stellar properties are used calibrate the distortions by the atmosphere and optical system (the point spread function or PSF), and the light from all celestial objects are calibrated against this PSF.

Several corrections have been proposed by Antilogus et al. (2014), Gruen et al. (2015), Coulton et al. (2018), and Astier, Pierre & Regnault, Nicolas (2023). Specifically, the two most recently proposed corrections (Coulton et al. 2018 and Astier, Pierre & Regnault, Nicolas 2023) are currently implemented in the LSST science pipelines, and Coulton et al. (2018) is currently used in the HSC data reduction pipeline and is currently planned to be used in the LSST science pipelines (Jurić et al. 2017; Bosch et al. 2018, 2019). These methods fall into two categories: scalar corrections and vector corrections. Scalar corrections model the BFE at the pixel-level while vector corrections model it at the sub-pixel level. All of the BFE corrections are derived from calibrations from measurements of flat field images, however, the corrections will ultimately matter for measurements of more realistic science images like star field images. However, the underlying electric fields in flat images might not approximate the underlying electric fields in stars images. The effectiveness of each correction depends on the correspondence between the assumptions they make and the true underlying physics of the sensors.

In this part of the thesis, we will directly measure the BFE in the LSST Camera and our ability to correct it and understand its impact on LSST science pathways. In §6, we will introduce the theory behind the BFE and its correction by Coulton et al. (2018) and weigh

its limitations in light of more recent findings by Astier et al. (2019). In §7, we will describe our laboratory measurements and image processing framework. In §8, we will explore the dynamic response of our sensors to charges in these data. In §9, we will construct an improved application of the correction by Coulton et al. (2018) (§9.1) and apply it to flat fields (§9.2) and artificial stars (§9.3). In §10, we will discuss the relative results between these two cases and directly test the underlying assumptions of the correction made by Coulton et al. (2018). Finally in §10.2, we discuss how residual BFEs could influence several of LSST's dark energy science goals and motivate future work on analysis structures to mitigate its impact. Throughout the text, we adopt the abbreviations "px" and "el" to refer to pixel and electron units, respectively.

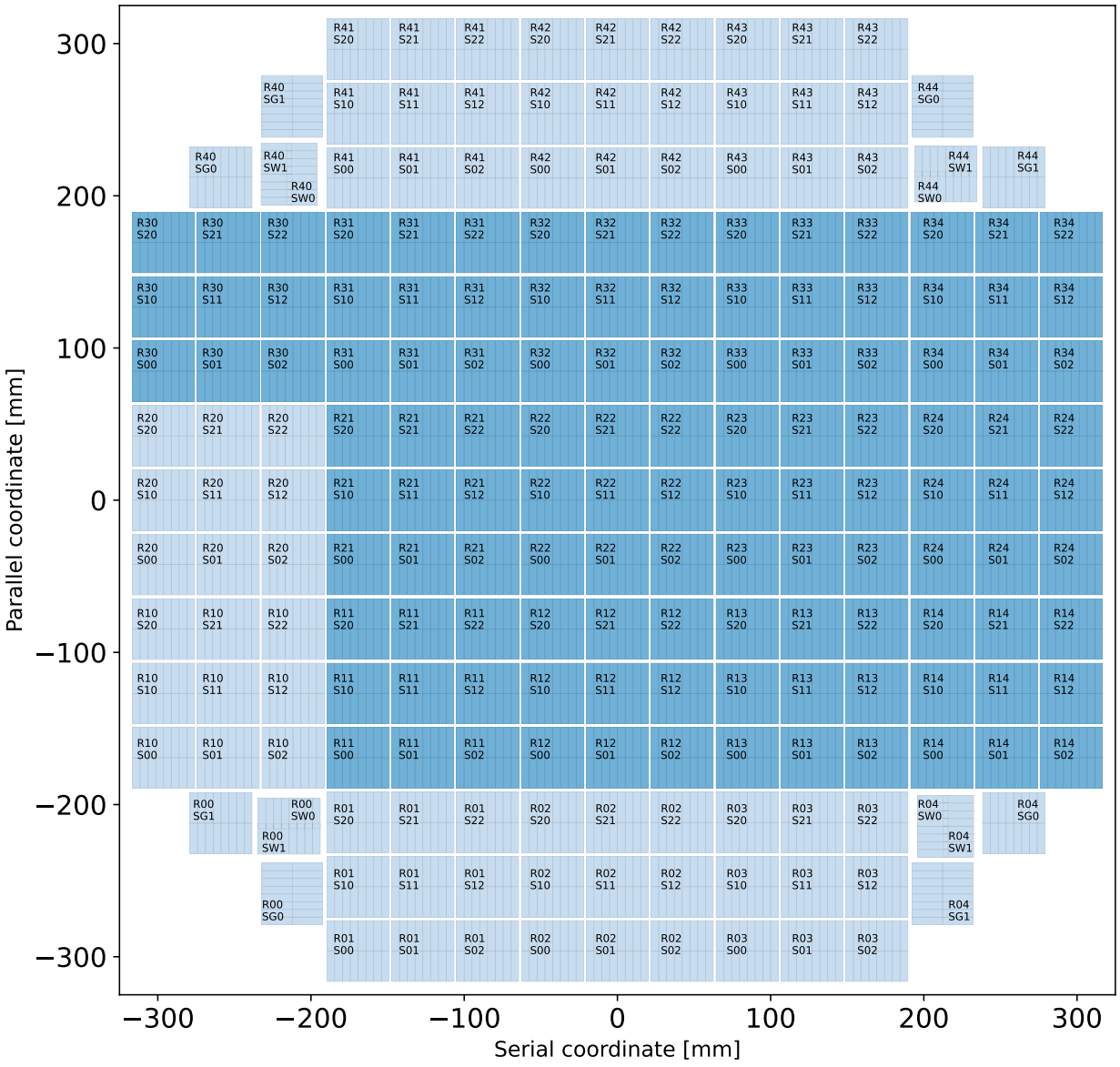


Figure 5.2: Map of the LSSTCam focal plane and coordinate system. The sensors are arranged into 3×3 groups called raft-tower modules (RTMs), labeled by the R/S numbers, respectively, an each sensor is made up of 16 channels (1 MP/channel). The darker colored sensors are E2V-type and the lighter colored sensors are ITL-type. The corner rafts (R00, R04, R40, R44) are used for wavefront sensing and guiding, and all other rafts are used for science.

Chapter 6

The Brighter-Fatter Effect

Several studies have shown that the pixel covariances in flat field images can be used to measure the BFE in the sensors of other optical and near-IR survey cameras (Antilogus et al. 2014; Gruen et al. 2015; Guyonnet et al. 2015; Coulton et al. 2018; Astier, Pierre & Regnault, Nicolas 2023; Hirata & Choi 2020; Freudenburg et al. 2020; Plazas et al. 2017; Plazas et al. 2018). The covariance can be numerically formalized between an arbitrary central pixel, $\mathbf{x} = (0, 0)$, assumed to be far from the edge of the detector, and a neighboring pixel $\mathbf{x}' = (i, j)$ in the difference of two flat field exposures, F_1 and F_2, with the same nominal signal level and $\mu = \text{avg}(F_1, F_2)$:

$$C_{ij} = \frac{1}{2}\text{Cov}(F_1(\mathbf{x}) - F_2(\mathbf{x}), F_1(\mathbf{x}') - F_2(\mathbf{x}')). \tag{6.1}$$

The directionality on each sensor is defined in regard to the CCDs' frontside architecture, where the vertical, $\hat{\mathbf{u}} = (0, 1)$, direction is along the "parallel" readout direction defined by orientation of the buried channel stops in the sensor, and the horizontal, $\hat{\mathbf{u}} = (1, 0)$, direction is along the "serial" readout direction defined by the orientation of the buried gate structure.

Figure 6.1 shows an example of a densely sampled measurement of the variance function

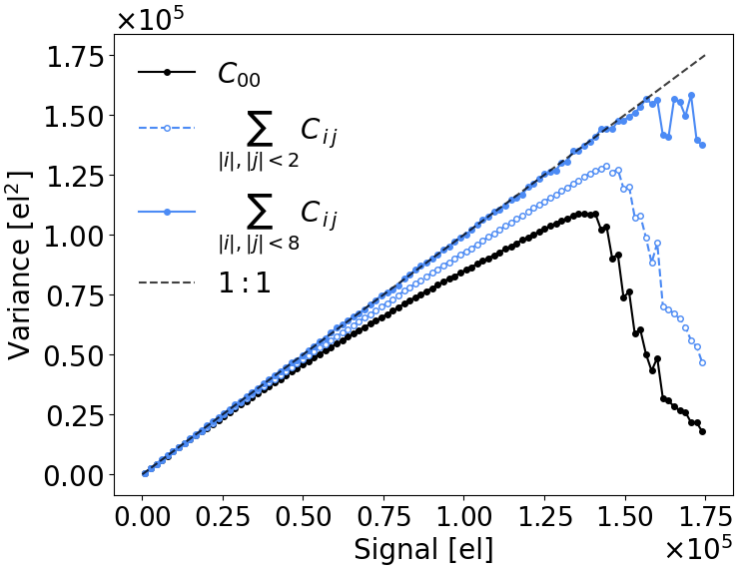


Figure 6.1: The variance and integrated covariance matrix vs. flux for a series of flat field images on an LSSTCam sensor, averaging the covariance matrix for all amplifiers. We assume that the covariance matrix has parity symmetry about the central pixel, and we sum the covariance matrix from $|i, j| < 2$ and $|i, j| < 8$, and we show that summing out to 8 pixels fully reconstructs the Poissonian behavior. We used 342 total exposure pairs with $\times 1.025$ log-spacing in signal space, which were corrected with the basic instrument signature removal defined in Appendix D. We group the data into 100 bins from 0 to 1.75×10^5 *el*. This was calculated for one sensor made by ITL on the LSST focal plane, which had an average gain of 1.702 *el* / ADU across all amplifiers.

$C_{00}(\mu)$ from a single sensor (ITL) on the LSST focal plane. If all the pixels are independent from each other, one can expect a simple linear relationship from the Poisson statistics as:

$$C_{ij}^{\text{Poisson}}(\mu) = (\delta_{i0}\delta_{j0})\mu/g + n_{ij}/g^2, \tag{6.2}$$

where δ_{ij} is Kronecker's delta, g (el/ADU) is the gain conversion between physical electrons and recorded counts (analog-digital units or ADU) by the analog-to-digital signal converter (ADC), and n_{ij} [el^2] is the electronic noise at a pixel (i, j). The measurement of C_{00} departs from the linear relationship at high signal levels as excesses due to Poisson noise are suppressed due to the BFE. Charges are deflected out of pixels with excess accumulated charge

and into neighboring pixels with less accumulated charge, resulting in a correlation between pixels at higher flat field signal levels. We measure a 30% loss in the expected variance function (photon transfer curve or PTC) near pixel saturation (Figure 6.1), which is consistent with the size of the effect observed by Astier et al. (2019), Coulton et al. (2018), Antilogus et al. (2014), Gruen et al. (2015), Guyonnet et al. (2015), Coulton et al. (2018), and Astier, Pierre & Regnault, Nicolas (2023).

If charge is conserved, the sum of all the covariances away from this central pixel recovers all the pixel-level Poissonian sensor behavior lost due to the BFE (first proposed by Downing et al. 2006). The central pixel and the immediate neighbors should contribute the most to recovering the noise budget and integrating further away from the central pixel should add vanishingly smaller contributions. Figure 6.1 also shows that accounting for the covariances 8px away from $(0, 0)$ in both directions on the focal plane allows us to reconstruct the Poissonian behavior in flat field exposures within statistical fluctuations up to fluxes above which the variance begins to drop off due to pixel saturation (around 1.35×10^5 el) for most of LSSTCam's sensors), suggesting that freed charges are being transported into neighboring pixels.

6.1 Modeling pixel-area changes from flat fields

We parameterize the evolution of pixel covariances with flux using the formalism proposed by Astier et al. (2019). The presence of charges in pixels distorts the electric field boundaries for incoming charges, which can be modeled as a change to the effective areas of the pixels that changes an incoming current of charge. In the case of flat fields, where it is safe to assume that the average charge current into each pixel (I) is constant, we model the effect of some accumulated "source" charge Q_{ij} in pixel (i, j) on the charge current into a pixel $(0, 0)$, which is assumed to far from the edge of the sensor. The incoming charge current will

affected by distortions of pixels areas from a rectilinear grid, and we consider alterations that are linearly proportional to the source charge and alterations that not linearly proportional to the source charge (such as would be the case if–but not only limited to–a hypothetical evolution of the of the charge cloud in the pixel as more charge arrives). The first component is parameterized using a matrix a_{ij} of area distortions to the pixel grid, and the second component is parameterized using a matrix b_{ij} of area distortions:

$$\dot{Q}_{00} = I[1 + \sum_{kl} a_{kl} Q_{kl}(1 + b_{kl}\ I\ t)]. \tag{6.3}$$

This equation follows the convention that the variable t is the time since the start of a flat field exposure at $t = 0$. The a_{ij} and b_{ij} matrices are scalars, meaning that each element of the coefficients is a single number that represents the net fractional change in area per unit charge, to pixel (i, j). In vector-based approaches, such as Antilogus et al. (2014), they separate the area shifts for each boundary of a pixel ($X \in$ [N, S, E, W]), and in this scalar-based approach we use $a_{ij} = \sum_X a^X_{ij}$. Equation 6.3 assumes parity symmetry such that $a^X_{i,j} = a^{-X}_{-i,-j}$ and translational invariance over the sensor, and of course similar assumptions apply for the b_{ij} coefficient. A value of $a_{ij} = 0$ means that there are no distortions to the pixel, and the charge current into pixel $(0, 0)$ is just equal to the influx of light. If $b_{ij} = 0$, then then the alteration of the current into pixel $(0, 0)$ from pixel (i, j) is simply proportional the source charge in pixel (i, j). Note that this does not assume that linear coefficient a_{ij} itself does not also vary with time. Also note that–although this equation does not explicitly require it–enforcing charge conservation naturally produces a sum rule ($\sum_{ij} a_{ij} b_{ij} = 0$ or $\sum_{ij} a_{ij} = 0$ when $b_{ij} = 0$).

Using the definition of covariance, this will ultimately affect the measured pixel covari-

ances as:

$$\begin{aligned}\dot{C}_{ij} &= \delta_{i0}\delta_{j0}I + \mathrm{Cov}[\dot{Q}_{00}, Q_{ij}] + \mathrm{Cov}[Q_{00}, \dot{Q}_{ij}] \\ &= \delta_{i0}\delta_{j0}I + 2I\sum_{kl} a_{kl}(1 + b_{kl}It)C_{i-k,j-l}\end{aligned} \tag{6.4}$$

To find the measured covariances after a specific amount of exposure time with a constant input flux and charge current (flat field), we integrate $\dot{C}_{ij}$ from $t = 0$ to $t = t$, where $\mu = \int_0^t I dt = It$ for a constant average flux of illumination. This is then expanded out into a Taylor series around $\mu = 0$. For covariances at a given signal level μ, both expressed in ADU:

$$\begin{aligned}C_{ij}(\mu) = \frac{\mu}{g}\Bigg[\delta_{i0}\delta_{j0} + a_{ij}\mu g + \frac{2}{3}[\mathbf{a}\otimes\mathbf{a} + \mathbf{ab}]_{ij}(\mu g)^2 \\ + \frac{1}{6}[2\mathbf{a}\otimes\mathbf{a}\otimes\mathbf{a} + 5\mathbf{a}\otimes\mathbf{ab}]_{ij}(\mu g)^3 + \cdots\Bigg] + n_{ij}/g^2,\end{aligned} \tag{6.5}$$

where a_{ij} (1/el) describes the fractional change in pixel area due to changes in the pixel boundaries from accumulated source charges, b_{ij} (1/el) describes smaller time-dependent processes that weaken or strengthen the BFE as we will explore in a later section, μ is a given mean signal over the image region, g is the sensor gain (el/ADU), $\mathbf{ab}$ is an element-wise matrix multiplication, and $\otimes$ refers to a discrete convolution operation. We have also added n_{ij}, which is a constant term that represents the mean square electronic noise. Typically, $n_{00} \sim 5.7\mathrm{el}^2$, and the non-(0,0) terms of the noise matrix are typically zero, but they could be contributed to by the readout electronics and overscan. We will refer to Equation 6.1 as the *covariance model* in the rest of the text.

This expansion allows for different types of BFE mechanisms: first order effects, which grow linearly with signal level ($\propto \mu^1$), and higher-order effects, which grow non-linearly with signal level ($\propto \mu^{2+}$). We will refer to the contribution of these non-linear terms as "non-

linear" effects, although there may yet be other physical effects–possibly related to BF–that also arise non-linearly with signal level, which might not be included in this model, for which we adopt the general term "higher-order" effects (we will discuss this further in §10).

In the following sections, we will parameterize the magnitudes of these different effects by the deviation of the covariances from a perfect Poissonian relationship:

$$\widetilde{C}_{ij} \equiv C_{ij} - \left(\frac{\mu}{g} \delta_{i0} \delta_{j0} + \frac{n_{ij}}{g^2} \right). \tag{6.6}$$

The impact of the BFE on flat field covariances is the expected Poissonian behavior modified by the change in pixel area so that $\widetilde{C}_{ij}$ / μ^2 is a useful quantity that models the *effective fractional area displacement matrix*. First order effects dominate at low signal levels, and in the limit of zero signal $\widetilde{C}_{ij}$ / $\mu^2 \rightarrow a_{ij}$. Tn practice, we infer the strength of these components by fitting the covariance model with these coefficients to our measured covariances as a function of flux.

6.2 Deriving a scalar correction with non-linear BFEs

We can derive the amplitude of the BFE from the measured pixel covariances in flat fields, and we can use this measurement to know how to redistribute the charge in a measured image to recover the true image. The deflection field from a collected charge changes the pixel boundary and therefore the total area of a pixel. Coulton et al. (2018) assumes that the deflection field caused by a charge can be described by the gradient of a unitless, 2D scalar kernel K. In analogy with electrostatics, they assume the effect of the kernel on the covariance matrix takes the form of Poisson's equation with the Laplacian of the kernel

(Equation 19 of Coulton et al. 2018):

$$\widetilde{C}_{ij} = -\mu^2 \nabla^2 K_{ij} \tag{6.7}$$

Thus the covariance matrix is simply Poisson noise modified by the change in area. The true image can therefore be recovered from the measured image using Equation 22 of Coulton et al. (2018):

$$\hat{F}(\mathbf{x}) = F(\mathbf{x}) + \frac{1}{2}\nabla \cdot [F(\mathbf{x})\nabla \left(K \otimes F(\mathbf{x})\right)] \tag{6.8}$$

where $\hat{F}$ is the true image and F is the measured image, and $\mathbf{x}$ runs over the pixel space in two dimensions. The expression includes a factor of 1/2 to average the BFE over the duration of the exposure as the first charge to enter the pixel experiences no BFE and the last charge is assumed to experience twice the average effect.

As seen by the form of Equation 6.7, the kernel is derived from the covariance matrix at a given signal level, therefore the correction explicitly assumes that the kernel does not evolve as a function of flux. However, this is inconsistent with non-zero non-linear terms in the covariance model. Astier et al. (2019) showed that an accurate fit to the PTC requires using their functional form (Equation 6.1) up to $O(\mu^4)$, citing $\chi^2/\mathrm{N_{dof}} = 1.23$ (note the number of observations is the number of flat pairs, and the parameters for just the variance element of the fit are g, n_{00}, and a_{00} so that $\mathrm{N_{dof}} = \mathrm{N(flat\ field\ pairs)} - 3$). This would suggests that the PTC shape is contributed by other non-linear terms included in the covariance model, which would cause the Coulton et al. (2018) algorithm to break down when applied to flat images at signal levels that are not equal to the exact signal level the kernel was derived from.

If the goal is to reconstruct the Poisson fluctuations in flat fields, the strength of the BFE can be derived from the measured flat fields at a specific signal level and at least theoretically

be nearly perfectly corrected from the data. Perhaps more realistically however, we need to correct the effect out of images with much larger pixel contrasts, such as images of stars, galaxies, and other scientifically interesting objects that we want to measure accurately. This is difficult because knowing how to include the higher-order BFEs in a correction or somehow weighting the correction toward higher signal levels (where the BFE correction might be most important) becomes ambiguous, as there is no special signal level at which to sample the pixel correlations and construct a correction. Furthermore, the maximum charge capacity of CCDs are not well-defined in practice or in literature. Initially, one might think to simply solve Poisson's equation for each whole image to fully correct the effect, but this is computationally expensive and impractical for the size of images that cosmological surveys like LSST will produce. Typically, surveys might derive the amplitude of the effect at low signal or at high signal or infer the effect from some average signal level (Jurić et al. 2017; Bosch et al. 2019), however this is itself arbitrary and in the worst case can bias the correction toward objects that are correspondingly dim, bright, or neither somewhere in-between.

In order to construct a correction from flat field statistics in an unbiased way, it is important to measure the signal range over which a given correction algorithm can reconstruct the BFE. Correcting the BFE in flat field images allows us to directly test the impact of higher-order effects at different signal levels. And correcting artificial star images will crucially allow us to test how well a given correction algorithm can reconstruct realistic PSFs at those signal levels. In addition, we expect the underlying sub-pixel electric fields and higher-order physics is different in flat images and star images. Comparing the impact of a given correction algorithm on both flat field images and star field images is therefore a direct test of the validity of the underlying assumptions of a given correction methodology.

Chapter 7

Data

Our measurements of the BFE are derived from datasets taken during the fifth electro-optical testing period for the LSST Camera (informally referred to as Run 5), which began on the integrated camera focal plane at the SLAC National Accelerator Laboratory in December 2021. In this section we describe the laboratory setup, data acquisition, and analysis methodology used to characterize the BFE in the LSSTCam. This includes the configurations of the instrument testbed, acquisitions of calibration datasets and artificial star data, image processing pipeline, and photometry.

7.1 Laboratory setup

We used the Bench for Optical Testing (BOT) as described in Newbry et al. (2018). This test bench envelops the LSST Camera cryostat in a dark box that suppresses ambient light to a level less than 0.01 electrons per pixel per second and includes a rig to swap different optical projectors for illuminating the focal plane with various light patterns. This includes a *spot grid projector*, which projects a 49×49 grid of stars (approximately 3cm $\times$ 3cm on the focal plane and approximately 65 pixels of rectilinear spacing between spots) to mimic

a realistic star field of various brightnesses, shapes, sizes, and positions (see Figure 7.1). The entire projector can be moved using remotely controlled motorized stages in all three axes for dithering (XY) across the focal plane and focusing (Z). This projector consists of a Nikon 105 mm f/2.8 Al-s Micro-Nikkor lens, and an HTA Photomask photo-lithographic mask etched with pinholes. The spot grid pattern is illuminated by an integrating sphere with a 1" opening, which is fed by a 3 mW fiber-coupled light emitting diode made by QPhotonics with a narrow band output peaked at 680 nm.

7.2 Sensors tested

For this study of the BFE, we selected one top-performing sensor of each type, one ITL (R03-S12) and one E2V (R24-S11), to independently test the BF correction. The R-number refers to the raft row and column location on the focal plane and the S-number refers to the sensor row and column location within the raft. We selected these sensors based on serial Charge Transfer Inefficiency (sCTI) criterion (Snyder & Roodman 2020). Both sensors have measured sCTI below 10^{-6} for all signal levels up to 10^5 el, which is within the requirements for LSST set in O'Connor et al. (2016). We also selected two other sensors, R02-S00 (ITL) and R22-S02 (E2V), which were used only as a secondary checks to compare broader performances between sensor types which we will explain in a later section, and the results we will show and discuss in this paper come from analyses of the two "primary" sensors. Kotov et al. (2016) shows that all sensors pass the acceptance tests defined by Radeka et al. (2009).

7.3 Electro-optical (EO) datasets

The datasets we collected during Run 5 testing campaign included two types of images: calibrations and artificial science images of star fields. The calibrations are used to find,

characterize, and remove instrument signatures in images, which we test on the artificial stars images. In this section we will describe both data types in detail.

7.3.1 Calibrations

Calibration images are used to remove other instrument artifacts and behaviors from images. Among the different types of calibrations, there are: (i) bias frames: 0-sec exposure used to measure fixed-pattern electronic noise; (ii) dark frames: longer exposure images taken with the camera shutter closed used to capture the spatially coherent thermal currents fingerprint of each sensor; and (iii) flat fields: longer exposure images with spatially uniform illumination across the sensor. Our calibration set includes 20 bias images to remove fixed-pattern electronic noise, 20 dark images to eliminate thermal currents, and 342 densely sampled pairs of flat images from low flux (50 el/px) to high flux (1.75×10^5 el/px) with a log-scale increment of $\times 1.025$ in the SDSS i band regime (700–800 nm) to measure the pixel covariances as a function of flux, deferred charge, and linearity. For the flat pair acquisitions, we also have National Institute of Standards and Technology (NIST)-calibrated photodiode measurements to provide an independent accurate measurement of the throughput of the light from the illuminator in the flat field acquisition.

While some calibrations (deferred charge, non-linearity, bias structure, etc.) are independent of different wavelengths and a single calibrated image of these types can be used to correct images of any color, flat fields are known to vary with color. We note that the SDSS i-band, which we use to calibrate flat fields, roughly corresponds to the range 700 nm–800 nm which covers a region of the electromagnetic spectrum outside the wavelength of the artificial star data at 680 nm. We will discuss the chromatic dependence of the BFE in §8.2.1, however based on preliminary studies by Astier, Pierre & Regnault, Nicolas (2023) in HSC and our own later studies, any potential differences of the BFE between light at these comparatively

close wavelengths is negligible.

7.3.2 Artificial stars

We projected artificial stars onto four sensors using the spot grid projector. An example exposure with artificial stars is shown in the top panel of Figure 7.1. A close-up example of a single star is shown in the bottom panel of Figure 7.1. We project these stars onto the sensors and fix the light level by adjusting the integration time. We took a sequence of images of artificial stars at increasing intensities by varying the integration time in 15 steps from 5 s to 200 s to cover the full range in flux of our flat field calibrations (the full dynamic range of our CCDs), with 40 images at each exposure level. We split these image acquisitions into four quadrants around each sensor, changing the projector position 4 times across the whole imaging sequence. This allows us to measure many similarly sized and shaped artificial stars at a variety of positions and signal levels. The projector tended to shake briefly after each change in its position, so we ignore the first image after the projector changes position to avoid images with distorted stars that could bias a consistent measurement of the size and shape of the stars.

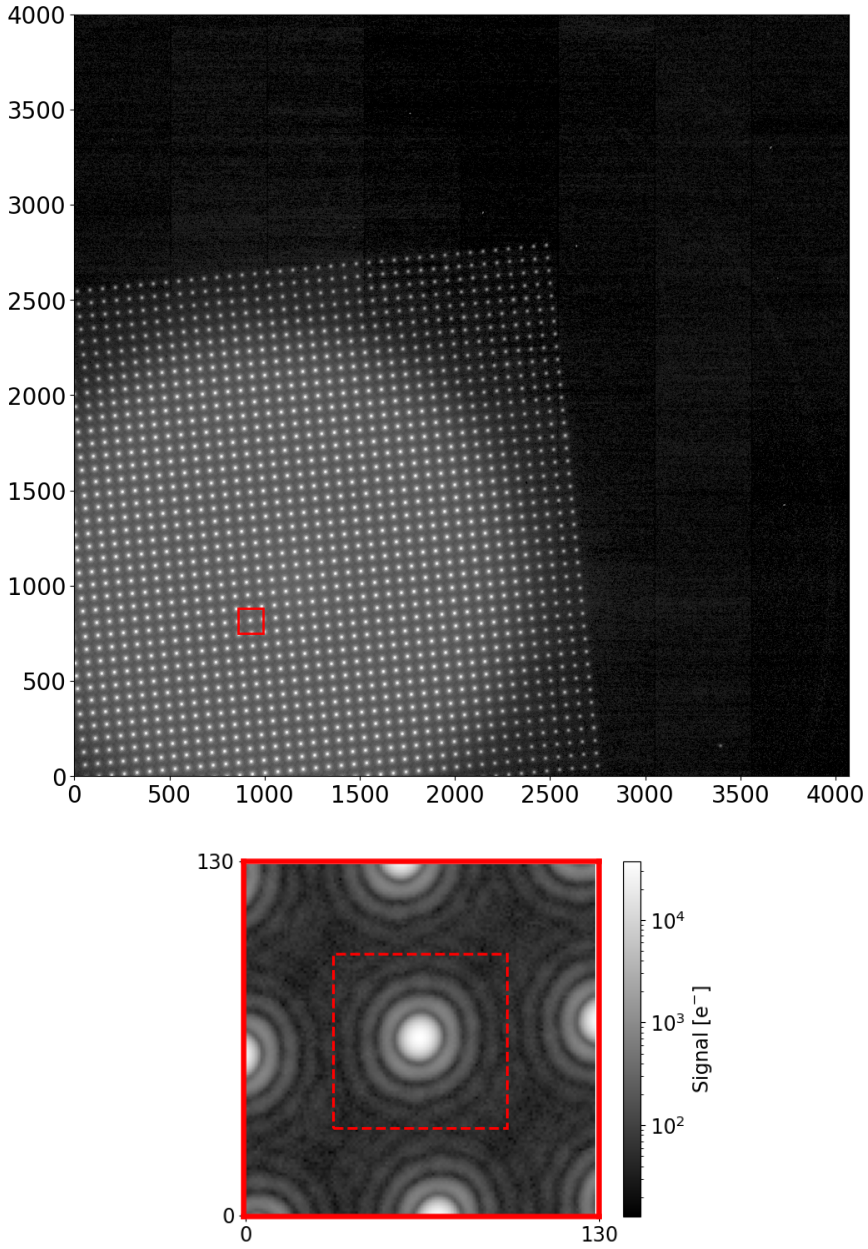


Figure 7.1: Top: an example of a spot image (20 s exposure with 680 nm LED) that has been processed by standard instrument signature removal discussed in §7.4.1. It shows the uneven (circularly symmetric) illumination pattern from the projector on a single sensor (ITL R03-S12). Bottom: an image subsection showing the Airy function due to the long wavelength regime (solid red), and the corresponding 65 px × 65 px stamp (dashed-red) which sets the bounds used to derive calculate each star's shape.

7.4 Image processing

Before any image analysis can take place, the images need to be cleaned of other instrument artifacts that are not related to the BFE, and we need to be able to measure the sources in the image and compile their useful properties into a catalog. In this section we give an overview of the methods we used to process these images (§7.4.1) and measure their sources (§7.4.2) in preparation for our analysis of the BFE.

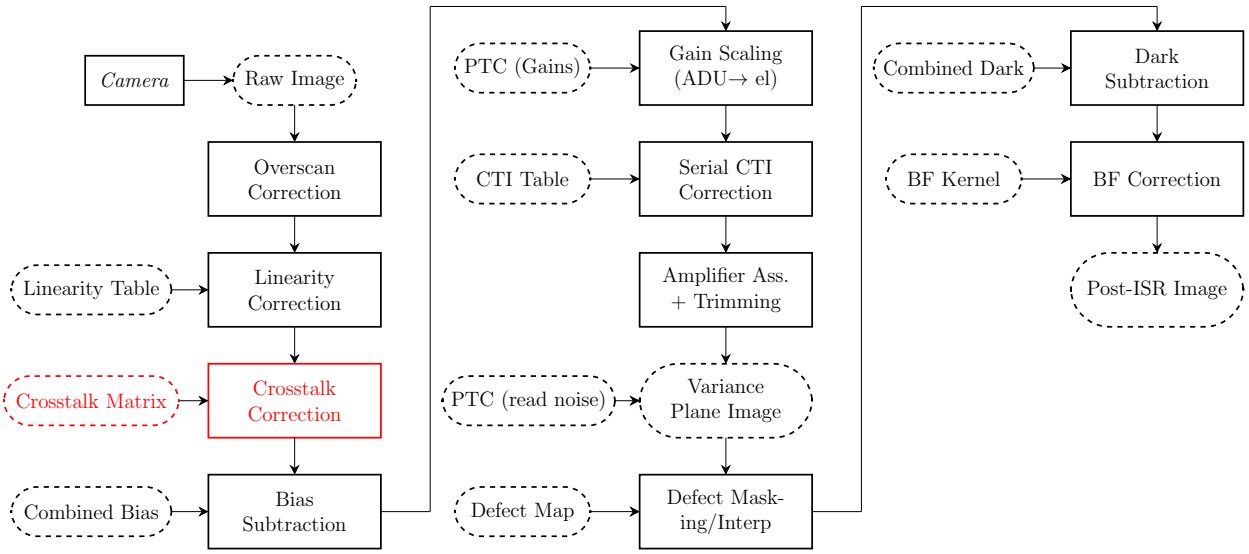


Figure 7.2: Simplified overview of Instrument Signature Removal (ISR) steps and necessary calibration products. The solid, square boxes are algorithmic steps ("tasks") in the pipeline. Dashed clouds are specific input/output data products. The steps in red are in the process of being implemented/or not applied in this analysis framework. The final data product here are the calibrated images ("Post-ISR Image"), and the other downstream steps (out-of-scope, not shown) include the measurement and analysis steps (including object detection). This flowchart is based on discussions at Rubin Calibration Workshop at SLAC, October 2023 (referenced here Plazas Malagón et al. 2024).

7.4.1 Instrument Signature Removal (ISR)

We processed each raw image using a complete sequence of the standard LSST Science Pipelines (w.2023.29 releases) and our calibrated Run 5 testing data products (Jurić et al.

2017; Bosch et al. 2019). Raw images are the unassembled images in analog-to-digital (ADU) units, and the output of the image processing pipeline are assembled, calibrated images in units of electrons. The full pipeline from raw data to a calibrated image is shown in Figure 7.2 alongside the necessary calibration products for each step. The standard Instrument Signature Removal (ISR) removes spatially characteristic systematic effects in the sensors using calibrated measurements and corrections. We do not apply all these steps to this analysis. The selected steps applied in this analysis, as well as the detailed configurations used in the software are given in Appendix D. The ISR steps correct different artifacts generally as follows:

1. Serial/parallel overscan: the overscan is the recorded part of an image that is not exposed to light (extra readouts along rows/columns of the sensor beyond the size of each imaging region), and it directly records row-wise/column-wise variations due to unstable bias levels across an image during readout. These variations can be measured and optionally fit and removed from each row/column.

2. Non-linearity (NL): a characteristic of the signal chain that results in the total incident signal on a sensor is no longer being proportional the recorded signal. We correct analog-to-digital converter (ADC) non-linearity (Downing et al. 2006) using a calibration correction informed by the densely sampled flat pairs compared with an independent measure of flux by the illuminator's photodiode measurements.

3. Crosstalk: electronic coupling between sensor channel readout electronics which creates ghost copies of bright sources within the same sensor and between sensors. It is a particular issue for LSST because it requires requires a fast synchronous readout of all 3024 video channels (the full focal plane needs reads out in under 2 seconds Roodman et al. 2018). We remove the features using a matrix of coefficients that captures the coupling between all combinations of source channels and target channels.

4. Bias: 2D structure that imprints on an image even without any integration time, such as spatially coherent electronic noise in the CCD and readout, can be removed by subtracting clipped per-pixel mean of many bias frames.

5. Charge Transfer Inefficiency (CTI): also referred to as deferred charge, it is the effect of the incomplete shifting of charge toward the readout amplifiers, resulting in charge getting trapped in silicon and released at a later time. We corrected the charge transfer inefficiency in the horizontal (serial) direction across the image using the same method as described in Snyder & Roodman (2020) and Astier et al. (2019).

6. Defects: pixels with outlier values in darks or flats, reside in a particular location on a sensor, or have a large value above a defined threshold can be flagged as BAD, SUSPECT, or SATURATED and can be ignored or interpolated over in downstream processing (e.g. hot columns, dark pixels, etc.).

7. Dark: 2D structure that imprints on an image even resulting from spatially coherent thermal currents. It can be removed by subtracting clipped per-pixel mean of many dark frames.

8. Brighter-Fatter Effect: discussed in this analysis.

During image processing, instrument artifacts are generally removed in the reverse order in which they occur in the signal chain. The BFE is one of the first instrument signatures to impact an exposure, and other instrument effects get mixed in with the image at later points in the signal chain. In order to isolate the effects of the BFE, these other effects need to be removed first, and the BFE correction will then applied near the end of the ISR pipeline. Extreme care must therefore be taken to remove these other effects and validate the quality of these other ISR steps before deriving calibrations for the BFE from flat fields and applying the correction to an image.

7.4.2 Photometry

We measured the light profile and shape statistics of each artificial star from each post-ISR exposure, deriving the source centroid, flux, and shape using the `Galsim` (Rowe et al. 2014) re-implementations of the Gaussian-weighted adaptive moments algorithm (HSM) defined by Hirata & Seljak (2003) and tested independently by Mandelbaum et al. (2005). The HSM photometrics were derived for a square stamp around each star's centroid, with a side length equal to the spot grid spacing, which should entirely capture the star and any diffraction rings without presuming anything about the stars' shapes (see the bottom of Figure 7.1). The metrics include the diagonal elements of the quadrupole image moment tensor (here we denote them I_{xx}, I_{yy}, and I_{xy}) where x is the serial direction along the orientation of the gates and y is the parallel direction along the orientation of the channel stops. These statistics are useful because they are directly related to metrics of cosmic shear distortions. More information about the calculation of these metrics can be found in Appendix E.

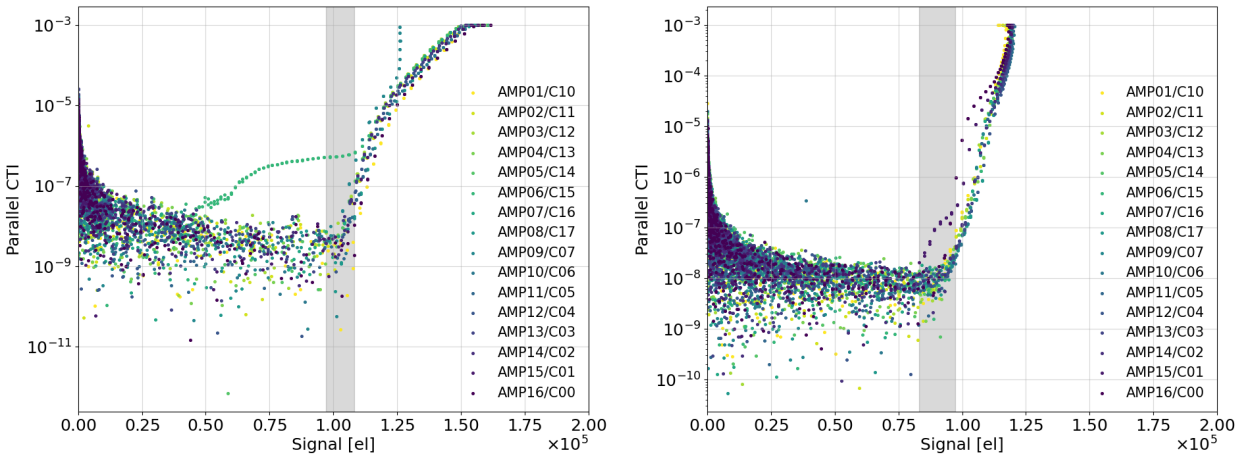
Chapter 8

Measurements of BFEs in LSSTCam

Using these calibration datasets, we can infer the properties of teh BFE in our sensors by fitting measured covariances to the covariance model. The terms of the covariance model represent the underlying mechanisms that affect charge transport in pixels as charge accumulates in those pixels. The full covariance model does not *a priori* assume the origins of these physical mechanisms, however in this section we will fit Equation 6.1 to our measured covariances and outline the possible underlying BF mechanisms that affect correlations between pixels linearly and non-linearly with signal level.

8.1 Pixel saturation effects

Above a specific empirical signal level, the pixel is considered saturated and other mechanisms besides transverse coulomb forces begin to dominate and produce large pixel correlations that grow rapidly with each additional charge that is accumulated. Before we can infer the magnitudes of the coefficients of our covariance model, we must first consider the signal range over which we can fit the covariance model to measured PTC data. Previous literature surrounding CCDs might refer to pixel saturation despite it being somewhat ambiguous, as



(a) ITL Sensor (R03-S12)

(b) E2V Sensor (R24-S11)

Figure 8.1: Measured parallel transfer CTI as functions of flux for the two sensors, with the range of turnoffs among the 16 channels highlighted. The anomalous amplifier in sensor R03-S12 (AMP06/C15) was ignored for all of our analyses due to abnormal pCTI behavior.

different mechanisms can influence pixels to varying degrees and at different signal levels.

We observe a correspondence between rapid changes of the covariance function with different metrics of pixel saturation or "turnoffs" which we further correlate with physical mechanisms of charge transport that result in correlations between pixels. Typically, the pixel saturation level is defined as the level above which the flat images dramatically lose variance ($\mathrm{d}C_{00}/\mathrm{d}\mu < 0$) and the PTC no longer monotonically increases (the "PTC turnoff" as described by Janesick 2001). In practice, we might also measure the maximum capacity more simply as the brightest recorded pixel ("maximum observed signal"). Another method defined by Snyder & Roodman (2020) assigns the limit as the flux above which charge can no longer be effectively transferred serially out of the sensor during readout. The serial charge transfer efficiency (sCTE or rather sCTI in the case of charge transfer in-efficiency described in Rhodes et al. 2010) is calculated using the extended pixel edge response (EPER) method:

$$\mathrm{CTI}(\mu) = \frac{F_{overscan}(\mu)}{N_T F_{lastpixel}(\mu)} \tag{8.1}$$

where $F_{overscan}/F_{lastpixel}$ is the ratio between the total charge left in the overscan region after an image is transferred out to the last image column and N_T is the number of pixel transfers during that readout, which is calculated for a specific flat image at a given flux level. The critical limit ("sCTI turnoff") is defined as the signal level at which the sCTI crosses above the 10^{-5} (unitless) threshold. In our investigation, we discovered another relevant limit, which falls below all of these other limits. It can be identified as a turnoff in the parallel transfer component of CTI ("pCTI turnoff"), as shown in Figure 8.1, which is calculated analogously to the sCTI in Snyder & Roodman (2020). We define this point of pCTI turnoff by fitting the pCTI(μ) to a flat line above the level of noise (2.5×10^4 el), and below any other features (5.0×10^4 el), and determine where the pCTI deviates by more than 3σ given by the standard error of the fitted data points. We typically measure this limit to be between $0.8 - 1.1 \times 10^5$ e^-, which is below any of our other metrics for CCD full-well capacity. Note that we will refer to the parallel EPER as parallel CTI since it's calculated analogously to serial CTI, however the origin of the turnoff in parallel CTI is still unknown. It is possible that this turnoff stems from the onset of a surface channel which persists from pixels in previous exposures that are above this critical signal level.

Each of these turnoff levels represents a different source of pixel correlations that begins to be relevant at different signal levels. And while we find the turnoff levels to vary from sensor to sensor, we find that their order is consistent in all the sensors we tested. The lowest is the pCTI turnoff ($0.83 - 1.13 \times 10^5$ el), followed by the PTC turnoff, the sCTI turnoff ($1.40 - 1.75 \times 10^5$ el), and the maximum observed signal ($> 1.50 \times 10^5$ el).

8.2 Photon transfer mechanisms

We fit the covariance model up to $O(\mu^3)$ (the "full covariance model") with an 8×8px covariance matrix ($i, j = 0, 1, \cdots 7$) as a function of mean flat field signal. The different

physical effects near pixel saturation that determine the full-well capacity of pixels would bias the fit of our covariance model to the PTC, and our estimation of the strength of the BFE. We therefore limit the ranges of our fits to only signal levels below the pCTI turnoff to avoid including these other sources of pixel correlations that could be misattributed to the BFE. A summary of the fitted values for each sensor is shown in Table 8.1. In the following sections we will use this theoretical model with our lab data to compare the role of first order ($\propto \mu^1$) and non-linear ($\propto \mu^{2+}$) effects within pixels.

$C_{ij}(\mu)$ Fit Parameters

Parameter	R03-S12 (ITL)	R24-S11 (E2V)	R02-S00 (ITL)	R21-S02 (E2V)
a_{00}	-1.80×10^{-6}	-3.11×10^{-6}	-1.80×10^{-6}	-3.10×10^{-6}
a_{10}	1.18×10^{-7}	1.63×10^{-7}	1.32×10^{-7}	1.64×10^{-7}
a_{01}	2.54×10^{-7}	3.83×10^{-7}	2.62×10^{-7}	3.78×10^{-7}
$\sigma(a_{00})$ / $\lvert a_{00}\rvert$	3.42%	3.44%	7.35%	3.94%
b_{00}	-8.69×10^{-7}	-2.95×10^{-7}	-8.10×10^{-7}	-5.38×10^{-7}
b_{10}	-5.36×10^{-6}	-3.65×10^{-6}	-5.97×10^{-6}	-3.01×10^{-6}
b_{01}	4.71×10^{-8}	2.04×10^{-6}	5.42×10^{-7}	1.06×10^{-6}
$\sigma(b_{00})$ / $\lvert b_{00}\rvert$	50.8%	171%	163%	133%
$\bar{g}$	1.70	1.50	1.68	1.50

Table 8.1: Fitted parameters of the full covariance model (Equation 6.1) to calibrated data (§7.3.1). We fit up to the pCTI turnoff, which was calibrated per amplifier, and then all parameters are averaged over all amplifiers for each sensor. Sigma parameters are given as the scatter over all amplifiers.

8.2.1 1st order effects

Pixel area distortions

To first order, the charge added to a pixel alters the effective area of the pixel by a constant amount. By fitting the covariance model to the covariances as a function of flux, we can

measure the a_{ij} matrix and infer how much the areas would be perturbed by a small amount of charge as a function of distance from the central perturbing pixel. The first three pixels are detailed in Table 8.1 and Figure 8.2 shows the observed a_{ij} matrix and profile of coefficients averaged over all channels. The value of a_{00} is of order 10^{-6}, and we observe it to be nearly double the size in E2V sensors as it is in ITL sensors. The value of a_{00} varies by 3.42% across amplifiers. Small variability in amplifier-dependent calibration levels (including non-linearity, overscan correction, etc.) can result in small empirical variation in a_{00}, which should be amplifier-independent. Similar levels are reported for R24-S11 (E2V) in Table 8.1. The a_{ij} matrix also shows a notable anisotropy between the cardinal directions of the pixel

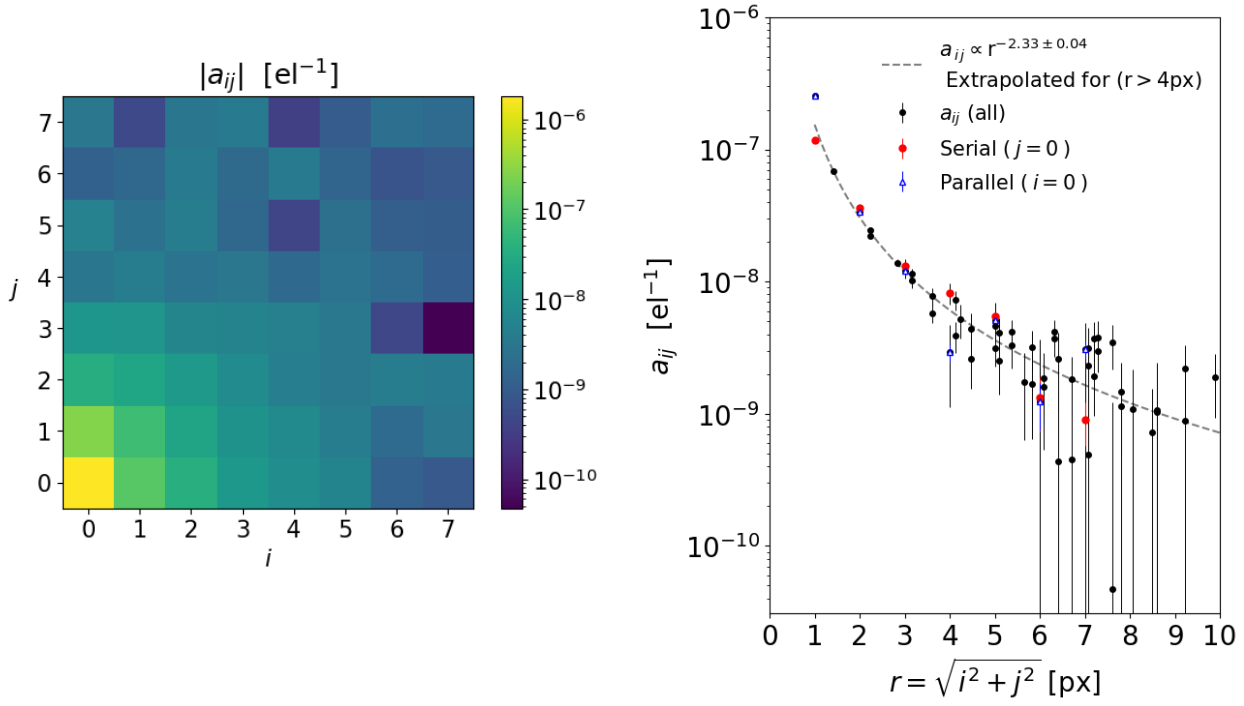


Figure 8.2: The pixel fractional area change as a matrix derived from the full covariance model (Equation 6.1). The top panel shows the asymmetry in direction ($a_{01}/a_{10} \sim 2.16$), and the bottom panel shows the decay of the matrix into noise past $3-4$ pixels. We also show the serial and parallel pixels in red and blue to show the difference in the model in both directions on the sensor. The central term is $a_{00} = -1.80 \times 10^{-6}$. Error bars are scatter across all 16 amplifiers. This was calculated for an ITL sensor on the LSST focal plane (R03-S12).

coordinate system ($a_{01}/a_{10} \sim 2.16$), indicating the BFE is stronger in the parallel direction.

Wavelength dependence

One possible additional variable is the wavelength dependence of the BFE. The depth at which a photon converts into an electron-hole pair in the sensor is determined by the wavelength of the light, and the BFE has a larger action length for charges that convert near the surface of the sensor. Figures 8.3 and 8.4 show that, for the typical wavelengths that LSSTCam will observe, the bluer wavelengths convert higher in the sensor than redder wavelengths. In fact approximately 10% of the near-infrared wavelengths pass fully through the sensor without getting absorbed. If the BFE has a wavelength dependence, one should expect to measure larger covariances in redder wavelengths than in bluer wavelengths.

Using PTC datasets from later LSSTCam testing (Run 6) taken in a similar way as the Run 5 flat pair datasets described in §7.3.1, we can compare the covariances and the fitted 1st order model coefficients between different wavelengths. The measured covariances for different wavelengths (measured in LSST filter bands and narrowband wavelengths) are shown in Figure 8.5, which shows that there are higher covariances, and therefore lower pixel correlations in the redder wavelengths. Slightly different operational settings, including bias and clocking voltages, can contribute to the discrepancy between the measured a_{00} shown in Figure 8.5 and that shown in 8.2, however we are able to compare the trend of the strength of the BFE as the evolution of a_{00} with wavelength.

This observed λ-dependence corresponds to a larger BFE at bluer wavelengths, measured by the first order components of the BFE (a_{00}). Among the bluer wavelengths, the effect is insignificant, however, at redder-wavelengths the band-to-band differences in the BF strength are more significant. One potential interpretation is that most of the deflection in the drift path of the electrons only occurs close to the charge cloud deep in the sensor, and this picture is consistent with 3D electrostatic simulations in LSST-type sensors by Lage et al. (2017).

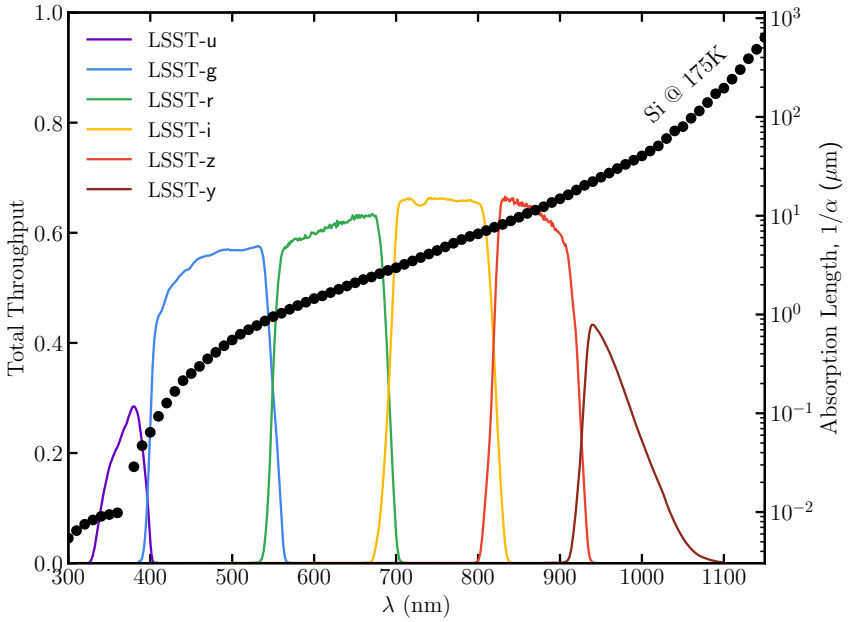


Figure 8.3: The combined hardware throughput of the LSSTCam optical system (left) compared to the absorption length in silicon (right). Silicon absorption length is taken from empirical measurements in Green (2022).

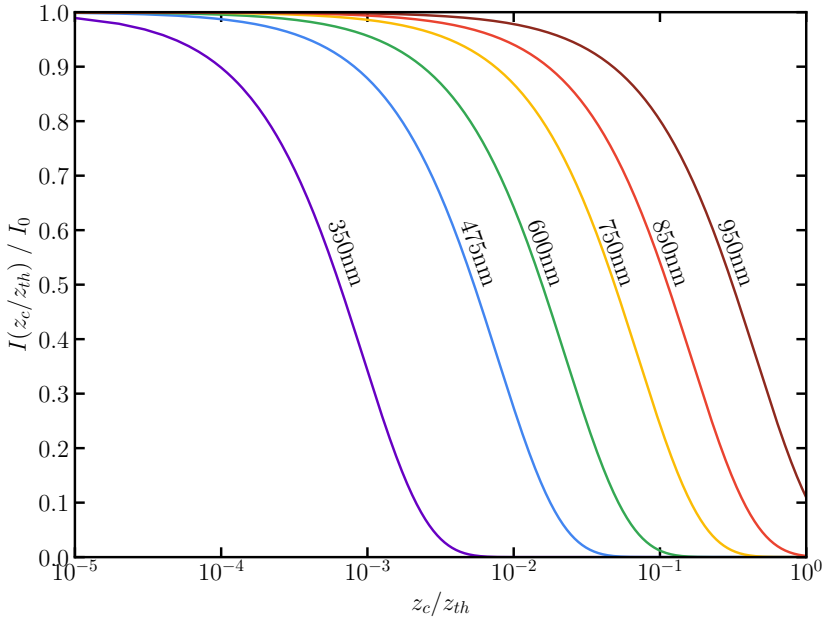


Figure 8.4: Cumulative absorption depth as a function of wavelength in silicon for LSSTCam sensors. We denote the full incident flux at each wavelength by I_0, and $I(z_c/z_{th})$ denotes the intensity of light at a given depth z_c/z_{th} in the silicon, where z_c is the conversion depth of the photon and z_{th} is the full thickness of the sensor.

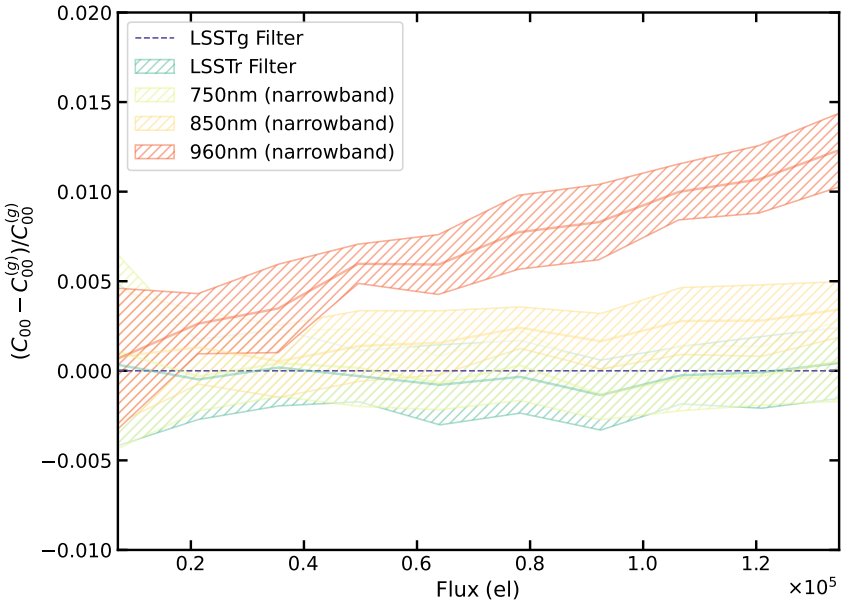


Figure 8.5: Measured covariance functions for different wavelengths. The data points are grouped into 10 flux bins, and we plot the mean and standard deviations at each bin center. Variances can add, so subtracting the g-band filter variance function ($C_{00}^{(g)}$) as a fiducial curve cancels out all other non-wavelength dependent and spatially static instrument systematics.

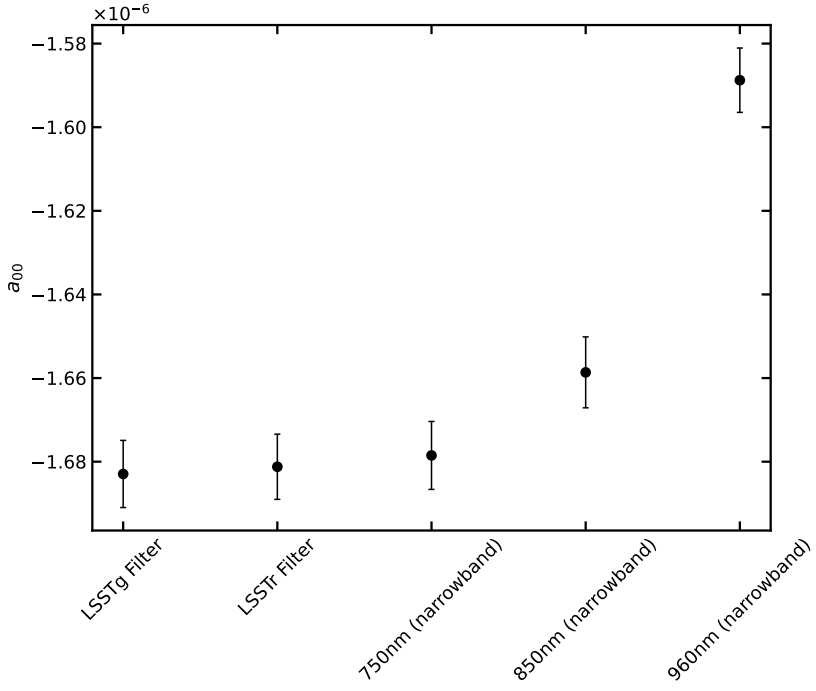


Figure 8.6: The magnitude of a_{00} at different wavelengths. The error bars represent the standard deviation across all amplifiers for this representative sensor.

8.2.2 Higher-order (HO) effects

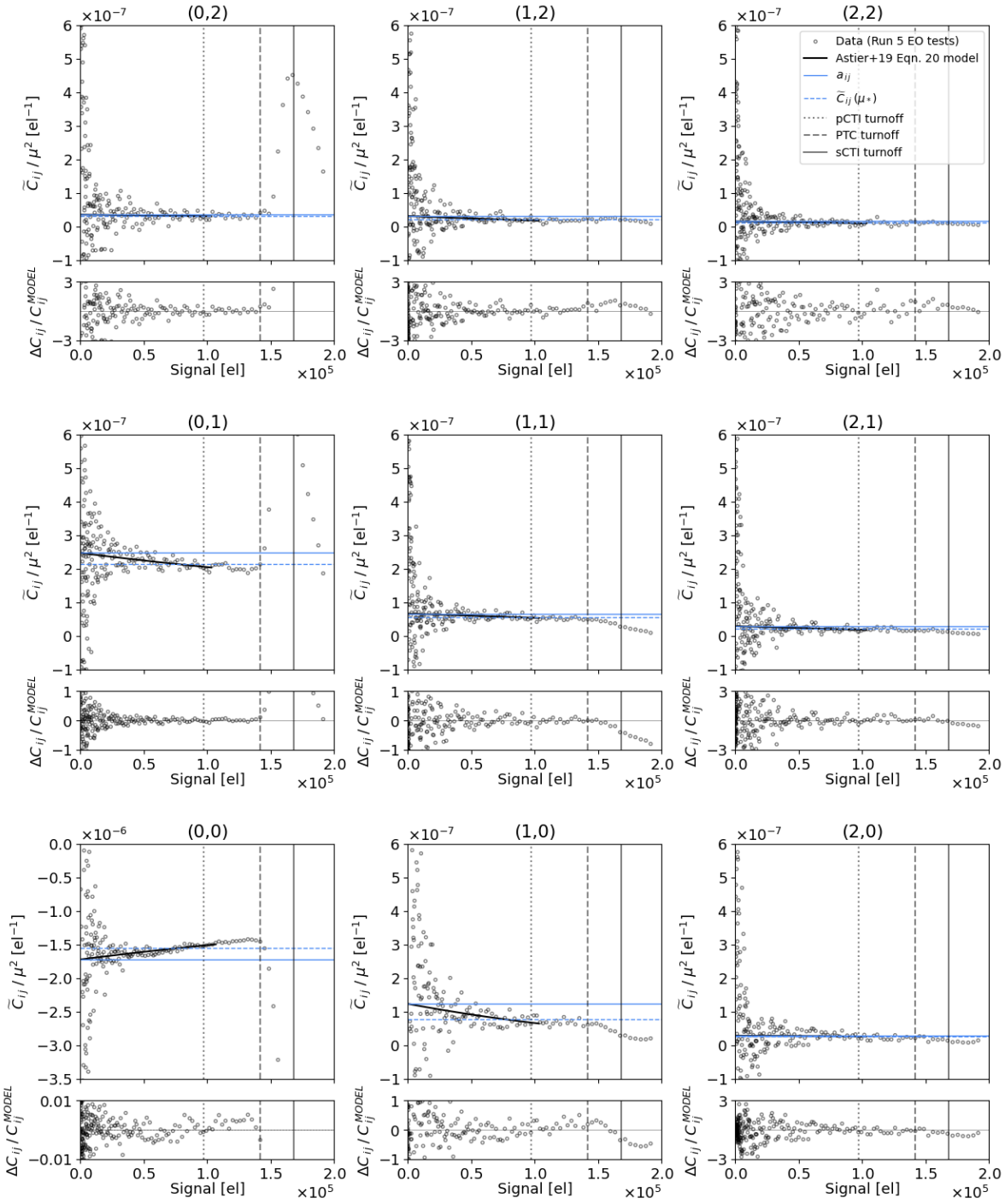
Non-linear effects

The main contributing non-linear term of the full covariance model has two components ($\mathbf{a} \otimes \mathbf{a} + \mathbf{ab}$) with two parameters ($\mathbf{a}$ and $\mathbf{b}$).

Figures 8.7 and 8.7 show that $\widetilde{C}_{ij}(\mu)/\mu^2$ deviates from the first order a_{ij} parameter for electron signal levels greater than zero, which must be due to contributions from the non-linear terms in Equation 6.1. In correspondence with Astier et al. (2019), we find that terms higher-order than μ^1 are a significant component of the covariance function, accounting for as much as 15–30% of the loss in pixel variance at 10^5 el, depending on the sensor tested, and are required for a good fit to the data. Excluding higher-order terms (μ^{4+}) in the fit contributes a negligible error, at least up to the pCTI turnoff level.

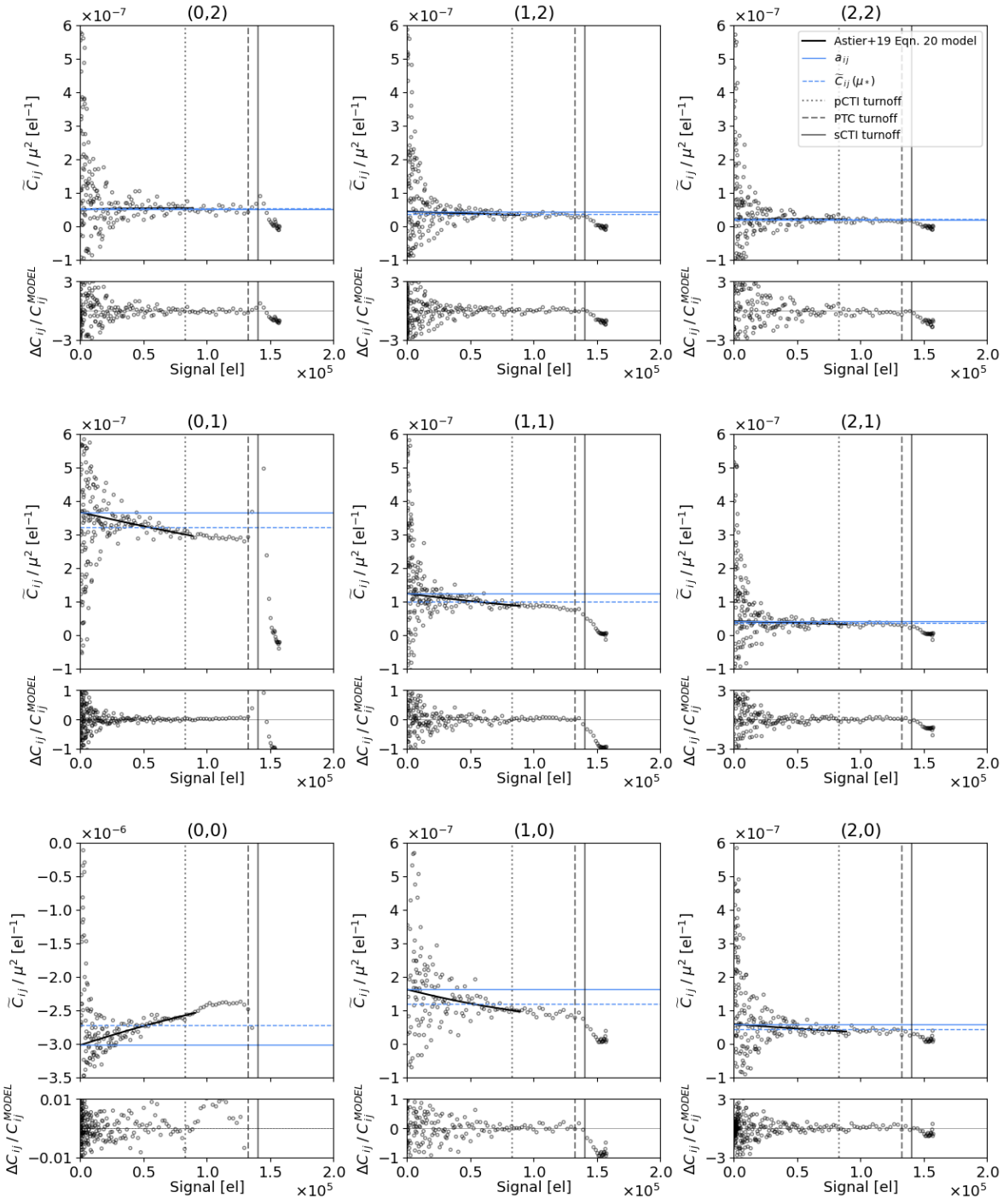
The other parameter in these higher-order terms, the b_{ij} matrix, was previously measured to be sub-dominant by Astier et al. (2019), however we find that it is almost the same size as the a_{00}. Only the b_{ij} values for $r_{ij} = \sqrt{i^2 + j^2} < 3$ are constrained by the fit to the data. Figure 8.8 shows the fitted value of the b matrix for R03-S12 (ITL). We measure $b_{00} = (-8.69 \pm 1.11) \times 10^{-7}$, $b_{10} = (-5.36 \pm 0.224) \times 10^{-6}$, and $b_{01} = (-4.71 \pm 14.5) \times 10^{-8}$, all in units of el^{-1}. The element b_{00} is 48% of the magnitude of a_{00}. The b matrix has the effect of decreasing a_{ij} if negative and increasing a_{ij} if positive, corresponding to decreasing or increasing effective pixel areas, and it has the same units as the a_{ij} matrix. It is not immediately clear how to physically interpret the value of the central b_{ij} matrix elements or other non-zero elements of b_{ij} other than as modifiers to the mathematical effect of a_{ij} that enter in as higher-order terms of $\widetilde{C}_{ij}(\mu)$ at non-zero signal levels. That is to say b_{ij} represents non-linear corrections to a_{ij} as charge accumulates.

The other non-(0,0) components of $\mathbf{b}$ are also negative except for the nearest parallel component. These elements are also asymmetric, with the best-fit measurement of $b_{10}/b_{01} =$



(a) ITL Sensor (R03-S12)

Figure 8.7: Top: measurements of the higher-order covariances and the corresponding fitted model. Bottom: the fractional residuals between the measured covariances and the corresponding covariance model. Vertical lines mark the various measures of full-well conditions.



(b) E2V Sensor (R24-S11)

Figure 8.7: Continued.

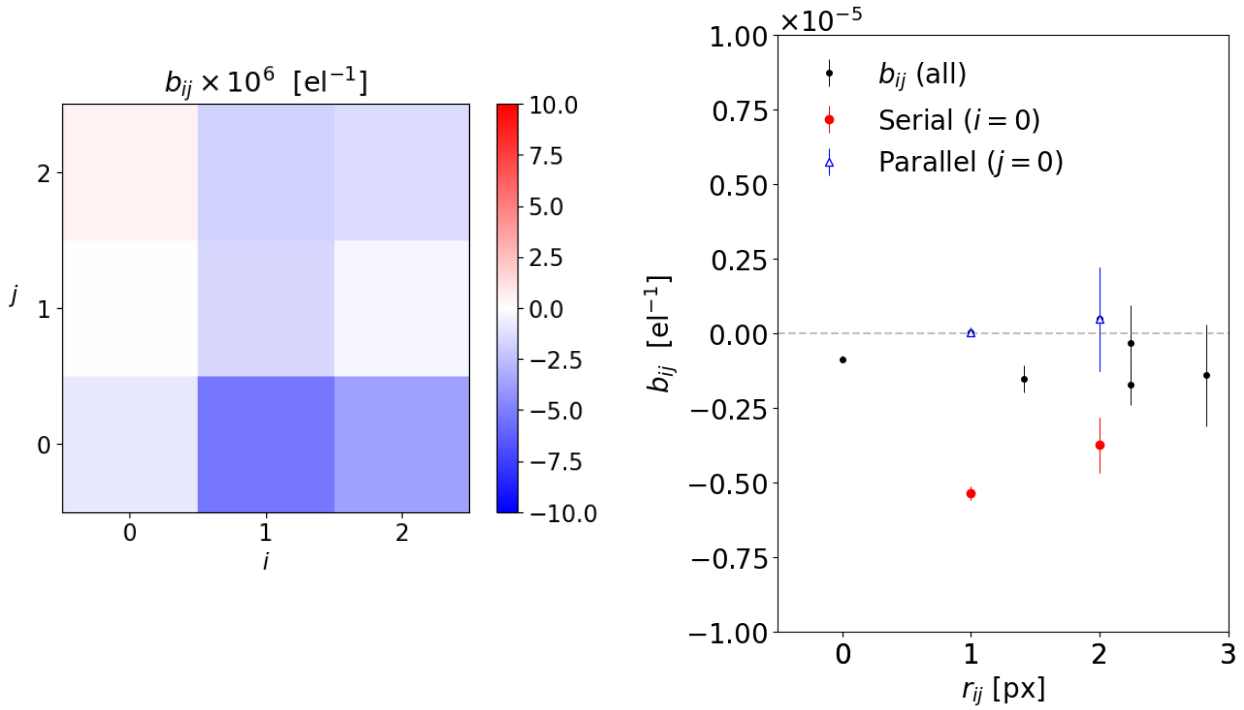


Figure 8.8: The measured b_{ij} charge-displacement component derived from the full covariance model (Equation 6.1). We also show the serial and parallel pixels in red and blue to show the difference in the model in both directions on the sensor. It shows the relative increase (positive) or decrease (negative) shift in the accumulated charges' centroids. The error bars represent the scatter over all amplifiers. This was calculated for an ITL sensor on the LSST focal plane (R03-S12).

2.82, which could result from a growing asymmetry in pixel boundary shifts with charge accumulation. As charge accumulates, if the stored charge cloud distorts anisotropically, further charges would preferentially be deflected in a particular direction, producing a non-linear charge feedback effect.

BF-induced diffusion effects

Diffusion of charge as it drifts into the collecting area of the potential well in pixels is known to exist. In the typical picture of charge diffusion, it does not produce correlations between pixels. However, several papers have hypothesized that collected charges could lower the

drift field enough to affect the drift time of incoming electrons and increase the scale of diffusion. BF-induced diffusion has been hypothesized by several works in current literature and would be expected to produce pixel correlations similar to the BFE. Magnier et al. (2018) investigated charge diffusion in Pan-STARRS 1 CCDs. The paper finds variations in charge diffusion resulting from variations in space charge density from the silicon doping structures, and they show the effect of this on observed stellar profiles, but they do this for a separate artifact called tree rings, which arise from variations in dopant concentration variations in the silicon bulk that look like rings due to how the silicon crystals are grown during the manufacturing process. The distribution of space charges that produce the tree ring effect are not the same distribution of charges that produce the BFE. And there is no formal foundation is laid out for how increased diffusion induced by the BFE would alter the variation in pixel correlations in flat fields with distance.

There have been other investigations on this topic by Holland et al. (2014). They use a simplified physics-based model to measure the impact of accumulated charges on diffusion. They measure the change in bias voltage and the diffusion scale as a function of accumulated charge for simulations of the DES's DECam sensors. They show that the effect on the diffusion scale is of the order 1% in their sensor's pixels at high signal levels. And show that the impact of accumulated charge on diffusion is even smaller when the change in bias voltage due to additional charge is much smaller than the bias voltage itself. However, their model might be over-simplified as they treat the accumulated charge as a point source, but a more realistic configuration is a more diffuse cloud of charge, which would likely result in an even smaller disturbance to the bias voltage, and an even smaller impact on diffusion. This is further supported by the fact that other works have actually described the measured covariances well by only considering alterations in charge trajectories without considering any impacts on diffusion (Lage 2019; Astier, Pierre & Regnault, Nicolas 2023).

There is a difference between the "typical" diffusion usually referred to in literature and

the "BF-induced" diffusion which has suggested. The shot noise that produces the variances in the PTC is produced by electrons independently falling into pixel bins, and typical diffusion does not change the fact that the electrons are independently falling into pixel bins, and therefore it does not produce a deviation from the Poisson variances in the PTC. However, if some source charge lowers the drift field and produces an increased diffusion-like effect, then the electrons no longer "diffuse" into neighboring pixels independently and could therefore produce a deviation from Poisson variances in the PTC. A pixel with some excess charge due to shot noise may have an increased diffusion scale, and a pixel with some smaller charge due to shot noise would have a smaller increase in diffusion scale. BF-induced diffusion would therefore affect the PTC in a way that typical diffusion does not. This BF-induced diffusion (while likely negligible) would be indistinguishable from the BFE itself and be picked up by the coefficients in the covariance model.

8.2.3 Model fit quality

A good metric of the quality of the fit is the "sum rule" of a_{ij}, which should be close to zero if the total net pixel area displacement is zero and we are fully capturing the range of the BFE with the chosen size of the a_{ij} matrix. Figure 8.9 shows the zero sum rule of the fitted **a** parameter for different sizes of the matrix in sensor R03-S12 (ITL). We find that beyond 4 px the fractional pixel area change fluctuations are largely dominated by noise, and that modeling out to 8 pixels gives $\sum a_{ij} < 1.04 \times 10^{-7}$ el^{-1}, which is 5.8% of $|a_{00}|$. Without including **b** in the fit, our sum rule is $\sum a_{ij} < 5.91 \times 10^{-9}$ el^{-1}, which is 0.345% of $|a_{00}|$. The sum rule can deviate from zero as the result of poorly constraining the noise parameter in the fit to Equation 6.1. The excess sum then results from summing up a noise offset, which dominates the covariance at large lags away from the $(0, 0)$ pixel. In the next section (§9.2) we describe how we overcome this issue and enforce the sum rule when ultimately deriving

the correction kernel.

Astier et al. (2019) showed their functional form (Equation 6.1) up to $O(\mu^4)$ fits the observed PTC well over the flux range they considered, citing $\chi^2/\mathrm{N_{dof}} = 1.23$. The number of observations is the number of flat pairs, and the parameters for just the variance element of the fit are g, n_{00}, and a_{00}, so they use $\mathrm{N_{dof}} = \mathrm{N(flat\ field\ pairs)} - 3$. Their good model fit suggests that the PTC shape is contributed by other non-linear terms included in the covariance model. Our fit to the model leaves an average $\chi^2/\mathrm{N_{dof}}$ of 13.7 in each amplifier, which is larger ($\chi^2/\mathrm{N_{dof}} = 14.9$) when $\mathbf{b} = 0$. This is large since the model fit leaves 0.5% fluctuation in the residual variance (C_{00}) above 5×10^4 ADU (bottom plots of the 00 terms in Figure 8.7). The fluctuation is characteristic of a residual un-modeled non-linearity in the Analog Signal Processing Integrated Circuit (ASPIC) (described in Astier et al. 2019;

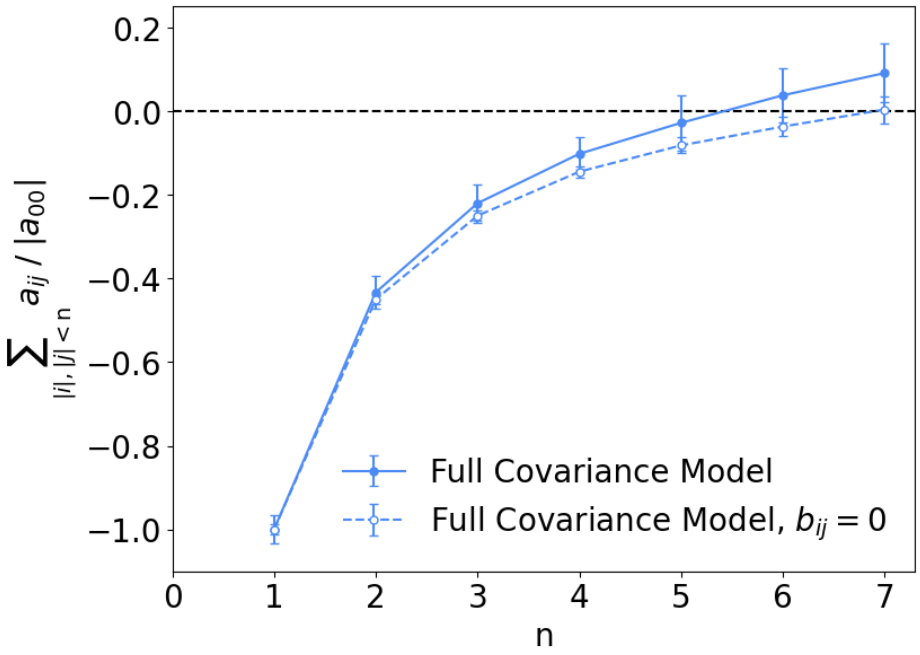


Figure 8.9: The charge conservation condition of the a_{ij} model for the cases where b_{ij} is allowed to vary or held fixed at 0. We calculate the sum over the $\mathbf{a}$ matrix. The error bars are the standard deviations across all amplifiers. This was calculated for the ITL sensor R03-S12.

Juramy et al. 2014). The non-linearity takes on the form of a regular periodic function in signal level above 7.5×10^4 el, and it might not have that much of an impact on our full covariance model if the surplus and deficit residuals balance the fit around the true PTC. However, by restricting the range of fluxes fit to those below the pCTI turnoff, we cut the fit range short enough that it might cause the non-linearity to bias the fit and give us a poor result. For this reason, the quality of the model fit is very sensitive to the flux range. We found that adjusting the maximum signal by $\pm 10^3$ el (around the pCTI turnoff), we could achieve arbitrarily better χ^2, but not necessarily a more accurate fit which models the BFE. On the other hand, the non-(0, 0) terms, which are also shown in Figures 8.7 and 8.7, exhibit good fits with residuals smaller than the level of read noise and statistical fluctuations. In the next section, when we use the PTC model to construct a BF correction, we will continue to use these PTC fits, which at least generally follow the curves of the PTCs. Regardless, we can still conclude that the higher-order terms are non-negligible and result in non-linear alterations of $\widetilde{C}_{ij}/\mu^2$ as charge distributions grow over time.

Chapter 9

Correction of BFEs in LSSTCam

In the following sections we will apply the scalar correction by Coulton et al. (2018) (§6.2) with non-linear BFEs included.

9.1 Adding non-linear corrections

As we saw in the previous section, the amplitude of the BFE is different at different signal levels due to the contribution of higher-order effects, and we know that any physically well-motivated correction algorithm would need to take higher-order terms into account at high signal levels. To include these higher-order components into K, we construct the kernel (using Equation 6.7) by sampling $\widetilde{C}_{ij}(\mu)$ from our full covariance model at a non-zero signal level, which would therefore include higher-order effects. We will do this in several steps:

1. Evaluate $\widetilde{C}_{ij}(\mu)$ from the full covariance model at several signal levels.

2. Use Equation 6.7, to calculate a kernel for each $\widetilde{C}_{ij}$.

3. Correct the flat pairs in the PTC using each kernel.

4. Calculate $\chi^2/\mathrm{N_{dof}}$ between each corrected PTC and the expected linear behavior, defined by the gain of the uncorrected PTC.

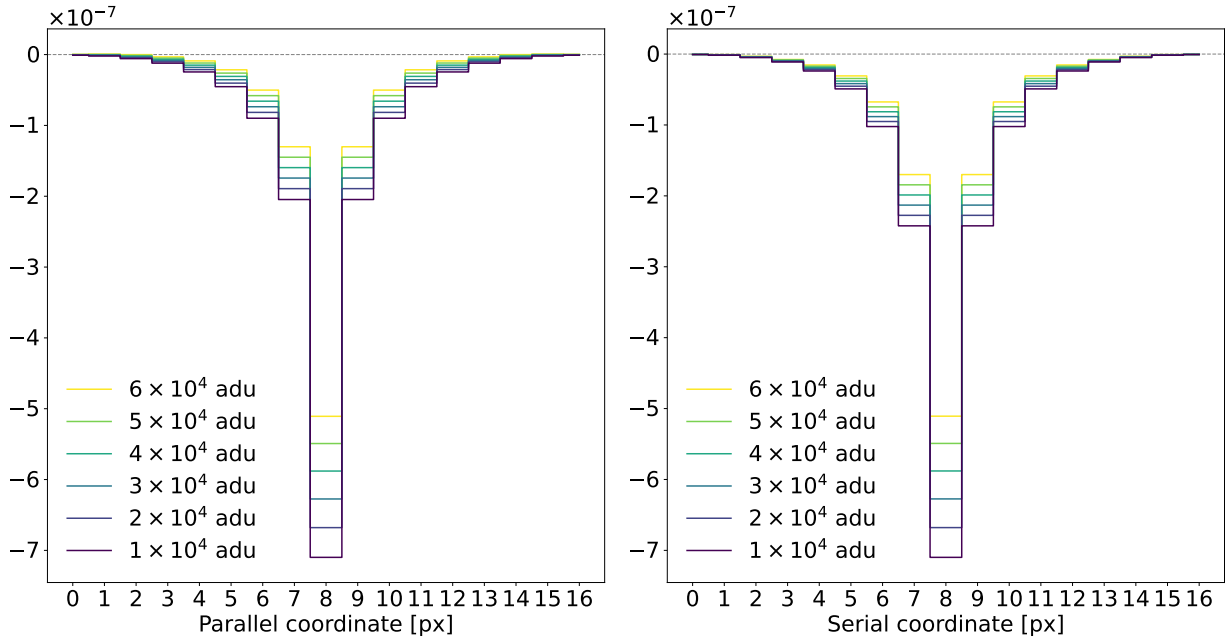
5. Determine the kernel that results in the minimum reduced χ^2 statistic.

By finding the signal level at which the corresponding kernel minimizes the reduced χ^2 between the resulting corrected PTC and the Poisson expectation, we define a characteristic signal level that can best correct the PTC. We will refer to this signal level as the "characteristic" signal level (μ_*) and the corresponding kernel as the "characteristic" kernel.

We search for μ_* over a range of $\mu \in [\,10^4\mathrm{ADU},\ \text{pCTI turnoff}\,)$. We derive corrected PTCs from kernels at six signal levels in this range, in steps of 10^4 ADU.

We also take a few practical steps to ensure each kernel is accurate. We normalized the kernel at a given signal level to enforce the zero sum rule (similar to charge conservation) by adjusting the value of the central pixel $\widetilde{C}_{00}$ so that $\sum_{ij} \widetilde{C}_{ij}/\mu^2 = 0$, for the measured correlations and the correlation model, integrating distances out to infinity. We risk over-fitting noise if we calculate the kernel boundaries beyond 4 pixels, so we model the correlations beyond 4px away from the center of the kernel with an empirically measured power law, $\widetilde{C}_{ij}/\mu^2 \propto r^{-\gamma}$ (typically $\gamma \sim 3.2$). However, we observed that inclusion of the correlation model makes only a negligible difference to the kernel and the final correction. The central term of the kernel (K_{00}) is typically decreased by approximately 10% as a result of enforcing this zero-sum rule, and K_{00} fluctuates across amplifiers by approximately 10% due to variations in calibrations between amplifiers. Enforcing this zero sum rule has a large impact on the correction since the $(0, 0)$ pixel dominates the correction amplitude as we will show in the following sections. In order to avoid discontinuities at amplifier boundaries, we average the kernels per sensor and use the average in the final correction, and these kernels are shown in Figure 9.1. Ultimately, as described and recommended in the implementation of the correction (Coulton et al. 2018), we recursively apply the kernel to each image until

the total added charge in each step falls below a threshold of 10 electrons (the "convergence condition"), which typically takes no more than 2–3 iterations.



(a) Center serial slice

(b) Center parallel slice

Figure 9.1: The average kernels calculated for ITL (R03-S12) at different signal levels. They are shown in profile as central slices along the parallel and serial directions.

We measure the quality of the correction by comparing the corrected PTC to the expected Poisson form. We calculate the reduced χ^2 as:

$$\chi^2/\mathrm{N_{dof}} = \frac{1}{\mathrm{N_{dof}}} \sum_{\mu} \frac{(C_{00}^{\mathrm{Corrected}} - C_{00}^{\mathrm{Poisson}})^2}{C_{00}^{\mathrm{Poisson}}} \tag{9.1}$$

The Poisson variance is determined using the gain derived from the covariance model fit of the uncorrected PTC. We also use $\mathrm{N_{dof}} = \mathrm{N(flat\,pairs)} - 2$ because we have N(flat pairs) observations and 2 parameters for gain and noise. We only calculate the residual χ^2-statistic from data below 5.0×10^4 el for E2V sensors and below 7.5×10^4 el for ITL sensors to avoid residual NL at higher signals from interfering with our test.

Figure 9.2 shows the residual reduced χ^2 as a function of the corresponding signal level. We perform a simple least-squares test by fitting a 3-parameter quadratic polynomial to the points across the flux space, and the location of the minimum of this quadratic polynomial determines the characteristic signal level μ_* that that minimizes the reduced χ^2 statistic for each sensor and best reconstructs the Poisson form of the PTC. The characteristic signal levels are not necessarily exactly half way between zero signal and pixel saturation at the PTC turnoff. The location of the characteristic signal level could be a balance between underestimating the role of higher-order BFEs at low signal and overestimating the role of higher-order BFEs at low signal. We also find that like-sensor types have similar characteristic signal levels within fitting errors between our 4 sensors, and while the reason is not

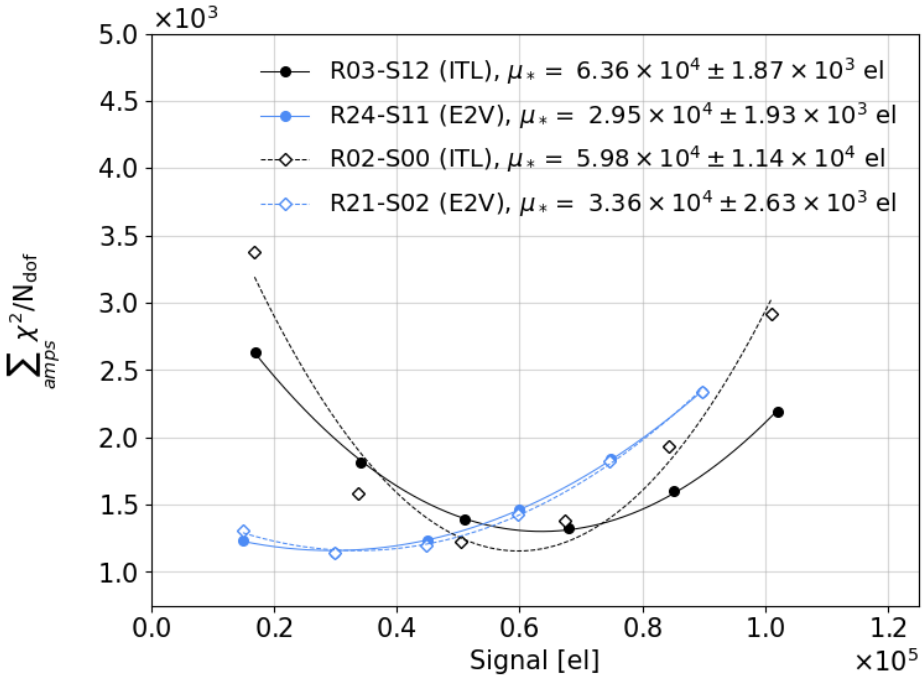


Figure 9.2: The measured reduced χ^2 values between corrected PTCs and the Poisson expectation using kernels at 6 different signal levels. We sum the reduced χ^2 over all amplifiers. A simple 3-parameter quadratic polynomial has been fit to the points across the parameter space to determine the signal level (μ_*) that minimizes the χ^2 statistic. This shows that we have a characteristic signal level that reconstructs the PTC, and this level is similar in sensors of the same type.

entirely clear, it could be the result of sensor-specific operating conditions or design specifications (e.g. similar doping levels in the substrate material as considered in Rasmussen et al. 2016).

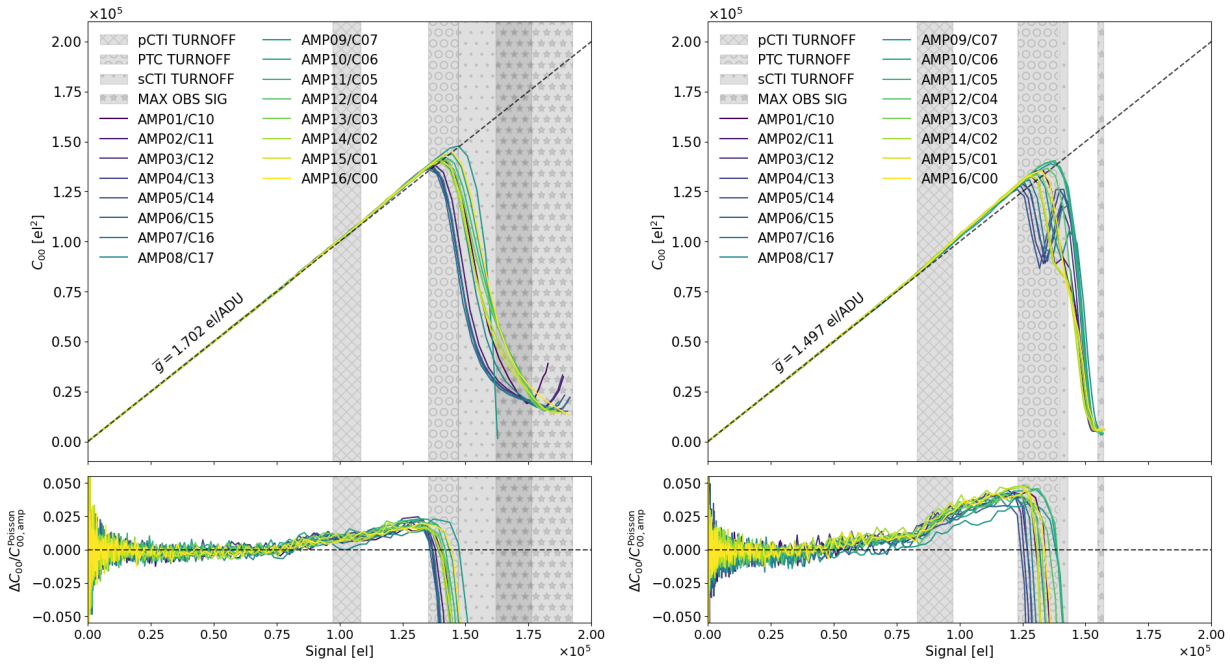
9.2 Correction of flat fields

In this section, we attempt to linearize the PTCs of our sensors by correcting the BFE out of the flat field images using the kernel K which includes higher-order BFE components. Figure 9.3 shows the ultimate result of applying the empirically determined best BF correction to the flat field images in order to reconstruct the linear PTC. The correction breaks down above the pCTI turnoff for each sensor. Since the kernels for each amplifier are averaged together, the amplifier with the lowest pCTI turnoff determines the maximum signal level that can be corrected. Above 7.5×10^4 el but below the pCTI turnoff, there is still some remaining uncorrected NL from the flat field projector, which prevents us from fully correcting the BFE above that signal level. This did not have an impact on our calculation of μ_* since we only used data points below this level in calculating our χ^2 statistic specifically to avoid it interfering with our analysis.

We fit the full covariance model (up to the pCTI turnoff) to the BF-corrected PTC for both sensors. In the ITL sensor, the average a_{00} decreased by 94.9% and the residual anisotropy was corrected closer to unity ($a_{01}/a_{10} = 0.735$). For the E2V sensor, a_{00} decreased by 97.1% and the residual anisotropy was also corrected closer to unity ($a_{01}/a_{10} = 0.942$). This shows that the BFE correction captures most of the strength and anisotropy of the BFE in flat fields below the pCTI turnoff.

We rely on the Coulton et al. (2018) algorithm that assumes the kernel is constant with respect to the flux. Our result finds this assumption is not true. Using a single kernel to correct multiple signal levels is intrinsically flawed because a single scalar kernel cannot

capture both the BFE at low signal and the BFE at high signal. However, we cannot use multiple kernels to correct images of PSF stars with more complicated light profiles. All previous literature that has investigated a correction of the BFE derived their correction from flat fields measurements. It is easily calibrated from flat fields and is far less computationally expensive than solving Poisson's equation for each image. If we use flat images to derive a correction for the BFE, it is important to find the kernel that can best reconstruct the Poisson noise of flat images and understand how well the BFE measured in flat fields can approximate the effect in star images.



(a) ITL Sensor (R03-S12)

(b) E2V Sensor (R24-S11)

Figure 9.3: Photon transfer curves reconstructed from corrected flat field images using the corresponding characteristic kernels for the ITL sensor (R03-S12) on the left and the E2V sensor (R24-S11) on the right. On the bottom, we include the residuals from the expected linear behavior. Although we use a detector averaged kernel, we compute this for each of the 16 amplifiers. The average gain, included as a parameter in the post-correction full covariance model fit, increased by 0.35% in the ITL sensor and 16% in the E2V sensor.

9.3 Correction of stars

In this section, we apply the characteristic correction derived from flat fields to correct images of artificial stars. The BFE in stars could be a stronger effect than in the earlier case of flat fields since the contrasts (pixel-to-pixel differences) are larger and therefore the gradients of the charge distribution and the divergence of the deflection fields are also larger. The BFE in stars could be a stronger effect than in the earlier case of flat fields since the contrasts (pixel-to-pixel differences) are larger and therefore the gradients of the charge distribution and the divergence of the deflection fields are also larger.

To start, we found several problematic systematic effects in the images which needed to be addressed. Firstly, there is non-uniform, disk-shaped background in the artificial star images that peaks around 0.5 el/px/s, which is caused by illumination through the photolithographic mask and evident by eye in the top image of Figure 7.1. The background light level is much smaller than the Poisson noise of each star, and it is unlikely to significantly affect our photometry of bright stars. However the shape-fitting algorithm is known to be sensitive to background light levels, especially in the low signal-to-noise (SNR) regime (Hirata & Seljak 2003; Refregier et al. 2012; Okura & Futamase 2018). We attempted to model and subtract the background in each exposure using a 65px box size and 5px median filter while masking out the footprints of the artificial stars. We attempted to remove the background in order to accurately measure the shape of the BFE at low signals, but there appeared to be a residual background that we failed to model by this procedure in the dimmest images, resulting in the increase of the second moment of stars in the low S/N region. We performed this background subtraction even though we ultimately ignored these low signal points in measuring the slope of the BFE in stars. The analyses with and without the background subtraction are not measurably different.

Secondly, we observe Airy diffraction rings on each of the stars due to the pinholes of

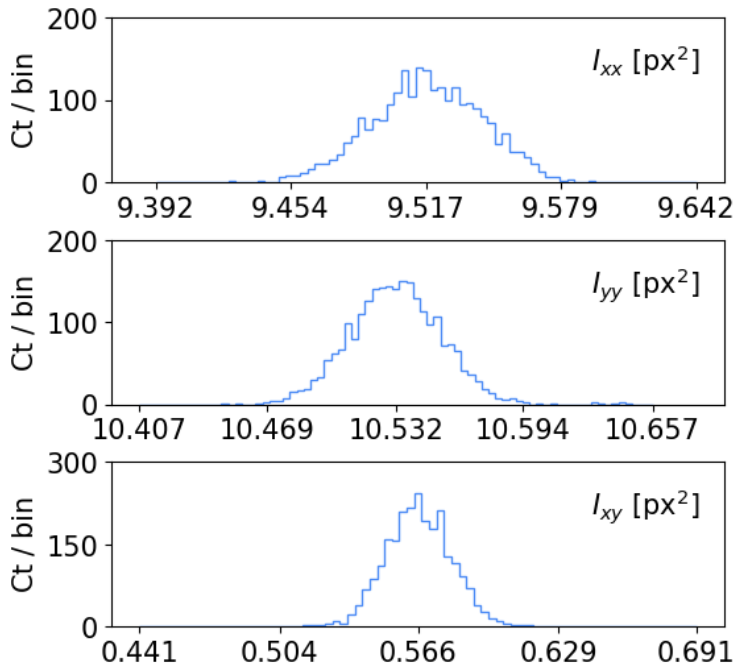


Figure 9.4: Distributions of the intrinsic shapes of the artificial stars used in these analyses (after selection cuts). Each distribution is centered on its mean and plotted with the same bin size. We derived these values from a collection of low-brightness images at the same exposure time (15 s), and these distributions contain about 3200 spots. The distributions indicate a small average ellipticity ($a/b < 0.07$) and correlated orientation angle ($\bar{\theta} = -24.1$ deg) due to the relative alignment of the spot projector with the mask and focal plane, indicating that the major axes are almost parallel with the serial gate structure. Angles are given by $\tan(2\theta) = 2I_{xy}/(I_{xx} - I_{yy})$ and measured relative to the serial transfer direction. At the zero signal level, the average shapes are $(I_{xx}, I_{yy}, I_{xy}) = (9.45, 10.45, 0.564)$ ([px^2]).

the spot projector mask, which can be seen in the bottom image of Figure 7.1. However, as we will discuss further, we will study the differences between spot sizes at any signal level with respect to the same star's intrinsic shape at low signal and the impact of the detailed structure will be cancelled out even if the diffraction rings of different spots have non-negligible overlap.

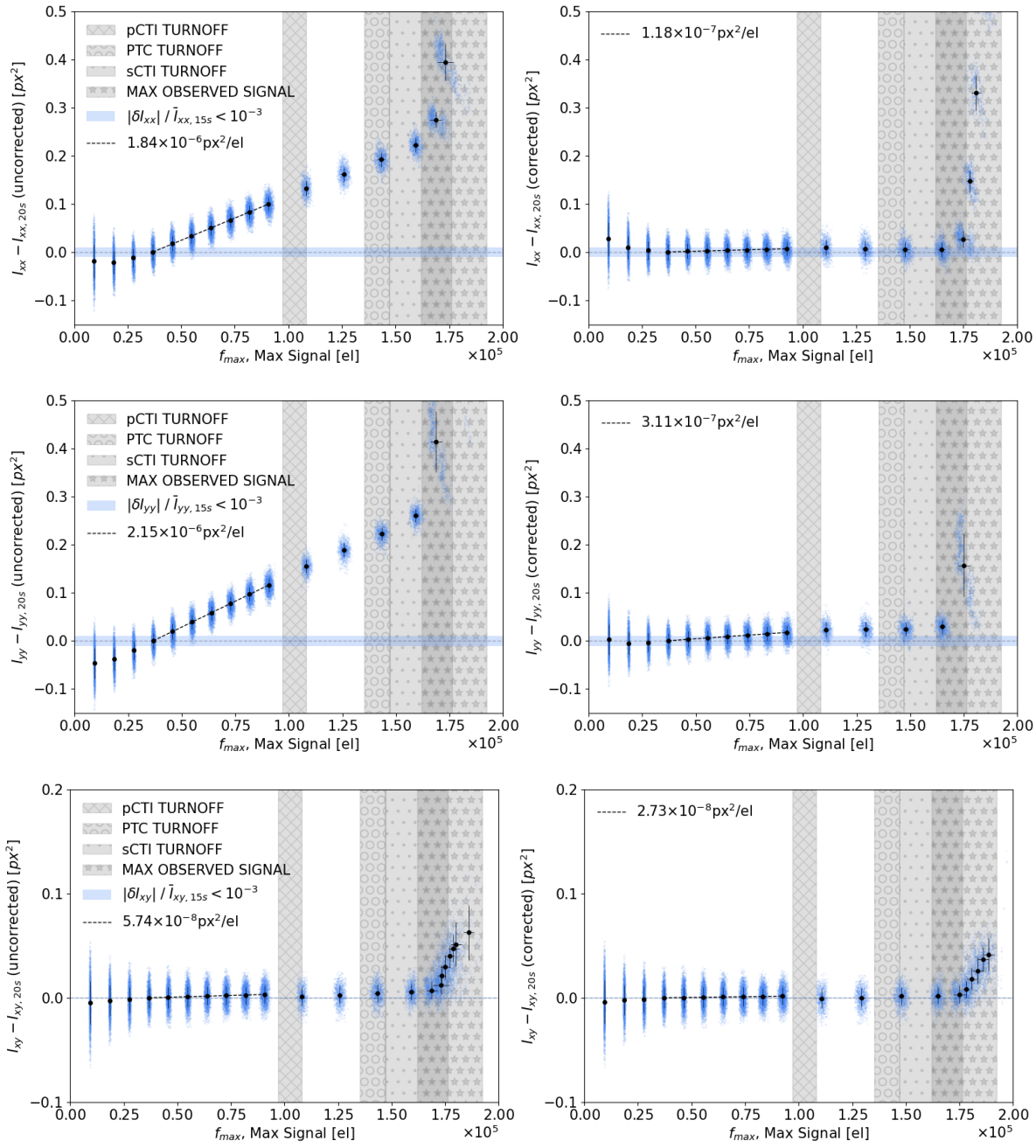
Thirdly, the focal-plane-to-projector orientation and optical aberration from the lens caused minor distortions of the artificial stars, especially away from the center of the grid, and they cannot be perfectly parameterized by isotropic 2D Gaussians. We observed that

the central and brightest stars of the grid hardly vary in intrinsic shape between exposures at the same or different projector positions. To avoid any other systematic effect related to the intrinsic alignment of the spots themselves, we therefore select only the top 5% brightest stars in total flux in each image. These stars happen to correspond to the same pinholes in the spot projector between in each image. In order to avoid including stars that land on the edges of the sensor, which skews their shapes, we also clipped stars beyond 3σ of the mean at each exposure time. This cut only removed O(10) data points in a few of the exposures. We ultimately selected approximately 80 spots in each exposure for our analysis. With 40 images at each exposure level, we use approximately 3200 spots/exposure level.

Figure 9.4 shows the approximate intrinsic shapes of the artificial stars used in this analysis, which we derive from their shapes at 15s of exposure, which corresponds to a relatively low peak signal level of 2.5×10^4el. The distortions are correlated across the grid with the major axes consistently 7% larger than the minor axes, and the major axes of all the spots are aligned at a -24 deg angle relative to the serial readout direction (horizontal and to the left in Figure 9.4). We do not believe this affects our result, as we will discuss further in §10.1.

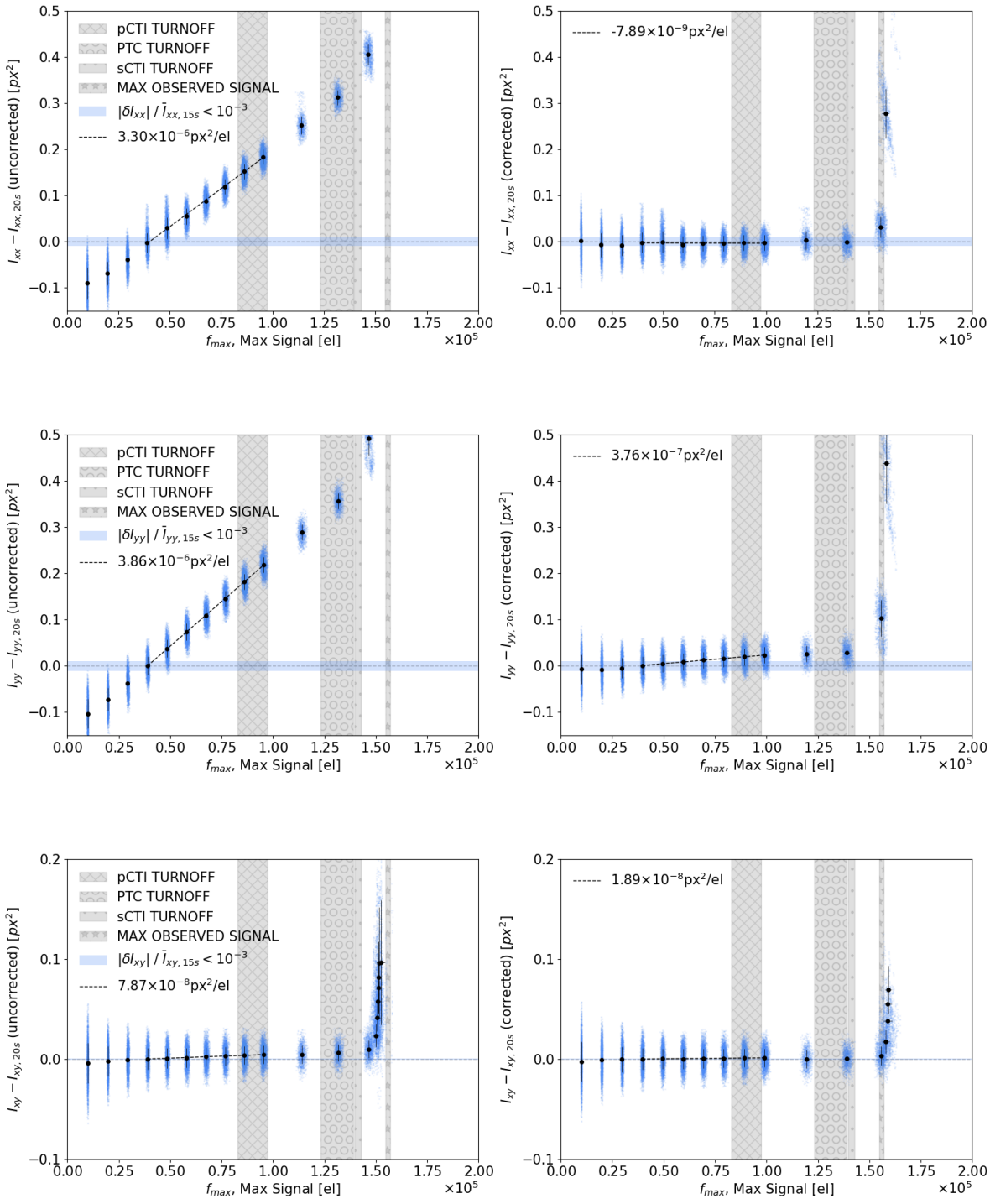
In Figure 9.5, we show the growth of all three independent components of the second-moments matrix (I_{xx}, I_{yy}, I_{xy}) separately for each sensor type. We chose the mean star shapes derived from the 20-second exposures as a fiducial point from which to show the growth of the size that result from the BFE since the lowest flux stars suffers from low signal-to-noise ratio due to the extended residual background light. We therefore measure the size-growth of each star, corresponding to an individual hole in the projector mask, relative to its intrinsic shape. We also use the peak signal of each spot instead of the integrated signal as a measure of the signal level so that we can relate our spot photometry to different measurements of the pixel full-well capacity.

To quantify the magnitude of the BFE, we determine the growth of the second moment



(a) ITL (R03-S12)

Figure 9.5: The second moments of the artificial stars taken from images vs. their peak pixel value for an ITL sensor (R03-S12). Left column is the uncorrected data and the right column is corrected data. Each blue point is a single spot from a single exposure. Error bars are standard errors of the distribution about the mean at each exposure time in moment and flux, and a line has been fitted to the 20–50 s points, which corresponds to $2.5 - 9.0 \times 10^4$ el (just below the pCTI turnoff). The vertical shaded regions represent the range of full-well capacities of the 16 amplifiers on the sensor using the four different methods.



(b) E2V (R24-S11)

Figure 9.5: Continued.

with respect to this peak signal level. We fit the data points above the influence of our background light ($f_{\max} > 2.5 \times 10^4$) and below the pCTI turnoff (which is our linear size-growth regime) with the function:

$$I_{\mu\nu} - I_{\mu\nu,20\mathrm{s}} = \alpha_{\mu\nu} f_{max} + c_{\mu\nu}, \tag{9.2}$$

where the indices run over the $\mathbf{x}$ and $\mathbf{y}$, and $\alpha_{\mu\nu}$ determines the relative strength of the BFE, and $c_{\mu\nu}$ just captures the difference between our fiducial point at $I_{\mu\nu,\mathrm{s}}$ and the true intrinsic shape of our spots in the limit of zero signal, which is important for fitting but unimportant for deriving the strength of the BFE.

We will refer to the size of stars using the canonical metric $T = I_{xx} + I_{yy}$. In addition, we will define the anisotropy of the BFE by the metric $\alpha_{yy} - \alpha_{xx}$. Note that this quantity measures the anisotropy of the BFE, not the anisotropy of the stars, which is typically defined using the pseudovector quantity of ellipticity (e_1, e_2), where the components are estimated as $e_1 = (I_{xx} - I_{yy})/(I_{xx} + I_{yy})$ and $e_2 = 2I_{xy}/(I_{xx} + I_{yy})$ in Hoekstra et al. (1998).

The results are shown in Figure 9.5. For the ITL sensor, the size of our artificial stars grow by 1.75% from their intrinsic size near the pCTI turnoff, and they grow by 3.25% for the E2V sensor. However, this size growth is asymmetric, just as we observed in the case of flat fields. For the sensors we measured, the BFE ($\alpha_{\mu\nu}$) is consistently 16% stronger in the parallel direction than in the serial direction for these stars.

Figure 9.5 also shows the result of applying the correction derived using the characteristic kernel. We correct $T = (I_{xx} + I_{yy})$ by 89.2% in ITL and 94.8% in E2V, which is worse than the correction on a_{00} in flat fields (95-97%), and we correct 31.4-37.8% of the anisotropy in the BFE. We measure the I_{xy} component to be poorly corrected only because it is a small component of the star shape and measurement of the BF strength along this component is also small. Table 9.1 breaks down the detailed fit parameters between the two sensors.

Summary of BF Correction on Artificial Stars

Sensor	α_{xx} [px^2/el]	α_{yy} [px^2/el]	α_{xy} [px^2/el]
R03-S12 (ITL)	$(1.84 \pm 0.010) \times 10^{-6}$	$(2.15 \pm 0.006) \times 10^{-6}$	$(5.74 \pm 0.441) \times 10^{-8}$
	$(1.18 \pm 0.098) \times 10^{-7}$	$(3.11 \pm 0.075) \times 10^{-7}$	$(2.73 \pm 0.429) \times 10^{-8}$
	93.5%	85.5%	52.5%
R24-S11 (E2V)	$(3.30 \pm 0.050) \times 10^{-6}$	$(3.86 \pm 0.009) \times 10^{-6}$	$(7.83 \pm 0.311) \times 10^{-8}$
	$(-7.89 \pm 0.385) \times 10^{-9}$	$(3.76 \pm 0.090) \times 10^{-7}$	$(1.89 \pm 0.297) \times 10^{-8}$
	100%	90.2%	76.0%

Table 9.1: Strength of the BFE before (top) and after (middle) correction, with the relative correction level (bottom) on artificial stars in different sensor types. Strengths are given as the fitted slope of the line from Equation 9.2 with 1-σ fitting errors. A value $\alpha_{\mu\nu} > 0$ means the correction is under-correcting the BFE.

We also show the different turnoff points of pixel full-well discussed in §8.1 to show that the BFE has impacts over the full dynamic ranges of our sensors. Interestingly, we observe that the PTC turnoff has little effect on the shapes of the corrected spots in our study compared to the physical effects at the sCTI turnoff and the pCTI turnoff, which results in a small deviation from the linear relation as shape of the spots increases due to deferred charge above this level.

9.4 Flux conservation

We should expect a well-motivated correction that correctly models the charge transport in an image to fully conserve the flux from the measured image. While we confirm zero sum rules on the inferred covariance model coefficients and enforce zero sum on the pixel correlations, we observed that the Coulton et al. (2018) correction does not conserve the PSF flux from the uncorrected image. Figure 9.6 shows the difference in the integrated flux of stars after applying the correction. We use the 35px pixel aperture photometry on each star on images before and after correction without estimating and removing the flat background.

This aperture was chosen to be large enough compared to the PSF size (6px FWHM) to capture the total footprint of each star without needing to assume a particular shape for the profile. The brightest unsaturated stars (below the pCTI turnoff) have a small surplus $\delta < 0.02\%$ of total integrated charge compared to the same star in the uncorrected image. When applying the kernel to an image, more flux is gained by the central pixels than is taken from the neighboring pixels. The correction only appears to conserve charge in the continuous limit ($\langle \delta F \rangle = 0$ integrating over the full image area) in flat fields.

Charge non-conservation was more apparent in high-contrast images such as our artificial stars than in flat fields, which would suggest that the error is the result of a local deviation of the correction from Gauss's law on short distance scales. Most of the correction

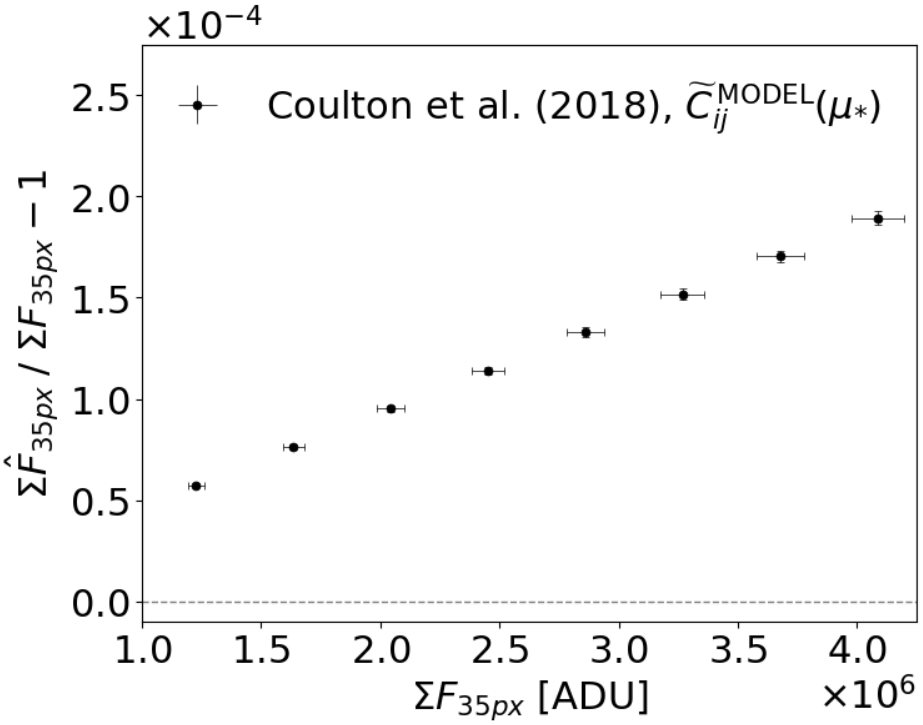


Figure 9.6: The level of charge conservation after applying the BFE correction. The quantity $f = \sum_{r<35\mathrm{px}} \hat{F} / \sum_{r<35\mathrm{px}} F$ is the ratio of integrated flux within 35px of the centroid of each spot, after/before the BFE correction is applied on each of the stars from the ITL sensor (R03-S12) dataset. We only plot spots with peak signals below the pCTI turnoff. Error bars are standard errors of the distribution at each exposure time.

is dominated by the central value of the kernel, but realistically most of the BFE is contributed by correlations more than 4px away. High-contrast images such as these images of star-fields have larger deflection field divergences than those in low-contrast images such as flat fields. Therefore it is more important to properly model the local physics when trying to correct realistic images. This could be problematic for on-sky photometry, particularly for exposures with good-seeing images or under-sampled PSFs, which are especially common to space-based telescopes.

The issue potentially lies with a disagreement between the theoretical formalism of the Coulton et al. (2018) correction and the application of the correction to real images. The kernel itself is derived from discrete covariances between pixels, however the mathematical framework from which Coulton et al. (2018) applies this kernel to an image assumes the flux distribution in the image is continuous and the pixels are infinitely small. Pixelization only occurs at the application of the kernel to the image, at which point we cannot be certain that the flux correction can hold. The flux distribution in an image is a discrete and discontinuous, as a result the gradient is also discontinuous. Around a local maximum in an image, the differences between neighboring pixels do not approximate the divergence around the peak. This is analogous to charge distributing around the area of the pixel and around the boundaries, which is unphysical. The result is that the application of the correction deviates from Gauss's Law on small scales, resulting in loss of charge per pixel conservation in stars. Since this determines how much charge is shifted in each iteration of a kernel, poor modeling of local charge transport could ultimately result in a overall worse amplitude of the correction.

9.5 Adding flux-conserving corrections

In this section, we will attempt to re-apply the scalar Coulton et al. (2018) correction with additional corrections to fully conserve charge at each iteration of applying the kernel. We do this by directly enforcing Gauss's Law in the application of the Coulton et al. (2018) correction in order to fully conserve the flux of the uncorrected image. In the scalar model defined by Coulton et al. (2018), we change the form of Equation 6.8 ($\hat{F} = F + \frac{1}{2}\nabla \cdot V$, where $V = F\nabla(K \otimes F)$) to enforce Gauss's law on small scales such that the charge lost over the area of one pixel is exactly matched by the sum of the charges lost over each boundary (X) of the pixel (such that $\int \nabla \cdot V dx^2 = \Sigma_X V_X$).

The result of applying flux corrections on the EO datasets for our two sensors is shown in Figures 9.7, the corresponding flux conservation is shown in Figure 9.8 and a detailed list of the associated strength coefficients (ref. Equation 9.2) are given in Table 9.2. This analysis conserves charge by construction, and the aperture flux containing the full footprint of the stars is also consistently conserved. We found that it corrects 94.4% of the total PSF size error in the ITL sensor and 97.4% of the total PSF size error in the E2V sensor, and it corrects 36.6% of the anisotropy in the ITL sensor and by 29.1% of the anisotropy in the ITL sensor in the E2V sensor.

The correction appears to re-distribute charge until it approximately converges along one axis of the focal plane and then stops iterating. This may be improved by better modeling of the BFE anisotropy. The residual anisotropy could stem from a conjunction of residual artifacts aligned with the gate/channel architecture or BFEs that were potentially misassociated with other coefficients in modeling the pixel correlations, or it could further suggest yet unmodeled physics at small scales.

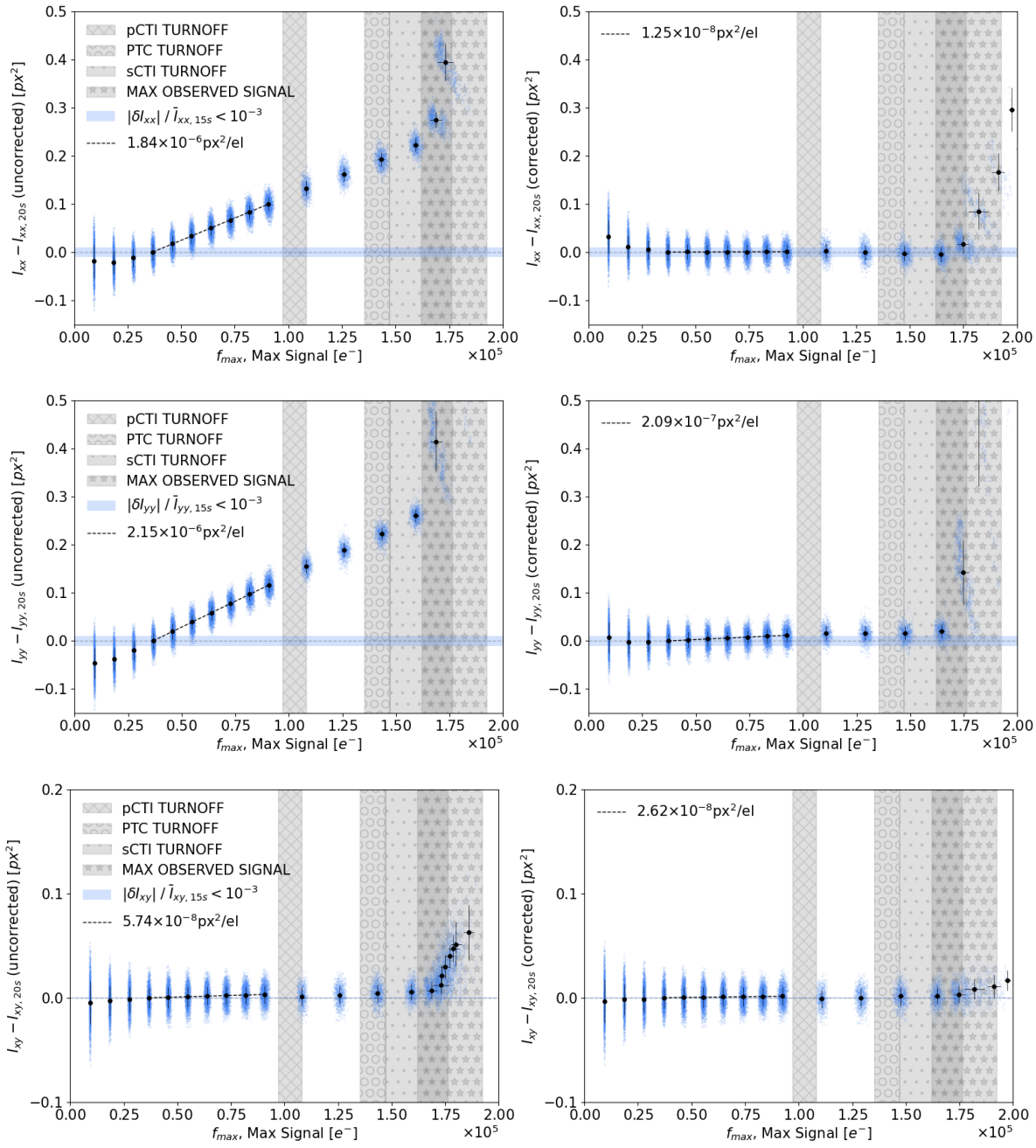
This flux-conserving correction should not be considered a physical solution, although enforcing flux conservation in this way better approximates the correct charge re-distribution

from BFEs. After all, a scalar correction can only consider pixel-scale variations. A more robust physical solution might require a vector approach that models the sub-pixel-scale physics on each pixel boundary. Additional modeling of the each of the four pixel boundary shifts from electrostatics may have the advantage of a more physical solution and may confirm other unmodeled sub-pixel physics.

Summary of Flux-Conserving BF Correction on Artificial Stars

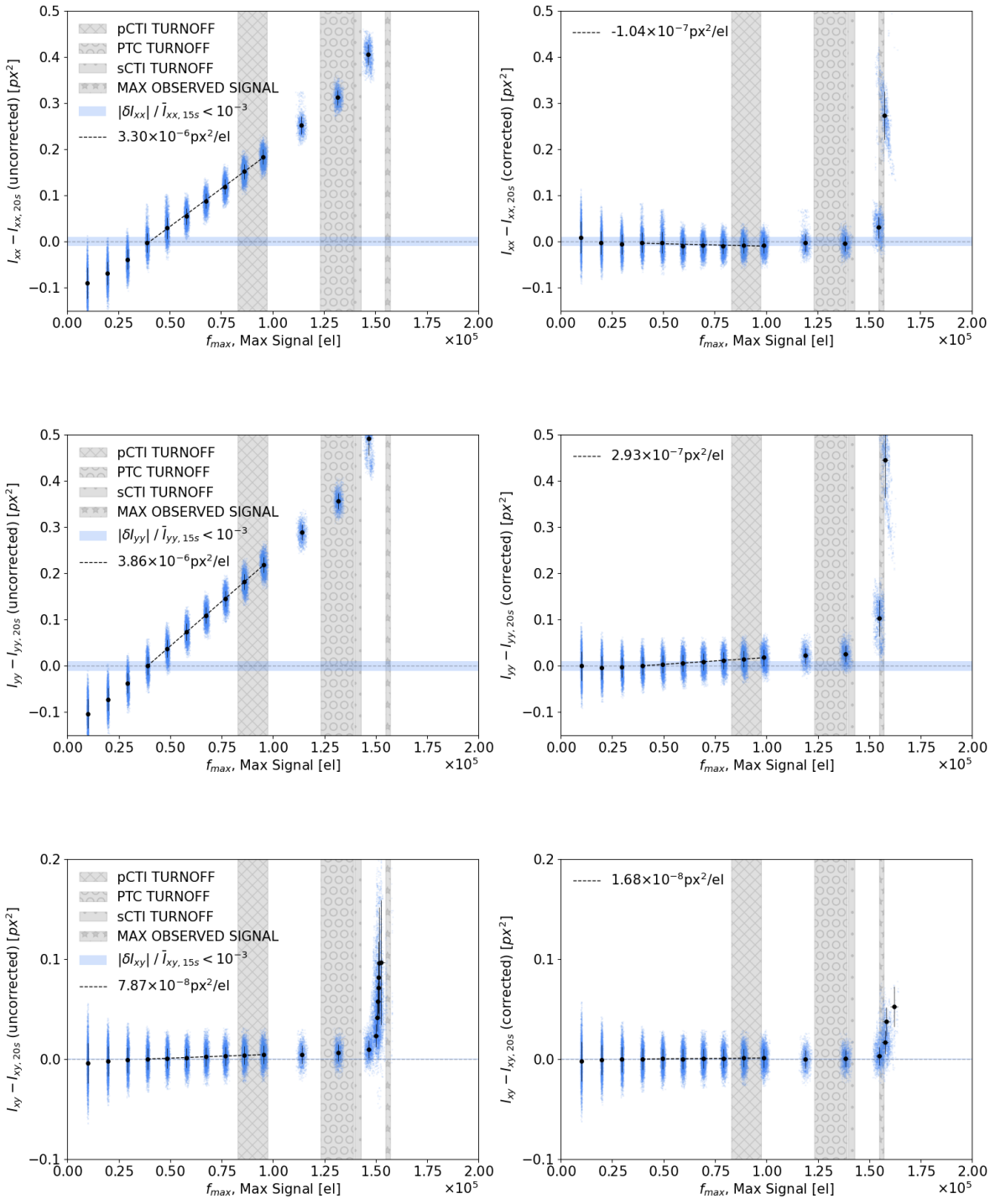
Sensor	α_{xx} [px^2/el]	α_{yy} [px^2/el]	α_{xy} [px^2/el]
R03-S12 (ITL)	$(1.84 \pm 0.010) \times 10^{-6}$	$(2.15 \pm 0.006) \times 10^{-6}$	$(5.74 \pm 0.441) \times 10^{-8}$
	$(1.25 \pm 0.098) \times 10^{-8}$	$(2.09 \pm 0.075) \times 10^{-7}$	$(1.68 \pm 0.429) \times 10^{-8}$
	99.3%	90.3%	70.7%
R24-S11 (E2V)	$(3.30 \pm 0.050) \times 10^{-6}$	$(3.86 \pm 0.009) \times 10^{-6}$	$(7.83 \pm 0.311) \times 10^{-8}$
	$(-1.04 \pm 0.385) \times 10^{-7}$	$(2.93 \pm 0.090) \times 10^{-7}$	$(1.68 \pm 0.297) \times 10^{-8}$
	103%	92.4%	78.5%

Table 9.2: Strength of the BFE before (top) and after (middle) correction, with the relative correction level (bottom) on artificial stars in different sensor types. Strengths are given as the fitted slope of the line from Equation 9.2 with 1-σ fitting errors. A value $\alpha_{\mu\nu} > 0$ means the correction is under-correcting the BFE.



(a) ITL (R03-S12)

Figure 9.7: The second moments of the artificial stars taken from images vs. their peak pixel value. Left column is the uncorrected data and the right column is corrected data using the flux-conserving corrections. Each blue point is a single spot from a single exposure. Error bars are standard errors of the distribution about the mean at each exposure time in moment and flux, and a line has been fitted to the 20–50 s points, which corresponds to $2.5 - 9.0 \times 10^4$ el (just below the pCTI turnoff). The vertical shaded regions represent the range of full-well capacities of the 16 amplifiers on the sensor using the four different methods.



(b) E2V (R24-S11)

Figure 9.7: Continued.

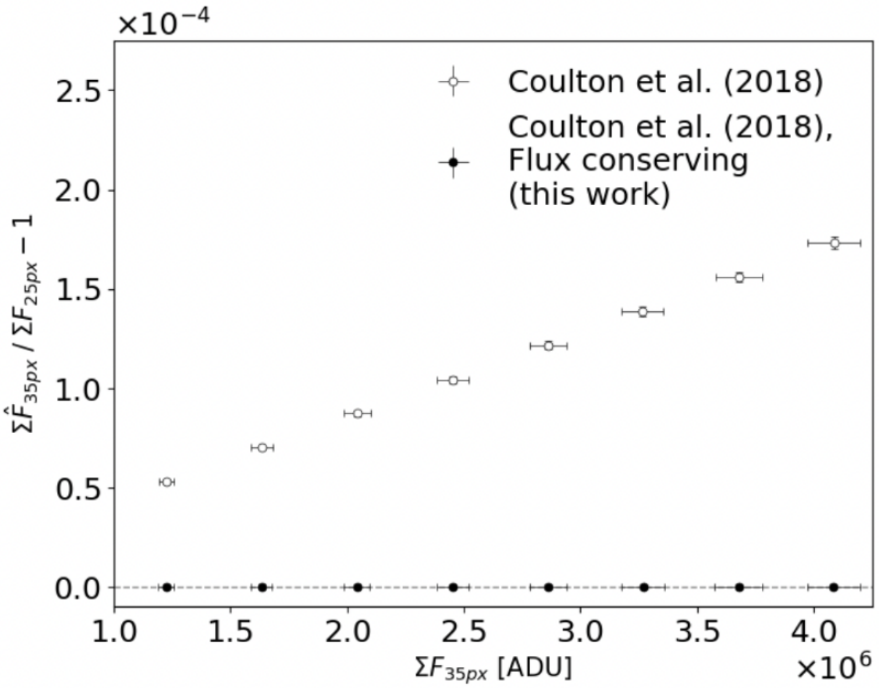


Figure 9.8: The level of charge conservation after applying the BFE with flux-conserving corrections. The quantity $f = \sum_{r<35\text{px}} \hat{F} / \sum_{r<35\text{px}} F$ is the ratio of integrated flux within 35px of the centroid of each spot, after/before the BFE correction is applied on each of the stars from the ITL sensor (R03-S12) dataset. We only plot spots with peak signals below the pCTI turnoff. Error bars are standard errors of the distribution at each exposure time.

Chapter 10

Discussion

10.1 Testing the assumptions of scalar BFE Theory

In this section, we examine the assumptions that allow $\widetilde{C}_{ij}(\mu)$ to take the form $\mu^2\nabla^2 K$. The Coulton et al. (2018) approach assumes (i) that there are no higher-order terms that result in non-linear BFEs, and (ii) the deflection field produced by accumulated charge in a pixel is conservative (zero curl).

1. Non-Linear BFE Components

   The implication of a fixed kernel K in the Coulton et al. (2018) model is that $\widetilde{C}_{ij}$ / μ^2 does not change with signal level, and therefore the higher-order terms of the covariance model are zero. The main contributing higher-order terms of the full covariance model has two components ($\mathbf{a} \otimes \mathbf{a} + \mathbf{ab}$) with two parameters ($\mathbf{a}$ and $\mathbf{b}$). Coulton et al. (2018) do not distinguish between $\mathbf{a}$ and $\mathbf{b}$ or include their higher-order components. Figure 8.7 shows the limited validity of the assumption that $\widetilde{C}_{ij}$ / μ^2 is independent of signal level. We show in the figure that $\widetilde{C}_{ij}$ / μ^2 agrees with the a_{ij} matrix by Astier et al. (2019) in Equation 6.1 only in the limit of zero signal, and that near the pCTI

turnoff it deviates by 20–30%, depending on the sensor. The deviation in the PTC is well-modeled by the non-linear terms of the full covariance model (Equation 6.1).

Figure 10.1 summarizes the correction of artificial stars in R03-S12 (ITL) for different kernels evaluated from different signal levels of the full covariance model, which encode the influence of varying amounts of the higher-order BFEs. The correction based on using $\widetilde{C}_{ij}(0) \sim a_{ij}$ over-corrects the artificial star images, and the correction based on $\widetilde{C}_{ij}(\mu_{\rm pCTI})$ under-corrects the images. Some kind of mechanism to alter the BFE from the pixel boundary shift at non-zero signal level is needed. An adjustment of the strength is needed by constructing a kernel from the measured covariance model at some higher signal level. Scanning the whole covariance model to find the characteristic signal level that reconstructs the expected flat field behavior (as we showed in §9.2) allows us to balance the higher-order BF contributions.

The characteristic kernel however, still does not fully correct the BFE in flat fields or artificial stars (Figures 9.3 and 9.5). The correction also leaves differing residuals in different sensors (Table 9.1). Treatment of the zero-sum rule, unmodeled sources of stray pixel-correlations (NL), or as-yet unmodeled physical complexities could all leave a diversity of additional, sensor-dependent systematic uncertainties in the PSF.

2. Non-Zero Curl

The correction in Coulton et al. (2018) is entirely based on a scalar kernel. Therefore, the calculation of displaced charge in a given pixel, in application of Equation 6.8, only depends on the divergence of the kernel at discrete pixel centers. While the scalar kernel can capture anisotropy in boundary displacements along the two directions, it cannot fully capture the realistic complexity of distortions to all four boundaries on a pixel. Flat fields only provide $(N+1)^2$ degrees of freedom for N pixels, but the boundary shifts represent $4(N+1)^2$ degrees of freedom, or $2(N+1)^2$ if one accounts

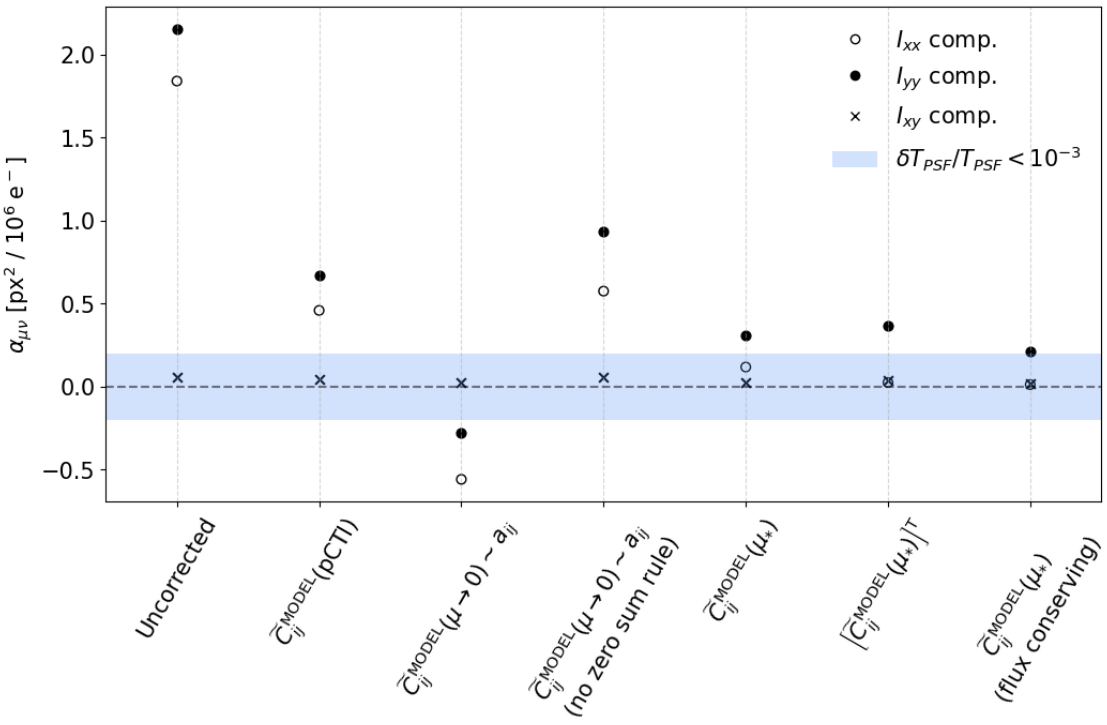


Figure 10.1: The strength of the BFE measured from artificial stars (fit to Equation 9.2) before/after applying scalar corrections derived from different configurations of $\widetilde{C}_{ij}$ for sensor R03-S12 (ITL). We show the fitted slopes for the uncorrected artificial stars as well as for the artificial stars corrected using the covariance model sampled at zero signal, sampled at the pCTI turnoff (minimum value of all amplifiers), the derived characteristic signal level from our full PTC scan, and for the last case, we also show the correction with the kernel transposed (purposefully applied in the wrong orientation), which shows how sensitive the correction is to the fixed-plane BF anisotropy. The shaded area represents the total allowed residual BFE ($\alpha_{xx} + \alpha_{yy}$) to meet the LSST science requirements (The LSST Dark Energy Science Collaboration et al. 2021) on $T_{\mathrm{PSF}} = I_{xx} + I_{yy}$ to the pCTI saturation limit (10^5 el) for stars of size $\langle T_{\mathrm{PSF}} \rangle = 20$ px^2, under the assumption that our artificially produced stars are real, measured PSF stars.

for shared boundaries. Flat fields alone cannot constrain enough degrees of freedom to calculate all four boundary shifts without additional modeling. Astier, Pierre & Regnault, Nicolas (2023) have investigated modeling all four boundary shifts in a pixel from 3D electrostatic solutions in HSC sensors and measured a non-zero and non-negligible, discrete curl-component of the displaced boundaries sourced by the intrinsic anisotropy in the BFE (of order 10^{-7} el^{-1}). The curl component was also measured from simulations in Rasmussen et al. (2016). However, no vector-like curl-component of the deflected pixel boundaries field is included or correctable in the Coulton et al. (2018) model. In fact, they explicitly assume that this component is zero.

The validity of the zero-curl assumption in the LSSTCam can be partially constrained by comparing the correction of flat field covariances and artificial stars. The large pixel-to-pixel contrasts in star images requires accurate modeling of the small-scale physics to converge to a good correction. On the other hand, the small pixel-to-pixel contrasts in flat fields does not require the same accuracy at small scales to achieve a good correction. In the application of the kernel to artificial stars, the correction reaches the total convergence condition before it can properly redistribute charge in the most adjacent pixels, whereas in flat fields, with small pixel contrasts, the correction in the adjacent pixels makes up a larger total fraction of the charge redistribution at each iteration of the kernel. The correction in flat fields therefore redistributes more charge in nearby pixels, in the same number of iterations, before meeting the convergence condition than it does in artificial star images.

The correction might be correcting the anisotropy of the BFE better in flat fields than in stars because the kernel models the local physics better where unmodeled components are small. The BFE in the parallel direction is consistently 215-234% larger than it is in the serial direction when measured from the a_{01}/a_{10} in flat fields.

And the parallel direction is consistently 16% stronger when measured by α_{yy}/α_{xx} in the size-growth of artificial stars. In Figure 10.1, only about 38% of the anisotropy in stars gets corrected by the algorithm. On the other hand, the anisotropy in flat field covariances is corrected by 60-65%. The Coulton et al. (2018) algorithm could be correcting the anisotropy better in the case of flat fields in part because it better models the local physics in flat fields, where the unmodeled curl component would likely be smaller.

If we take an extreme case and purposefully apply the transposed kernel to the raw images–that means purposefully applying the correction with the anisotropic component of the kernel in the wrong direction (Figure 10.1)–the final anisotropy is roughly unchanged with the uncorrected case. This suggests that in artificial star images, most of the correction is dominated by the central pixels of the kernel, however we can see from charge conservation in the **a** matrix (Figure 8.9) that most of the charge redistribution that occurs is due to the pixels that are not immediately adjacent to the central pixel. This could be further evidence that the kernel is not accurately capturing the physics of the BFE at small distances. While enforcing flux conserving corrections appears to better approximate the true charge redistribution in an image on small scales, the correction itself is also not exact. This result still leaves room for other potentially unmodeled physics at small scales.

We cannot conclude for certain the presence of a non-zero or nonnegligible curl component of the BFE distortions in LSSTCam from these data since there are not enough degrees of freedom to derive all four boundary shifts on a pixel from flat field statistics alone, even accounting for shared boundaries. It would require further modeling as is done in Astier, Pierre & Regnault, Nicolas (2023), which includes modeling the 3D electrostatics in a pixel with each boundary shift as a free parameter, in order to fully

constrain the curl component as a potential systematic.

We would like to note here that these measurements of the anisotropy in the BFE in the sensors could be affected by the intrinsic shape of the stars and their alignment on the pixel grid if the BFE is sensitive to the gradient of light between pixels. The stars are intrinsically shaped and oriented on the pixel grid as shown in Figure 9.4 (based on their shapes at low signal level). Although there is little variation among the spots shapes in our sample, we attempt to measure how the BFE changes with respect to this intrinsic star shape in Figure 10.2. We measure the slope of the BFE individually for each star rather than the ensemble, as a function of intrinsic star shape, and we find that there is a small (but slightly beyond estimated error) decrease in the strength of the BFE in more extended sources. Since our stars are intrinsically larger in the y-direction than the x-direction, we would expect that α_{xx} in our sensors is slightly smaller than what we report and α_{yy} in our sensors is slightly larger than we report, therefore the overall anisotropy in our sensors is likely larger than what we report. However, given that there is only a small variation in intrinsic spot shapes, the error bars are on the same order of magnitude as our measurements in Figure 9.4, and it is impractical to extrapolate our measurements for perfectly round sources.

10.2 Impact of BFEs on cosmological probes

It is of practical importance to LSST's cosmological analyses to know how these residual BFEs will impact observables of survey targets and forecast its impact on inferred cosmological outcomes. BFEs can enter into cosmological analyses by contaminating PSF estimation or by directly biasing survey targets (Liaudat et al. 2023).

The impact of BF will rely on the measurements of artificial stars in this study. It it important to note that the artificial stars (1".48, $T \sim 20$ px^2, 7.42px FWHM) were nearly 2×

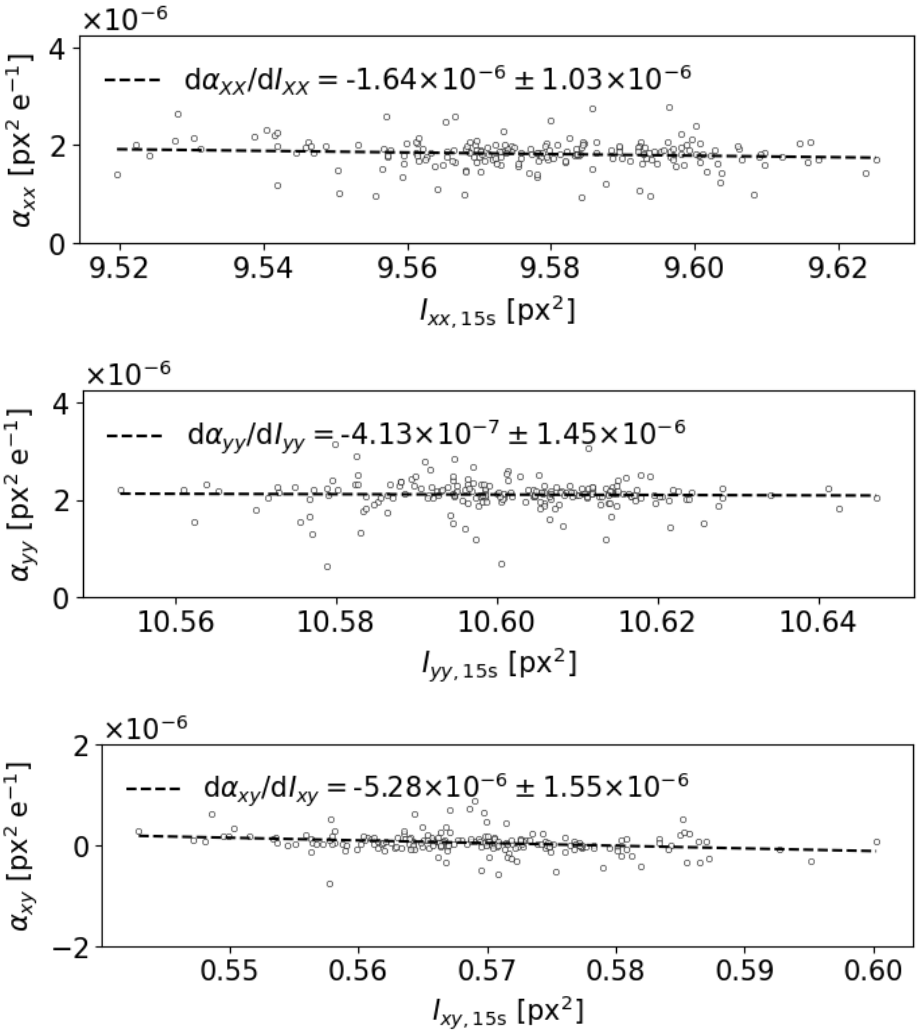


Figure 10.2: The decrease in the influence of the BFE in more extended sources, which would have smaller flux gradients in any direction of the pixel plane. We show the slope of the BFE by component, derived from the 81 individual spots used in our analysis, where each point corresponds to a single hole in the spot mask, relative to their intrinsic shapes. We derived these slopes from the same calculation in Equation 9.2. The errors shown are the fitting errors of the slope. These were calculated for the ITL sensor (R03-S12).

larger than the expected LSST PSF size (0".7, $T \sim 3$ px^2, or 3.5px FWHM), and we would expect the BFE and the residual BFE to be worse for smaller PSFs due to the stronger charge gradients between pixels. We also anticipate that bluer bands will experience an augmented BFE due to a shorter photon conversion depth within the sensor, thereby resulting in a strengthened BFE from the increased drift length (observed by Astier, Pierre & Regnault, Nicolas 2023, in HSC and ongoiong work to measure the effect in LSSTCam sensors).

In addition, the outcome of cosmological analyses depends on a myriad of survey, hardware, and analysis choices throughout the development and lifetime operation of a survey, and we cannot hope capture the full complexity of this in a single thesis section. We will therefore employ simplified models to obtain only an order of magnitude estimate of how BF systematic effects enter into these different analyses.

10.2.1 PSF constraining power

For a single star, the modeling bias induced by BFEs is a function of $\delta\vec{e}_{\rm PSF} = f(\vec{e}_{\rm PSF}, \mathcal{S})$, where $\mathcal{S}$ is the signal-to-noise ratio (S/N) based on the definition in Paulin-Henriksson et al. (2008). The average PSF modeling bias over a star field in a single-visit is given by a function $\langle\delta\vec{e}_{\rm PSF}\rangle = f(\langle\vec{e}_{\rm PSF}\rangle, \mathcal{S}_{\rm eff})$ where $\mathcal{S}_{\rm eff}$ is the effective signal-to-noise ratio and depends on the number density function of the signal-to-noise ratio over the imaged area ($dn/d\mathcal{S}$) as well as the stellar selection function. The selection function should depend on the fact that dim stars contain relatively little PSF constraining power and are often difficult to measure accurately while bright stars carry more PSF constraining power, but they are particularly impacted by BFEs (Paulin-Henriksson et al. 2008). Cosmic shear surveys typically employ a variety of selection functions (see Paulin-Henriksson et al. 2008; Bertin 2011; Jarvis et al. 2016; de Jong et al. 2017; Aihara et al. 2018), the simplest is to apply minimum and maximum signal-to-noise ratio cut when selecting which stars to use for PSF estimation. The framework

proposed by Paulin-Henriksson et al. (2008) uses:

$$\mathcal{S}_{\rm eff}^2 = N_{\rm TOT}^{-1} \int_{\mathcal{S}_{\rm min}}^{\mathcal{S}_{\rm max}} \frac{dn}{d\mathcal{S}} \mathcal{S}^2 d\mathcal{S}, \tag{10.1}$$

where $N_{\rm TOT}$ is the total number of stars used to estimate the PSF between $\mathcal{S}_{\rm min}$ and $\mathcal{S}_{\rm max}$. They also write the *effective* number of stars as:

$$N_* = \mathcal{S}_{\rm eff}^{-2} \sum_i S_i^2 \tag{10.2}$$

where i runs over all the stars used for PSF calibration. A perfectly corrected BFE would enable surveys to use stars all the way up to pixel saturation, enabling better PSF constraining power as $\sigma_{\rm sys}^{\rm PSF} \propto 1/\sqrt{N_* \mathcal{S}_{\rm eff}^2}$ with each new effective star. Improvements to the BF correction will allow for higher thresholds and the inclusion of more stars with higher statistics and result more accurate and precise PSF estimation.

We will therefore need to apply the methods in this paper to all LSSTCam sensors to determine sensor-specific the maximum signal thresholds that will inform PSF star selection across the focal plane. The requirements laid out for LSST by The LSST Dark Energy Science Collaboration et al. (2021) requires us to reconstruct the true PSF size to a level of $\delta T_{\rm PSF}/T_{\rm PSF} < 10^{-3}$ in the co-added PSF in Y10, which in our lab-simulated stars is only satisfied for stars with peak signals below 4.6×10^4 el in the ITL sensor and 5.5×10^4 el in the E2V sensor, which covers only about 50% and 70% of the dynamic range of these sensors (up to pCTI turnoff). With the flux-conserving correction, the systematic requirement is fully satisfied below 9.1×10^4 el in the ITL sensor and 1.1×10^5 el in the E2V sensor, which covers only more than 90% of the dynamic range of the ITL sensor and all of the dynamic range of the E2V sensor. Therefore, we conclude that the correction of these artificial stars is good enough for LSST science requirements below a certain peak signal level, which is

chosen to be chosen with regard to pre-defined accuracy requirements when selecting a PSF star sample.

10.2.2 Type Ia supernovae

Observationally, SN Ia are resolved like PSFs, however the maximum S/N of SN Ia in their brightest wavelength band is typically S/N < 100 in single visit-exposures, which is dimmer than the stars we use to calibrate the PSF in the same image (typically S/N > 100). We employ forced-PSF photometry at the locations of SN Ia to avoid the risk of noise/background divergences associated with aperture-integrated photometry. However, in the presence of BFEs, we would expect the size of the SN Ia to be smaller than the model PSF, which we would expect to result in an overestimation of the SN Ia's flux.

The flux estimator we are using is the unweighted least-squares fit of the PSF image (F_{ij}^{PSF}) to the cutout image of the SN Ia (F_{ij}):

$$\chi^2 = \sum_{ij}(F_{ij} - \Phi F_{ij}^{\mathrm{PSF}})^2. \tag{10.3}$$

The best flux estimator ($\hat{\Phi}$) is just the value of Φ that minimizes the chi-squared, and is simply the solution to $d\chi^2/d\Phi = 0$:

$$\hat{\Phi} = \frac{\sum_{ij}(F_{ij}F_{ij}^{\mathrm{PSF}})}{\sum_{ij}(F_{ij}^{\mathrm{PSF}})^2}. \tag{10.4}$$

Using the GALSIM (Rowe et al. 2014) framework, we simulate a forced-PSF photometry on arbitrary SN Ia images with BF-biased PSFs similar to the ones measured in LSSTCam in §9.3. The PSF and the SN Ia are both modeled as 2D Gaussians centered on a 65×65px grid. They are both drawn on the pixel grid without noise and with no sky-background, host-galaxies, or neighbors; the PSF is also normalized to a signal level of 1 and the SN Ia

is given an arbitrarily large flux (Φ).

To inject BF systematics into the simulation, each image is given an ellipticity (e_1, e_2) selected at a given signal level ($f_{\rm max}$) from the measurements of artificial stars, using the intrinsic ellipticities of the stars and the measured BF slopes in Equation 9.2:

$$\begin{aligned} \sigma(f_{\rm max}) &= \sqrt{I_{xx}(f_{\rm max}) + I_{yy}(f_{\rm max})} \\ e_1(f_{\rm max}) &= \frac{I_{xx}(f_{\rm max}) - I_{yy}(f_{\rm max})}{I_{xx}(f) + I_{yy}(f_{\rm max})} \\ e_2(f_{\rm max}) &= \frac{2I_{xy}(f_{\rm max})}{I_{xx}(f_{\rm max}) + I_{yy}(f_{\rm max})}, \end{aligned} \tag{10.5}$$

Note that $f_{\rm max}$, as in Equation 9.2, is the peak pixel value of the SN in the brightest image taken of it and *not* a measurement of the total apparent brightness of the SN Ia at the maximum of its light curve.

To test the worst case BF scenario, we fix the PSF to the same ellipticity as the artificial stars at the maximum signal level before the pCTI turnoff, as if it were a real analysis where all the stars used to estimate the PSF in an image were nearly saturated ($\langle f_{\rm max}^{\rm PSF} \rangle =$ pCTI turnoff).[1] On the other hand, the SN Ia ellipticity is allowed to vary by selecting signal levels somewhere between 0 and the brightness of the stars used to calibrate the PSF: $f \equiv f_{\rm max}/\langle f_{\rm max}^{\rm PSF} \rangle \in [0.01, 1]$.

In the case where the SN is as bright as the stars used to calibrate the PSF ($f = 1$), we would expect BF to affect them both equally, and they should have that same size/shape and the flux estimator should be theoretically perfect ($\hat{\Phi}/\Phi = 1$). In the case where the SN is dimmer than the stars used to calibrate the PSF ($0 < f < 1$), we would expect the PSF model to be larger than the SN, and therefore the flux will be overestimated. In the limit of a very dim SN and a very bright set of PSF stars ($f \to 0$), we would expect to observe the largest photometric errors in that sample.

[1] We choose $\langle f_{\rm max}^{\rm PSF} \rangle = 9.0 \times 10^4$el for the ITL sensor and $\langle f_{\rm max}^{\rm PSF} \rangle = 7.5 \times 10^4$el for the E2V sensor.

We perform this test using the slopes without any BF correction, with the BF correction, and the slopes with the flux-conserving BF correction, whose fitted values are in Table 9.1 and Table 9.2. We also perform this test for both the ITL and E2V sensor values. The results are shown in Figure 10.3. The LSST Dark Energy Science Collaboration et al. (2021) (see also Ivezić et al. 2019) requires that SN Ia photometry to an accuracy better than 10^{-3}. We conclude that BF correction is necessary to achieve these levels using PSF forced photometry, but that any of the corrections tested out in this analysis can meet these goals with this sample of artificial PSFs. There are some important caveats, however: (i) this analysis does not consider non-Gaussian profiles, (ii) the PSF size was nearly 2× the size of the expected seeing for LSST, and it is possible that a sharper PSF will have an correspondingly larger BFE slope, (iii) this analysis also did not consider the impact of sky background or host galaxy light, and (iv) this analysis only considers flux errors associated with modeling errors on the PSF and it does not take into account the additional loss of flux conservation due to the application of the one BF correction (green in Figure 10.3).

10.2.3 Weak gravitational lensing

Most galaxies observed by LSST within will have an apparent magnitude $m > 21$ in r-band in a single visit exposure (LSST Science Collaboration et al. 2009). For comparison, a Gaussian source at that apparent magnitude with a size of 0".9 will have a maximum peak signal below 1000 el in a 30s exposure (Gee et al. 2007). On the other hand, stars at a typical LSST PSF size of 0".7 that would likely be used for PSF estimation (S/N > 100) will have a minimum apparent magnitude around $m = 21$ and begin to saturate at the pCTI turnoff around $m = 16.5$ in a single-visit image (LSST Science Collaboration et al. 2009). Observationally, stars used for PSF estimation are typically much brighter than galaxies used for WL analyses, and the primary effect of BF will be through misrepresentation of the

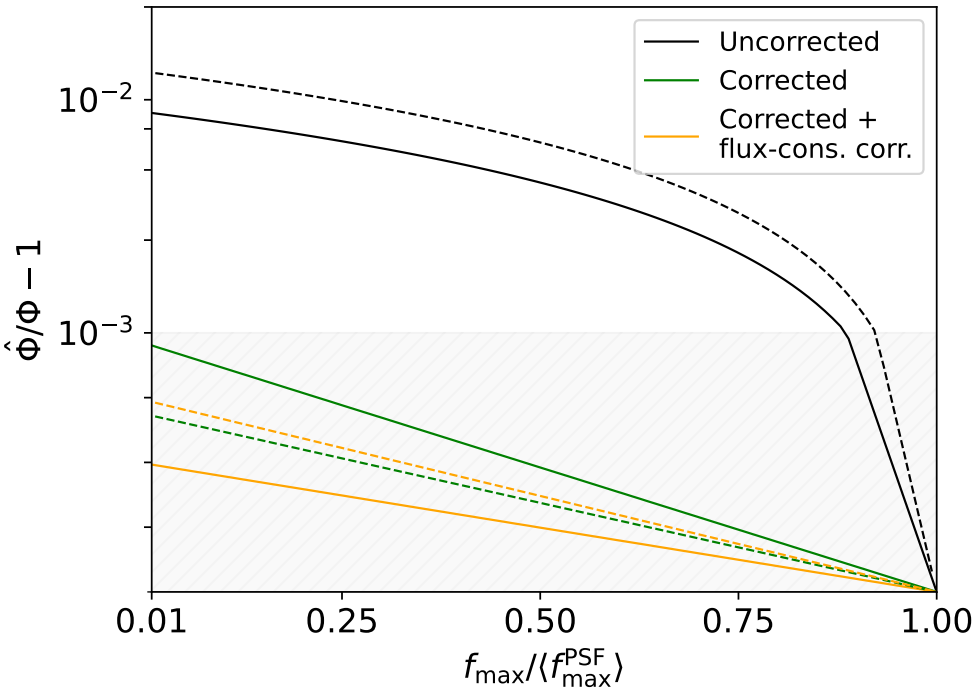


Figure 10.3: Accuracy of the forced PSF flux estimation ($\hat{\Phi}$) compared with truth (Φ) as a function of relative peak signal level between the SN ($f_{\max}$) and the stars used to calibrate the PSF ($\langle f^{\mathrm{PSF}}_{\max}\rangle$). We show this for the ITL sensor (solid lines) and the E2V sensor (dashed lines) for the BF slopes in the uncorrected case (black), the corrected case (green), and the flux-conserving correction case (orange). The shaded region is plotted in linear scale and the non-shaded above it is in log-scale; the shaded region is also within the 10^{-3} requirement for SN Ia photometric accuracy for LSST (Ivezić et al. 2019; The LSST Dark Energy Science Collaboration et al. 2021). For example, the photometric accuracy of a SN Ia whose peak signal is 20% of the stellar peak signal cannot achieve the required photometric accuracy without additional correction. However, a SN Ia whose peak signal is 95% of the stellar peak signal can achieve an accurate forced PSF photometry without a BF correction.

size and shapes of stars. BF distortion effects on the bright stars in our study are of O(1%), which is larger than the O($0.1 - 1\%$) distribution of shape distortions to galaxies caused by weak gravitational lensing from galaxy clusters.

The impact of the imperfect PSF shape estimation on inferences of cosmological parameters is still a major challenge for Stage IV surveys (Paulin-Henriksson et al. 2008; Rowe 2010; Jarvis et al. 2016; Zhang et al. 2021, 2022). Any error in the PSF model can result in an error in the PSF-corrected ellipticity of a galaxy. As described in §4.3.1, weak lensing analyses typically use a linear model for describing biases to a single galaxy's shear estimate ($\hat{\gamma}$) from its measured shear (γ):

$$\hat{\gamma} = (1 + m)\gamma + c, \tag{10.6}$$

where biases can enter multiplicatively (m) as a fraction of the galaxy ellipticity or additively (c) independently of the galaxy ellipticity. Understanding how a change in the PSF will affect the shear estimate of course depends on the chosen shear estimator algorithm (e.g. see Sheldon et al. 2020, 2023; Li et al. 2023). Here we will assume the shear estimator is perfectly well-calibrated and unbiased for a perfectly modeled PSF ($m = c = 0$) so that we only consider errors that enter into the shear estimation due to BF-induced PSF modeling errors.

Cosmological theories most directly predict the matter power spectrum (two-point correlation function in k-space), so to compare measurements to theoretical predictions, we need to know how errors will alter the two-point correlation functions over an ensemble of many galaxies. The shear auto-correlation function becomes:

$$\langle \hat{\gamma}\hat{\gamma} \rangle = (1 + m)^2 \langle \gamma\gamma \rangle + 2(1 + m)\langle c\gamma \rangle + \langle cc \rangle \tag{10.7}$$

We would expect that $\langle c\gamma \rangle = 0$ since the additive error should not correlate with the shear of the galaxy over a large region of the sky with images that have been rotated over a large range of angles.

Biases in the estimated PSF size (trace of the second moment matrix) result in multiplicative biases in the weak lensing shear signal since the PSF size enters into shear estimation via convolution and the bias is therefore proportional to the galaxy size. A multiplicative bias will result in a shear correlation function error that is proportional to the true underlying shear correlation function by a factor of about $2m$ for small biases. Poorly correcting the amplitude of the BF correction contaminates the size of the PSF. For example, an overestimated PSF size results in an underestimated galaxy size and an overestimated shear. The framework in Mandelbaum et al. (2015); Paulin-Henriksson et al. (2008); Massey et al. (2007) proposes a simplified model with a Gaussian PSF and a Gaussian galaxy, and the error in the estimated shear (of a single galaxy) due to a change in PSF size is given by:

$$m \equiv \frac{\delta\hat{\gamma}}{\hat{\gamma}} = \frac{\delta\hat{e}_+}{\hat{e}_+} = (R_2^{-1} - 1)\frac{\delta T_{\rm PSF}}{T_{\rm PSF}} \tag{10.8}$$

The resolution factor (R_2), written in terms of second moments, defines the relative size of the PSF ($T_{\rm PSF}$) and the size of the PSF-convolved image of a galaxy (T_I).

$$R_2 = 1 - \frac{T_{\rm PSF}}{T_I}. \tag{10.9}$$

The resolution factor approaches 1 for well-resolved galaxies and is 0 for poorly resolved galaxies. Simulations by Almoubayyed et al. (2020), using the COSMOS catalog galaxies (Mandelbaum et al. 2015) down to an i-band limiting magnitude of 25.2, find that the arithmetic averaged ratio of $T_{\rm PSF}/T_{\rm gal}$ for LSST with a Gaussian PSF of FWHM 0".6 (typical of Cerro Pachón seeing measured from the Differential Image Motion Monitor or DIMM, The

LSST Dark Energy Science Collaboration et al. 2021) is approximately 2.10. For comparison, Jarvis et al. (2016) reported similar values (2.42) for the Dark Energy Survey (a Stage III survey), although the value for LSST (a Stage IV survey) reflects that next generation surveys be larger and deeper, and we expect to include more galaxies at higher redshifts, which are more likely to be smaller and poorly resolved. If we approximate $M_{\mu\nu}^{(I)} = M_{\mu\nu}^{(\mathrm{gal})} + M_{\mu\nu}^{(\mathrm{PSF})}$, then then we should expect $\langle R_2 \rangle \approx 0.32$ for LSST's 10 year survey.[2] If we have a PSF like the one determined by our artificial stars ($T_{\mathrm{PSF}} \sim 20$ px^2), then we should expect $R_2 \sim 0.07$, which means galaxies would be very poorly resolved, and modeling errors in the PSF will have a large impact on the observed shapes of galaxies.

Following Equation 10.8, if we used nearly saturated stars to determine the PSF ($f_{\mathrm{max}}^{\mathrm{PSF}}$ = pCTI turnoff), then we should expect a PSF error of roughly $\delta T_{\mathrm{PSF}}/T_{\mathrm{PSF}} = 10^{-3}$ with the flux-conserving corrections and ultimately a multiplicative error of $m = 0.0132$. To achieve a multiplicative error $|m| < 10^{-3}$ using the flux-conserving correction, we would need to know T_{PSF} to an accuracy of better than 7.5×10^{-5}, which is only true for stars less than $f_{\mathrm{max}} < 340$ el (which is impractically dim). Therefore, we conclude is that the artificial stars in this study are far too large to make any realistic description of the propagation of real multiplicative errors from BF in a weak lensing analysis.

Biases in the PSF shape (ellipticity) result in additive errors to the estimated galaxy shapes, since the PSF shape bias is independent of the galaxy shape. Improperly correcting the anisotropy of the BF effect in the focal plane coordinate system can result in PSF ellipticity errors in the sky coordinate system. The additive systematics act as a spurious, independent offset to the two-point correlation function. The LSST Dark Energy Science Collaboration et al. (2021) does not place any requirements on the PSF ellipticity or PSF ellipticity error (c) since additive biases are typically easy to identify from null tests (Heymans

[2]This is true for unweighted second moments ($w(\vec{x}) = 1$ in Equation 4.20), but it is a decent approximation here.

et al. 2012; Mandelbaum et al. 2018; Li et al. 2022; Giblin et al. 2021) because they create an anomalous power of cross-correlation between galaxy shear and star shapes compared to the same cross-correlation for a random lens catalog over the same region of the sky. The shear auto-correlation function can then be corrected by forward modeling the additive bias systematics (for examples, see Heymans et al. 2012; Zhang et al. 2023). In practice, weak lensing analyses also use co-added images to reduce noise contamination in shear calibration (Mandelbaum et al. 2023). The PSFs used at the location of a galaxy are the inverse-variance weighted sum of all the PSFs in all exposures. In a typical survey, these co-added images will combine exposures with multiple rotation angles of the camera focal plane to avoid imprints of the focal plane coordinate system on the sky coordinate system. This means that coherent additive errors in many exposures rotated over many angles can become multiplicative errors in the co-added images (Almoubayyed et al. 2020; Ivezić et al. 2019; Almoubayyed et al. 2020).

Deviations from the assumed Gaussianity are quantified by changes to the standardized weighted radial fourth moment (kurtosis or $\rho^{(4)}$) in Zhang et al. (2021, 2022), and this algorithm is implemented in GALSIM (Rowe et al. 2014). It is defined as follows:

$$\rho^{(4)} = \frac{\int (r/\sigma)^4 I(\vec{x}) w(\vec{x}) \mathrm{d}\vec{x}}{\int I(\vec{x}) w(\vec{x}) \mathrm{d}\vec{x}} \tag{10.10}$$

where $r = |\vec{x}|$ and $\sigma = \sqrt{(I_{xx} + I_{yy})/2} = \sqrt{T/2}$, and $w(x, y)$ is the adaptive Gaussian weight defined by Hirata & Seljak (2003). For a purely Gaussian profile, we would expect this measurement of kurtosis to be $\rho^{(4)} = 2$. Errors in kurtosis can affect errors in second moment statistics through the weight function, leading to both multiplicative and additive shear biases. A study by Zhang et al. (2021), using a combination of the HSC Y1 dataset and image simulations in GALSIM, found that kurtosis modeling errors of the PSF produced multiplicative bias in HSC Y1 cosmic shear on the order of 10^{-3} and additive errors of order

10^{-2}. A later–more comprehensive–study by Zhang et al. (2022) found that other higher moments (higher powers of r, up to 6th order) can also induce multiplicative and additive biases. Although, for a constant bias, even-numbered moments bias shear estimates $2-3\times$ more than odd-numbered moments. The same study found that the total multiplicative errors produced by these higher moments could produce a 0.1σ uncertainty in cosmological parameters, and the total additive errors could produce a 1σ uncertainty, if not properly taken into account. We did not measure any deviation in higher-order moments in the stars in our study, however, we acknowledge that this could be a topic for further study. The higher-order moments are determined by higher-powers on the distance coordinate, and as such are very sensitive to the modeling of the tails of the PSF distribution; given the large size of the artificial stars compared to real stars, as well as the presence of our background, it is possible that any changes to the higher-order moments were overwhelmed by the variations in these other systematic variables. It is also important to note that The LSST Dark Energy Science Collaboration et al. (2021) do not yet place constraints on PSF higher-order moment errors for LSST science.

10.3 Proposed improvements

There are several steps that could be added and tested to the LSST Science Pipelines that could improve the correction:

1. We will need to apply these methods individually to all LSSTCam science sensors to determine sensor-specific calibrated thresholds that will inform PSF star selection rather than constructing a correction in a monolithic way for the full focal plane as was done in analyses by Aihara et al. (2018) and Miyatake et al. (2023) for HSC Y3 ISR.

2. The BFE is small, and the calibration of the BFE can be very sensitive to the calibration of other effects that can impact pixel correlations, such as charge transfer inefficiency (CTI), non-linearity, overscan subtraction, as well as other sensor-specific calibrations, such as gain estimation. Extreme care must therefore be taken to remove these other effects and validate these other ISR steps before deriving calibrations for the BFE from flat fields and applying the correction to an image. In particular, our best-fit model parameters of the PTC are somewhat sensitive to the residual NL for these sensors since we fit up to the pCTI turnoff, which is above the signal level we start to observe significant signal NL. In addition, the PTC scanning method we describe is sensitive to the signal range on which we calculate the χ^2 for the same reason. While we mitigate this by limiting the signal range, the method we propose here does require the adequate removal of NL and other sources of pixel correlations.

3. Include a curl component of the deflection field. It would not be possible to implement further "curl" corrections into the Coulton et al. (2018) model. However, this could be done by using an electrostatically derived vector-based model fit to the pixel-pixel correlations in flat fields, with each pixel boundary shift being a free parameter. Such

a method is proposed in Astier, Pierre & Regnault, Nicolas (2023) and tested for HSC, but it should be considered and tested in LSSTCam in future work.

Chapter 11

Conclusion

In this chapter, we measured the BFE in two types of sensors in the LSSTCam focal plane from flat field and artificial stars in the lab. We then quantified and qualified the scalar BF correction proposed by Coulton et al. (2018) with improvements to account for sources of higher-order pixel correlations and charge conservation:

1. We measured dense flat pair data to characterize the PTC, and we found that the BFE gets stronger with signal due to higher-order, non-linear BFE components in a model by Astier et al. (2019), which includes flux-dependent drift fields and feedback mechanisms on accumulating charges. The BF correction proposed in Coulton et al. (2018), which uses a scalar kernel derived from the pixel covariance matrix, relies on the assumption of linearity. The merit of the correction therefore depends on the signal level from which we calculate the covariance matrix and derive the kernel, which is a priori ambiguous.

2. We introduced a new procedure to derive the kernel in a way that includes the contribution of non-linear BFE terms in the correction. We determined the kernel that best reconstructs the linear PTC. We then used this kernel to correct the BFE in both

sensor types by 95-97%, but this was only able to correct up to a flux we identify as the parallel deferred charge turnoff at 10^5 el, where we can no longer efficiently transfer charge along an image column during readout. The flux range of pixels that can be used for weak lensing science is severely limited by this cutoff.

3. We corrected the BFE in artificially produced stars by $94-100\%$ in the serial direction and by $86-90\%$ in the parallel direction. The BFE was observed to be anisotropic (16%) relative to the CCD coordinate system, and the correction preserved $32-38\%$ of this anisotropy. We also observed that the correction violated charge conservation by (0.02%) near saturation at the pCTI turnoff level.

4. We introduced another procedure to apply the kernel in a way that fully conserves charge, and tested the correction again on the same artificial stars. We corrected the BFE in artificially produced stars by $99-103\%$ in the serial direction and by $90-93\%$ in the parallel direction. The BFE was observed to be anisotropic (16%) relative to the CCD coordinate system, and the correction preserved $36-38\%$ of this anisotropy. We observed that the new correction fully conserved flux across the dynamic range of the sensors.

5. We eliminated other systematics that could have interfered with these results, including non-circular artificial stars and charge-conservation, to identify the source of these effects as intrinsic BFEs in the sensor itself and limitations of the assumptions of the Coulton et al. (2018) correction. Our data suggests that the higher-order BFEs are not negligible and that the correction is not properly modeling the physics of charge transport at small distances. Part of this could be deviations in the correction model from Gauss's Law along local pixel boundaries or potentially non-zero curl components in the deflected field boundaries, which are also not included in the Coulton et al. (2018) model.

6. Finally, we explored how residual errors from BFEs could ultimately impact cosmology and other science goals within the system requirements defined by LSST's primary science drivers. While we found the stars in this study were too large to realistically describe the impact of PSF modeling errors on galaxy shapes and shear estimation for cosmic shear analyses, we did show how improvements to the BF correction could improve the PSF constraining power by extending the available signal range for the stellar selection function, and we found that the two correction schemes we showed are both able to achieve photometric accuracy required to SN Ia cosmological probe is fully satisfied with the artificial PSFs by

Our findings motivate a detailed study on more realistic PSF stars and how measurement errors from BF could ultimately impact cosmology and other science goals. Ultimately, it is important to characterize the sensitivity of cosmological parameters to observables biased with BFEs. Even with state-of-the-art correction techniques, the residual effects could represent a significant component of the systematics error budget for cosmological analyses of LSST observations.

Astier et al. (2019) pointed out that all of the correction methods reported in literature only correct 90% of the initial effect in the images (Guyonnet et al. 2015; Gruen et al. 2015) and that all of the proposed correction techniques assume no contribution from higher-order terms. Several different correction algorithms have been proposed by Antilogus et al. (2014), Gruen et al. (2015), Coulton et al. (2018), and Astier, Pierre & Regnault, Nicolas (2023), and they all require calibrated inputs of the amplitude of the BFE from flat field statistics, and none allow for the BFE to evolve with signal level. The flux-scanning methodology might be applicable to these other correction techniques as well, but we reserve this for a more careful and detailed study in future work.

Chapter 12

Epilogue

The current state of our understanding about the Dark Universe relies heavily on our knowledge of the physical sub-processes out there and the analytical methods employed to measure those things. Observations confirm that the Universe is primarily made up of dark matter and dark energy, however the true identity of these components remains elusive. In addition, recent observations by several recent experiments using a variety of different methods appear to disagree with some predictions of this model which suggests a potential paradigm shift in this field (see DESI Collaboration et al. 2024; Amon et al. 2023; Planck Collaboration et al. 2020a; Asgari et al. 2021, and for an thorough discussional review, see McGaugh 2023). Our confidence in our theories, tested with these types of analyses, depends on very accurate measurements of positions (astrometry), fluxes (photometry), and shapes of large samples of astronomical objects to find small adjustments to these observables against the noise of their many possible intrinsic properties. In general, the quality of all science that will be conducted with modern surveys is dependent on well-tuned survey properties, sensor operational configurations, and/or data-processing algorithms that mitigate systematic errors at the pixel level.

I will leave on one final note. The fact that we can measure single photons from galaxies

of $O(10^9)$ly away, and derive *anything* scientifically useful is impressive. I recall an audience member once say "It's a wonder precision cosmology is even possible at all." And I think this sentence captures the current state of the field perfectly. These achievements never could have been made without the diligence, ingenuity, and curiosity of so many people. Who knows what scientific analyses will look like to physicists in 10, 50, or 100 years, but I do know that we are tantalizingly close to the sensitivity thresholds needed to begin distinguishing fundamental models of the Dark Universe.

Bibliography

Aalbers, J., Akerib, D. S., Akerlof, C. W., et al. 2023, PhRvL, 131, 041002, doi: 10.1103/PhysRevLett.131.041002

Abazajian, K., Canac, N., Horiuchi, S., & Kaplinghat, M. 2014, Physical Review D, 90, doi: 10.1103/PhysRevD.90.023526

Abazajian, K. N. 2011, Journal of Cosmology and Astroparticle Physics, 2011, 010, doi: 10.1088/1475-7516/2011/03/010

Abazajian, K. N., & Kaplinghat, M. 2012, Phys. Rev. D, 86, 083511, doi: 10.1103/PhysRevD.86.083511

Abbott, T. M. C., Aguena, M., Alarcon, A., et al. 2022, PhRvD, 105, 023520, doi: 10.1103/PhysRevD.105.023520

Abdollahi *et al*, S. 2020, The Astrophysical Journal Supplement Series, 247, 33, doi: 10.3847/1538-4365/ab6bcb

—. 2022, The Astrophysical Journal Supplement, 260, 53, doi: 10.3847/1538-4365/ac6751

Ackermann, M., Ajello, M., Atwood, W., et al. 2012, The Astrophysical Journal, 750, 3

Ackermann, M., Ajello, M., Albert, A., et al. 2017a, ApJ, 840, 43, doi: 10.3847/1538-4357/aa6cab

—. 2017b, ApJ, 836, 208, doi: 10.3847/1538-4357/aa5c3d

Agostinelli, F., Hoffman, M., Sadowski, P., & Baldi, P. 2014, arXiv preprint arXiv:1412.6830

Aihara, H., Arimoto, N., Armstrong, R., et al. 2018, PASJ, 70, S4, doi: 10.1093/pasj/psx066

Ajello, M., Albert, A., Atwood, W., et al. 2016, The Astrophysical Journal, 819, 44

Ajello, M., Albert, A., Atwood, W. B., et al. 2016, ApJ, 819, 44, doi: 10.3847/0004-637X/819/1/44

Albrecht, A., Bernstein, G., Cahn, R., et al. 2006, arXiv e-prints, astro, doi: 10.48550/arXiv.astro-ph/0609591

Almoubayyed, H., Mandelbaum, R., Awan, H., et al. 2020, arXiv e-prints, arXiv:2006.12538, doi: 10.48550/arXiv.2006.12538

Almoubayyed, H., Mandelbaum, R., Awan, H., et al. 2020, Monthly Notices of the Royal Astronomical Society, 499, 1140, doi: 10.1093/mnras/staa2879

Amon, A., Robertson, N. C., Miyatake, H., et al. 2023, MNRAS, 518, 477, doi: 10.1093/mnras/stac2938

Antilogus, P., Astier, P., Doherty, P., Guyonnet, A., & Regnault, N. 2014, Journal of Instrumentation, 9, C03048–C03048, doi: 10.1088/1748-0221/9/03/c03048

Argyriou, I., Lage, C., Rieke, G. H., et al. 2023, arXiv e-prints, arXiv:2303.13517, doi: 10.48550/arXiv.2303.13517

Asgari, M., Lin, C.-A., Joachimi, B., et al. 2021, A&A, 645, A104, doi: 10.1051/0004-6361/202039070

Astier, P., Antilogus, P., Juramy, C., et al. 2019, A&A, 629, A36, doi: 10.1051/0004-6361/201935508

Astier, Pierre, & Regnault, Nicolas. 2023, A&A, 670, A118, doi: 10.1051/0004-6361/202245407

ATLAS Collaboration, Aad, G., Abbott, B., et al. 2022, Journal of High Energy Physics, 2022, 104, doi: 10.1007/JHEP08(2022)104

Atwood, W. B., Abdo, A. A., Ackermann, M., et al. 2009, ApJ, 697, 1071, doi: 10.1088/0004-637X/697/2/1071

Balaji, B., Cholis, I., Fox, P. J., & McDermott, S. D. 2018, PhRvD, 98, 043009, doi: 10.1103/PhysRevD.98.043009

Baldi, P. 2021, Deep Learning in Science (Cambridge, UK: Cambridge University Press)

Baldi, P., & Sadowski, P. 2014, Artificial Intelligence, 210C, 78

Ballet, J., Bruel, P., Burnett, T. H., Lott, B., & The Fermi-LAT collaboration. 2023, arXiv e-prints, arXiv:2307.12546, doi: 10.48550/arXiv.2307.12546

Ballet, J., Burnett, T. H., Digel, S. W., & Lott, B. 2020, arXiv e-prints. https://arxiv.org/abs/2005.11208

Bartels, R., Krishnamurthy, S., & Weniger, C. 2016, Phys. Rev. Lett., 116, 051102, doi: 10.1103/PhysRevLett.116.051102

Bartels, R., Storm, E., Weniger, C., & Calore, F. 2018, Nature Astronomy, 2, 819, doi: 10.1038/s41550-018-0531-z

Becker, M. R. 2013, Monthly Notices of the Royal Astronomical Society, 435, 1547, doi: 10.1093/mnras/stt1396

Bernstein, G. M. 2010, MNRAS, 406, 2793, doi: 10.1111/j.1365-2966.2010.16883.x

Bernstein, G. M., & Jarvis, M. 2002, The Astronomical Journal, 123, 583–618, doi: 10.1086/338085

Bertin, E. 2011, in Astronomical Society of the Pacific Conference Series, Vol. 442, Astronomical Data Analysis Software and Systems XX, ed. I. N. Evans, A. Accomazzi, D. J. Mink, & A. H. Rots, 435

Blackwell, R., Burton, M., & Rowell, G. 2017, in The Multi-Messenger Astrophysics of the Galactic Centre, ed. R. M. Crocker, S. N. Longmore, & G. V. Bicknell, Vol. 322, 164–165, doi: 10.1017/S1743921316012035

Bosch, J., Armstrong, R., Bickerton, S., et al. 2018, PASJ, 70, S5, doi: 10.1093/pasj/psx080

Bosch, J., AlSayyad, Y., Armstrong, R., et al. 2019, in Astronomical Society of the Pacific Conference Series, Vol. 523, Astronomical Data Analysis Software and Systems XXVII, ed. P. J. Teuben, M. W. Pound, B. A. Thomas, & E. M. Warner, 521, doi: 10.48550/arXiv.1812.03248

Braiding, C., Wong, G. F., Maxted, N. I., et al. 2018, Publications of the Astronomical Society of Australia, 35, doi: 10.1017/pasa.2018.18

Branch, D., & Tammann, G. A. 1992, ARA&A, 30, 359, doi: 10.1146/annurev.aa.30.090192.002043

Brandt, T. D., & Kocsis, B. 2015, ApJ, 812, 15, doi: 10.1088/0004-637X/812/1/15

Brout, D., Scolnic, D., Popovic, B., et al. 2022, ApJ, 938, 110, doi: 10.3847/1538-4357/ac8e04

Burton, M. G., Braiding, C., Glueck, C., et al. 2013, Publications of the Astronomical Society of Australia, 30

Buschmann, M., Rodd, N. L., Safdi, B. R., et al. 2020, PhRvD, 102, 023023, doi: 10.1103/PhysRevD.102.023023

C. Karwin, A. Broughton, S. Murgia, A. Shmakov, M. Tavakoli, and P. Baldi. 2022, Improved Modeling of the Discrete Component of the Galactic Interstellar γ-ray Emission and Implications for the *Fermi*–LAT Galactic Center Excess: Supplemental Material

Calore, F., Cholis, I., & Weniger, C. 2015, Journal of Cosmology and Astroparticle Physics, 2015, 038, doi: 10.1088/1475-7516/2015/03/038

Calore, F., Donato, F., & Manconi, S. 2021, PhRvL, 127, 161102, doi: 10.1103/PhysRevLett.127.161102

Choi, A., & Hirata, C. M. 2020, PASP, 132, 014502, doi: 10.1088/1538-3873/ab4504

Cirelli, M., Gaggero, D., Giesen, G., Taoso, M., & Urbano, A. 2014, Journal of Cosmology and Astroparticle Physics, 2014, 045, doi: 10.1088/1475-7516/2014/12/045

Cirelli, M., Serpico, P. D., & Zaharijas, G. 2013, Journal of Cosmology and Astroparticle Physics, 2013, 035, doi: 10.1088/1475-7516/2013/11/035

Clemens, D. P. 1985, ApJ, 295, 422, doi: 10.1086/163386

Cormier, D., Bigiel, F., Jiménez-Donaire, M. J., et al. 2018, Monthly Notices of the Royal Astronomical Society, 475, 3909, doi: 10.1093/mnras/sty059

Coulton, W. R., Armstrong, R., Smith, K. M., Lupton, R. H., & Spergel, D. N. 2018, The Astronomical Journal, 155, 258, doi: 10.3847/1538-3881/aac08d

Crenshaw, J. F., Connolly, A. J., Meyers, J. E., et al. 2024, The Astronomical Journal, 167, 86, doi: 10.3847/1538-3881/ad1661

Dame, T. M., Hartmann, D., & Thaddeus, P. 2001, ApJ, 547, 792, doi: 10.1086/318388

Dark Energy Survey Collaboration, Abbott, T., Abdalla, F. B., et al. 2016, MNRAS, 460, 1270, doi: 10.1093/mnras/stw641

Daylan, T., Finkbeiner, D. P., Hooper, D., et al. 2016, Physics of the Dark Universe, 12, 1, doi: https://doi.org/10.1016/j.dark.2015.12.005

de Jong, J. T. A., Verdoes Kleijn, G. A., Erben, T., et al. 2017, A&A, 604, A134, doi: 10.1051/0004-6361/201730747

DES Collaboration, Abbott, T. M. C., Acevedo, M., et al. 2024, arXiv e-prints, arXiv:2401.02929, doi: 10.48550/arXiv.2401.02929

DESI Collaboration, Adame, A. G., Aguilar, J., et al. 2024, arXiv e-prints, arXiv:2404.03002, doi: 10.48550/arXiv.2404.03002

Di Teodoro, E. M., McClure-Griffiths, N. M., Lockman, F. J., et al. 2018, ApJ, 855, 33, doi: 10.3847/1538-4357/aaad6a

Di Valentino, E., Mena, O., Pan, S., et al. 2021, Classical and Quantum Gravity, 38, 153001, doi: 10.1088/1361-6382/ac086d

Downing, M., Baade, D., Sinclaire, P., Deiries, S., & Christen, F. 2006, in SPIE Conference Series, Vol. 6276, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, ed. D. A. Dorn & A. D. Holland, 627609, doi: 10.1117/12.671457

Eden, D. J., Moore, T. J. T., Currie, M. J., et al. 2020, MNRAS, 498, 5936, doi: 10.1093/mnras/staa2734

Esteves, J. H., Utsumi, Y., Snyder, A., et al. 2023

Euclid Collaboration, Mellier, Y., Abdurro'uf, et al. 2024, arXiv e-prints, arXiv:2405.13491, doi: 10.48550/arXiv.2405.13491

Freudenburg, J. K. C., Givans, J. J., Choi, A., et al. 2020, arXiv e-prints, arXiv:2003.05978, doi: 10.48550/arXiv.2003.05978

Gaitskell, R., Barnes, P., DaSilva, A., Sadoulet, B., & Shutt, T. 1996, Nuclear Physics B - Proceedings Supplements, 51, 279, doi: https://doi.org/10.1016/S0920-5632(96)00515-4

Gee, P., Tyson, J. A., Pinto, P., et al. 2007, in American Astronomical Society Meeting Abstracts, Vol. 211, American Astronomical Society Meeting Abstracts, 137.28

Génolini, Y., Maurin, D., Moskalenko, I. V., & Unger, M. 2018, PhRvC, 98, 034611, doi: 10.1103/PhysRevC.98.034611

Giblin, B., Heymans, C., Asgari, M., et al. 2021, A&A, 645, A105, doi: 10.1051/0004-6361/202038850

Gnedin, O. Y., Kravtsov, A. V., Klypin, A. A., & Nagai, D. 2004, ApJ, 616, 16, doi: 10.1086/424914

Goodenough, L., & Hooper, D. 2009, Possible Evidence For Dark Matter Annihilation In The Inner Milky Way From The Fermi Gamma Ray Space Telescope, arXiv, doi: 10.48550/ARXIV.0910.2998

Gordon, C., & Macías, O. 2013, Phys. Rev. D, 88, 083521, doi: 10.1103/PhysRevD.88.083521

Green, M. A. 2022, Progress in Photovoltaics: Research and Applications, 30, 164, doi: https://doi.org/10.1002/pip.3474

Grenier, I. A., Casandjian, J.-M., & Terrier, R. 2005, Science, 307, 1292, doi: 10.1126/science.1106924

Gruen, D., Bernstein, G. M., Jarvis, M., et al. 2015, Journal of Instrumentation, 10, C05032, doi: 10.1088/1748-0221/10/05/C05032

Guy, J., Astier, P., Baumont, S., et al. 2007, A&A, 466, 11, doi: 10.1051/0004-6361:20066930

Guyonnet, A., Astier, P., Antilogus, P., Regnault, N., & Doherty, P. 2015, A&A, 575, A41, doi: 10.1051/0004-6361/201424897

Hamana, T., Hikage, C., Oguri, M., Shirasaki, M., & More, S. 2022, PASJ, 74, 923, doi: 10.1093/pasj/psac046

He, K., Zhang, X., Ren, S., & Sun, J. 2015, CoRR, abs/1502.01852

Hertel, L., Collado, J., Sadowski, P., & Baldi, P. 2018, in Sherpa: Hyperparameter Optimization for Machine Learning Models, Machine Learning Open Source Software 2018: Sustainable communities (openreview.net). https://openreview.net/forum?id=HklSUMyJcQ

Hertel, L., Collado, J., Sadowski, P., Ott, J., & Baldi, P. 2020, SoftwareX

Heymans, C., Van Waerbeke, L., Bacon, D., et al. 2006, MNRAS, 368, 1323, doi: 10.1111/j.1365-2966.2006.10198.x

Heymans, C., Van Waerbeke, L., Miller, L., et al. 2012, MNRAS, 427, 146, doi: 10.1111/j.1365-2966.2012.21952.x

Hilbert, S., Hartlap, J., White, S. D. M., & Schneider, P. 2009, A&A, 499, 31, doi: 10.1051/0004-6361/200811054

Hirata, C., & Seljak, U. 2003, Monthly Notices of the Royal Astronomical Society, 343, 459, doi: 10.1046/j.1365-8711.2003.06683.x

Hirata, C. M., & Choi, A. 2020, PASP, 132, 014501, doi: 10.1088/1538-3873/ab44f7

Hoekstra, H., Franx, M., Kuijken, K., & Squires, G. 1998, ApJ, 504, 636, doi: 10.1086/306102

Holland, S. E., Bebek, C. J., Kolbe, W. F., & Lee, J. S. 2014, Journal of Instrumentation, 9, C03057–C03057, doi: 10.1088/1748-0221/9/03/c03057

Hooper, D., & Goodenough, L. 2011, Phys. Lett. B, 697, 412, doi: 10.1016/j.physletb.2011.02.029

Hooper, D., & Linden, T. 2011, Phys. Rev. D, 84, 123005, doi: 10.1103/PhysRevD.84.123005

Hooper, D., & Slatyer, T. R. 2013, Physics of the Dark Universe, 2, 118, doi: https://doi.org/10.1016/j.dark.2013.06.003

Horiuchi, S., Kaplinghat, M., & Kwa, A. 2016, JCAP, 2016, 053, doi: 10.1088/1475-7516/2016/11/053

Hubble, E. 1929, Proceedings of the National Academy of Science, 15, 168, doi: 10.1073/pnas.15.3.168

Huff, E., & Mandelbaum, R. 2017, arXiv e-prints, arXiv:1702.02600, doi: 10.48550/arXiv.1702.02600

Huterer, D. 2023, A course in cosmology: from theory to practice (Cambridge University Press)

Ioffe, S., & Szegedy, C. 2015, in Proceedings of the 32nd International Conference on International Conference on Machine Learning - Volume 37, ICML'15 (JMLR.org), 448–456

Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, ApJ, 873, 111, doi: 10.3847/1538-4357/ab042c

Janesick, J. R. 2001, Scientific charge-coupled devices (Bellingham: SPIE), 1 online resource (xvi, 906 pages : illustrations), doi: 10.1117/3.374903

Jarvis, M., Meyers, J., Leget, P.-F., & Davis, C. 2021, Piff: PSFs In the Full FOV, Astrophysics Source Code Library, record ascl:2102.024

Jarvis, M., Sheldon, E., Zuntz, J., et al. 2016, Monthly Notices of the Royal Astronomical Society, 460, 2245, doi: 10.1093/mnras/stw990

Johannesson, G., Porter, T. A., & Moskalenko, I. V. 2018, Astrophys. J., 856, 45, doi: 10.3847/1538-4357/aab26e

Jóhannesson, G., et al. 2016, Astrophys. J., 824, 16, doi: 10.3847/0004-637X/824/1/16

Juramy, C., Antilogus, P., Bailly, P., et al. 2014, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9154, High Energy, Optical, and Infrared Detectors for Astronomy VI, ed. A. D. Holland & J. Beletic, 91541P, doi: 10.1117/12.2055175

Jurić, M., Kantor, J., Lim, K. T., et al. 2017, in Astronomical Society of the Pacific Conference Series, Vol. 512, Astronomical Data Analysis Software and Systems XXV, ed. N. P. F. Lorente, K. Shortridge, & R. Wayth, 279, doi: 10.48550/arXiv.1512.07914

Kaiser, N., Squires, G., & Broadhurst, T. 1995, ApJ, 449, 460, doi: 10.1086/176071

Kalberla, P. M. W., Kerp, J., & Haud, U. 2020, A&A, 639, A26, doi: 10.1051/0004-6361/202037602

Karwin, C., Broughton, A., Murgia, S., et al. 2022, Improved Modeling of the Discrete Component of the Galactic Interstellar γ-ray Emission and Implications for the *Fermi*–LAT Galactic Center Excess. https://arxiv.org/abs/2206.02809

Karwin, C., Murgia, S., Tait, T. M., Porter, T. A., & Tanedo, P. 2017, Physical Review D, 95, 103005

Karwin, C. M., Murgia, S., Campbell, S., & Moskalenko, I. V. 2019, The Astrophysical Journal, 880, 95

Kenworthy, W. D., Jones, D. O., Dai, M., et al. 2021, ApJ, 923, 265, doi: 10.3847/1538-4357/ac30d8

Kotov, I. V., Haupt, J., O'Connor, P., et al. 2016, in High Energy, Optical, and Infrared Detectors for Astronomy VII, ed. A. D. Holland & J. Beletic, Vol. 9915, International Society for Optics and Photonics (SPIE), 99150V, doi: 10.1117/12.2231925

Kowalski, M., Rubin, D., Aldering, G., et al. 2008, ApJ, 686, 749, doi: 10.1086/589937

Krause, E., & Hirata, C. M. 2010, A&A, 523, A28, doi: 10.1051/0004-6361/200913524

Lage, C. 2019, arXiv e-prints, arXiv:1911.09567, doi: 10.48550/arXiv.1911.09567

Lage, C., Bradshaw, A., & Tyson, J. 2017, Journal of Instrumentation, 12, C03091, doi: 10.1088/1748-0221/12/03/c03091

Lauer, T. R. 1999, PASP, 111, 1434, doi: 10.1086/316460

Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, arXiv e-prints, arXiv:1110.3193, doi: 10.48550/arXiv.1110.3193

Law, N. M., Mackay, C. D., & Baldwin, J. E. 2006, A&A, 446, 739, doi: 10.1051/0004-6361:20053695

Leach, R. W., Schild, R. E., Gursky, H., et al. 1980, PASP, 92, 233, doi: 10.1086/130654

Leane, R. K., & Slatyer, T. R. 2019, Phys. Rev. Lett., 123, 241101, doi: 10.1103/PhysRevLett.123.241101

Leane, R. K., & Slatyer, T. R. 2020, PhRvL, 125, 121105, doi: 10.1103/PhysRevLett.125.121105

Leane, Rebecca K., S. T. R. 2020, PhRvD, 102, 063019, doi: 10.1103/PhysRevD.102.063019

Leavitt, H. S., & Pickering, E. C. 1912, Harvard College Observatory Circular, 173, 1

Lee, S. K., Lisanti, M., & Safdi, B. R. 2015, JCAP, 2015, 056, doi: 10.1088/1475-7516/2015/05/056

Lee, S. K., Lisanti, M., Safdi, B. R., Slatyer, T. R., & Xue, W. 2016, Phys. Rev. Lett., 116, 051103, doi: 10.1103/PhysRevLett.116.051103

Li, X., Katayama, N., Oguri, M., & More, S. 2018, MNRAS, 481, 4445, doi: 10.1093/mnras/sty2548

Li, X., Mandelbaum, R., Jarvis, M., et al. 2023, arXiv e-prints, arXiv:2309.06506, doi: 10.48550/arXiv.2309.06506

Li, X., Miyatake, H., Luo, W., et al. 2022, Publications of the Astronomical Society of Japan, 74, 421, doi: 10.1093/pasj/psac006

Liaudat, T., Starck, J.-L., Kilbinger, M., & Frugier, P.-A. 2023, arXiv e-prints, arXiv:2306.07996, doi: 10.48550/arXiv.2306.07996

Linden, T., Rodd, N. L., Safdi, B. R., & Slatyer, T. R. 2016, Phys. Rev. D, 94, 103013, doi: 10.1103/PhysRevD.94.103013

List, F., Rodd, N. L., & Lewis, G. F. 2021, PhRvD, 104, 123022, doi: 10.1103/PhysRevD.104.123022

Lopez, M., Marshall, S., Bond, T., et al. 2018, in Ground-based and Airborne Instrumentation for Astronomy VII, ed. C. J. Evans, L. Simard, & H. Takami, Vol. 10702, International Society for Optics and Photonics (SPIE), 107022C, doi: 10.1117/12.2312200

LSST Science Collaboration, Abell, P. A., Allison, J., et al. 2009, arXiv e-prints, arXiv:0912.0201, doi: 10.48550/arXiv.0912.0201

MacCrann, N., Becker, M. R., McCullough, J., et al. 2022, MNRAS, 509, 3371, doi: 10.1093/mnras/stab2870

Macias, O., Gordon, C., Crocker, R. M., et al. 2018, Nature Astronomy, 2, 387, doi: 10.1038/s41550-018-0414-3

Macias, O., Horiuchi, S., Kaplinghat, M., et al. 2019, Journal of Cosmology and Astroparticle Physics, 2019, 042, doi: 10.1088/1475-7516/2019/09/042

Magnier, E. A., Tonry, J. L., Finkbeiner, D., et al. 2018, PASP, 130, 065002, doi: 10.1088/1538-3873/aaaad8

Mandelbaum, R. 2015, Journal of Instrumentation, 10, C05017, doi: 10.1088/1748-0221/10/05/C05017

Mandelbaum, R., Jarvis, M., Lupton, R. H., et al. 2023, The Open Journal of Astrophysics, 6, 5, doi: 10.21105/astro.2209.09253

Mandelbaum, R., Hirata, C. M., Seljak, U., et al. 2005, Monthly Notices of the Royal Astronomical Society, 361, 1287, doi: 10.1111/j.1365-2966.2005.09282.x

Mandelbaum, R., Rowe, B., Armstrong, R., et al. 2015, MNRAS, 450, 2963, doi: 10.1093/mnras/stv781

Mandelbaum, R., Miyatake, H., Hamana, T., et al. 2018, PASJ, 70, S25, doi: 10.1093/pasj/psx130

Massey, R., & Refregier, A. 2005, MNRAS, 363, 197, doi: 10.1111/j.1365-2966.2005.09453.x

Massey, R., Rhodes, J., Leauthaud, A., et al. 2007, ApJS, 172, 239, doi: 10.1086/516599

McDaniel, A., Ajello, M., Karwin, C. M., et al. 2024, PhRvD, 109, 063024, doi: 10.1103/PhysRevD.109.063024

McGaugh, S. S. 2023, Research Notes of the American Astronomical Society, 7, 20, doi: 10.3847/2515-5172/acba9a

McLean, I. S. 2004, in Astrophysics and Space Science Library, Vol. 300, Scientific Detectors for Astronomy, The Beginning of a New Era, ed. P. Amico, J. W. Beletic, & J. E. Belectic, 1–8, doi: 10.1007/1-4020-2527-0_1

Meerburg, P. D., Meyers, J., van Engelen, A., & Ali-Haïmoud, Y. 2016, PhRvD, 93, 123511, doi: 10.1103/PhysRevD.93.123511

Melchior, P., Viola, M., Schäfer, B. M., & Bartelmann, M. 2011, MNRAS, 412, 1552, doi: 10.1111/j.1365-2966.2010.17875.x

Miller, L., Heymans, C., Kitching, T. D., et al. 2013, MNRAS, 429, 2858, doi: 10.1093/mnras/sts454

Mishra-Sharma, S., & Cranmer, K. 2022, PhRvD, 105, 063017, doi: 10.1103/PhysRevD.105.063017

Miyatake, H., Sugiyama, S., Takada, M., et al. 2023, PhRvD, 108, 123517, doi: 10.1103/PhysRevD.108.123517

Moskalenko, I. V., & Strong, A. W. 1998, Astrophys. J., 493, 694, doi: 10.1086/305152

—. 2000, Astrophys. J., 528, 357, doi: 10.1086/308138

Murgia, S. 2020a, Annual Review of Nuclear and Particle Science, 70, 455, doi: https://doi.org/10.1146/annurev-nucl-101916-123029

—. 2020b, Annual Review of Nuclear and Particle Science, 70, 455, doi: 10.1146/annurev-nucl-101916-123029

—. 2020c, Ann. Rev. Nucl. Part. Sci., 70, 455, doi: 10.1146/annurev-nucl-101916-123029

Nakajima, R., & Bernstein, G. 2007, AJ, 133, 1763, doi: 10.1086/511957

Neill, D., Angeli, G., Claver, C., et al. 2014, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9150, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 91500G, doi: 10.1117/12.2056553

Newbry, S. P., Lange, T., Roodman, A., et al. 2018, in Ground-based and Airborne Instrumentation for Astronomy VII, ed. H. Takami, C. J. Evans, & L. Simard (SPIE)

O'Connor, P., Antilogus, P., Doherty, P., et al. 2016, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9915, High Energy, Optical, and Infrared Detectors for Astronomy VII, ed. A. D. Holland & J. Beletic, 99150X, doi: 10.1117/12.2232729

Okura, Y., & Futamase, T. 2018, MNRAS, 479, 4971, doi: 10.1093/mnras/sty1746

Paulin-Henriksson, S., Amara, A., Voigt, L., Refregier, A., & Bridle, S. L. 2008, Astronomy &amp Astrophysics, 484, 67, doi: 10.1051/0004-6361:20079150

Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, ApJ, 517, 565, doi: 10.1086/307221

Petrović, J., Serpico, P. D., & Zaharijas, G. 2015, Journal of Cosmology and Astroparticle Physics, 2015, 023, doi: 10.1088/1475-7516/2015/02/023

Pineda, J. E., Caselli, P., & Goodman, A. A. 2008, ApJ, 679, 481, doi: 10.1086/586883

Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020a, A&A, 641, A6, doi: 10.1051/0004-6361/201833910

Planck Collaboration, Akrami, Y., Arroja, F., et al. 2020b, A&A, 641, A10, doi: 10.1051/0004-6361/201833887

Plazas, A. A., Shapiro, C., Smith, R., Huff, E., & Rhodes, J. 2018, Publications of the Astronomical Society of the Pacific, 130, 065004, doi: 10.1088/1538-3873/aab820

Plazas, A. A., Shapiro, C., Smith, R., Rhodes, J., & Huff, E. 2017, Journal of Instrumentation, 12, C04009, doi: 10.1088/1748-0221/12/04/C04009

Plazas Malagón, A. A., Shapiro, C., Choi, A., & Hirata, C. 2023, arXiv e-prints, arXiv:2310.01920, doi: 10.48550/arXiv.2310.01920

Plazas Malagón, A. A., Waters, C., Broughton, A., et al. 2024, arXiv e-prints, arXiv:2404.14516, doi: 10.48550/arXiv.2404.14516

Pohl, M., Macias, O., Coleman, P., & Gordon, C. 2022, ApJ, 929, 136, doi: 10.3847/1538-4357/ac6032

Porter, T. A., Johannesson, G., & Moskalenko, I. V. 2017, Astrophys. J., 846, 23pp

Ptuskin, V. S., Moskalenko, I. V., Jones, F. C., Strong, A. W., & Zirakashvili, V. N. 2006, Astrophys. J., 642, 902, doi: 10.1086/501117

Radeka, V., Frank, J., Geary, J. C., et al. 2009, Journal of Instrumentation, 03, 03002, doi: 10.1088/1748-0221/4/03/P03002

Rasmussen, A., Guyonnet, A., Lage, C., et al. 2016, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9915, High Energy, Optical, and Infrared Detectors for Astronomy VII, ed. A. D. Holland & J. Beletic, 99151A, doi: 10.1117/12.2234482

Refregier, A. 2003, Monthly Notices of the Royal Astronomical Society, 338, 35, doi: https://doi.org/10.1046/j.1365-8711.2003.05901.x

Refregier, A., & Bacon, D. 2003, MNRAS, 338, 48, doi: 10.1046/j.1365-8711.2003.05902.x

Refregier, A., Kacprzak, T., Amara, A., Bridle, S., & Rowe, B. 2012, MNRAS, 425, 1951, doi: 10.1111/j.1365-2966.2012.21483.x

Rhodes, J., Leauthaud, A., Stoughton, C., et al. 2010, arXiv e-prints, arXiv:1002.1479, doi: 10.48550/arXiv.1002.1479

Rhodes, J., Refregier, A., & Groth, E. J. 2000, ApJ, 536, 79, doi: 10.1086/308902

Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, AJ, 116, 1009, doi: 10.1086/300499

Roodman, A. J., et al. 2018, Proc. SPIE Int. Soc. Opt. Eng., 10705, 107050D, doi: 10.1117/12.2314017

Rowe, B. 2010, Monthly Notices of the Royal Astronomical Society, doi: 10.1111/j.1365-2966.2010.16277.x

Rowe, B., Jarvis, M., Mandelbaum, R., et al. 2014, GalSim: The modular galaxy image simulation toolkit, arXiv, doi: 10.48550/ARXIV.1407.7676

Rubin, D., Aldering, G., Betoule, M., et al. 2023, arXiv e-prints, arXiv:2311.12098, doi: 10.48550/arXiv.2311.12098

Schneider, P., Eifler, T., & Krause, E. 2010, A&A, 520, A116, doi: 10.1051/0004-6361/201014235

Sheldon, E. S., Becker, M. R., Jarvis, M., Armstrong, R., & LSST Dark Energy Science Collaboration. 2023, The Open Journal of Astrophysics, 6, 17, doi: 10.21105/astro.2303.03947

Sheldon, E. S., Becker, M. R., MacCrann, N., & Jarvis, M. 2020, ApJ, 902, 138, doi: 10.3847/1538-4357/abb595

Sheldon, E. S., & Huff, E. M. 2017, ApJ, 841, 24, doi: 10.3847/1538-4357/aa704b

Smith, B. A. 1976, in Charge-Coupled Device Technology and Applications, 135–138

Snyder, A., & Roodman, A. 2020, arXiv e-prints, arXiv:2001.03223, doi: 10.48550/arXiv.2001.03223

Spergel, D., Gehrels, N., Baltay, C., et al. 2015, arXiv e-prints, arXiv:1503.03757, doi: 10.48550/arXiv.1503.03757

Srivastava, N., Hinton, G. E., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. 2014, Journal of Machine Learning Research, 15, 1929

Strong, A. W., & Moskalenko, I. V. 1998, Astrophys. J., 509, 212, doi: 10.1086/306470

Strong, A. W., Moskalenko, I. V., & Ptuskin, V. S. 2007, ARNPS, 57, 285, doi: 10.1146/annurev.nucl.57.090506.123011

Strong, A. W., Moskalenko, I. V., & Reimer, O. 2000, Astrophys. J., 537, 763, doi: 10.1086/309038

Stubbs, C. W. 2013, arXiv e-prints, arXiv:1312.2313, doi: 10.48550/arXiv.1312.2313

Stubbs, C. W., & Tonry, J. L. 2006, ApJ, 646, 1436, doi: 10.1086/505138

Tait, T. 2021. https://indico.cern.ch/event/822029/contributions/4346550/attachments/2245200/3808491/Tait-PPC21.pdf

Tavakoli, M., Agostinelli, F., & Baldi, P. 2021, Neural Networks, 140, 1

The LSST Dark Energy Science Collaboration, Mandelbaum, R., Eifler, T., et al. 2021, The LSST Dark Energy Science Collaboration (DESC) Science Requirements Document, arXiv, doi: 10.48550/ARXIV.1809.01669

Urquhart, J. S., König, C., Giannetti, A., et al. 2018, MNRAS, 473, 1059, doi: 10.1093/mnras/stx2258

Vladimirov, A. E., Digel, S. W., Johannesson, G., et al. 2011, Comput. Phys. Commun., 182, 1156, doi: 10.1016/j.cpc.2011.01.017

Wilson, T. L., Rohlfs, K., & Hüttemeister, S. 2009, Tools of Radio Astronomy, doi: 10.1007/978-3-540-85122-6

Zhang, J., & Komatsu, E. 2011, MNRAS, 414, 1047, doi: 10.1111/j.1365-2966.2011.18436.x

Zhang, T., Mandelbaum, R., & Collaboration, T. L. D. E. S. 2021, Monthly Notices of the Royal Astronomical Society, 510, 1978, doi: 10.1093/mnras/stab3584

Zhang, T., Almoubayyed, H., Mandelbaum, R., et al. 2022, Monthly Notices of the Royal Astronomical Society, 520, doi: 10.1093/mnras/stac3350

Zhang, T., Li, X., Dalal, R., et al. 2023, MNRAS, 525, 2441, doi: 10.1093/mnras/stad1801

Zhong, Y.-M., McDermott, S. D., Cholis, I., & Fox, P. J. 2020, Phys. Rev. Lett., 124, 231103, doi: 10.1103/PhysRevLett.124.231103

Appendix A

Separating the Gas into Galactocentric Annuli

The H_2 gas is placed at Galactocentric radii based on its velocity, assuming uniform circular motion around the Galactic center (GC), described with the rotation curve $V(R)$. The velocity with respect to the local standard of rest ($V_{\rm LSR}$) of a region with Galactocentric distance R, viewed toward Galactic coordinates l, b, is given by Ackermann et al. (2012)

$$V_{\rm LSR} = (\frac{R_\odot}{R} V(R) - V_\odot)\sin(l)\cos(b). \tag{A.1}$$

We use the parametrized rotation curve of Clemens (1985), with $R_\odot = 8.5$ kpc for the distance from the GC to the Sun, and $V_\odot = 220$ km s^{-1} for the velocity of the Sun around the GC (e.g. see Ackermann et al. (2012) and references therein). Since the Mopra data is within 0.5° from the Galactic plane, we use the small angle approximation $\cos(b) \approx 1$. Following the same procedure as Ackermann et al. (2012), we bin the gas in 17 discrete radial bins, as summarized in Table A.1. Figure A.1 shows the local standard of rest velocity as a function of Galactocentric radius, calculated for longitudes between $300° - 350°$. We have

Table A.1: Radial Bins

Bin	$R_{\rm min}$ [kpc]	$R_{\rm max}$ [kpc]	$X_{\rm CO}(\times 10^{-20})$ [cm^{-2} (K km s^{-1})$^{-1}$]
1	0	1.5	0.36
2	1.5	2.0	1.01
3	2.0	2.5	1.04
4	2.5	3.0	1.06
5	3.0	3.5	1.11
6	3.5	4.0	1.09
7	4.0	4.5	1.15
8	4.5	5.0	1.19
9	5.0	5.5	1.2
10	5.5	6.5	1.22
11	6.5	7.0	1.33
12	7.0	8.0	1.4
13	8.0	10.0	0.72
14	10.0	11.5	7.0
15	11.5	16.5	24.55
16	16.5	19.0	131.79
17	19.0	50.0	532.47

verified that the resulting gas distributions for the radial bins are in good agreement with those from Ajello et al. (2016) and Karwin et al. (2019), which use the same radial binning with the CO maps from Dame et al. (2001).

Appendix A. Separating the Gas into Galactocentric Annuli

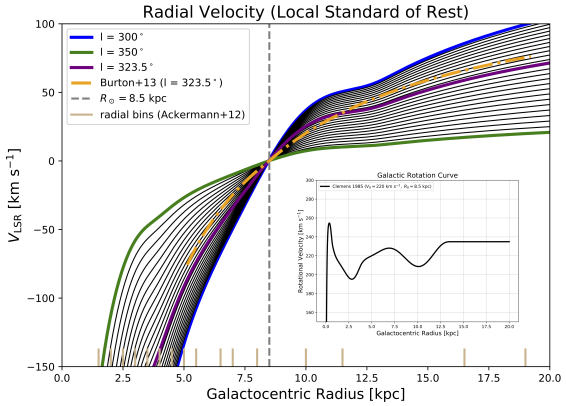


Figure A.1: Local standard of rest velocity (V_{LSR}) as a function of Galactocentric radius, calculated for Galactic longitudes between $300° - 350°$. The calculations are made assuming uniform circular rotation described by the Galactic rotation curve of Clemens (1985), shown in the inset. The Solar radius is assumed to be 8.5 kpc, indicated with the dashed grey line. The purple curve is calculated at a longitude of 323.5°, which is compared to the result from Burton et al. (2013), shown with a dash dot orange curve (note, however, that they use a different Galactic rotation curve, which leads to slight differences in V_{LSR}). The 17 radial bins used in the analysis are shown with the tan bins along the x-axis.

Appendix B

Calculating the H_2 column density

We employ the Mopra data to calculate the corresponding H_2 column density separately for both CO and ^{13}CO, which we label as $N(H_2)_{CO12}$ and $N(H_2)_{CO13}$, respectively. We note that in both cases we correct for the beam efficiency using a value of 1/0.55 Braiding et al. (2018). To calculate $N(H_2)_{CO12}$ we use

$$N(H_2)_{CO12} = W(CO) \times X_{CO}, \tag{B.1}$$

where we employ a radially-dependent conversion factor (X_{CO}) for the 17 radial bins, as determined in Ackermann et al. (2012), based on a maximum likelihood fit to the γ-ray data. The values are given in the last column of Table A.1.

In general, $N(H_2)_{CO13}$ can be calculated from the ^{13}CO column density, N(CO13), and the abundance ratio, $[\frac{H_2}{CO13}]$:

$$N(H_2)_{CO13} = N(CO13) \times \Big[\frac{H_2}{CO13}\Big]. \tag{B.2}$$

We follow the procedure from Wilson et al. (2009); Braiding et al. (2018), which we summarize below. We derive N(CO13) assuming that the gas is in local thermodynamic equilibrium

at a fixed excitation temperature of 10 K:

$$N(\mathrm{CO13}) = \frac{3.0 \times 10^{14}}{1 - \mathrm{e}^{-5.3/T}} \times T \int \tau^{13}(\nu) d\nu, \tag{B.3}$$

where τ^{13} is the ^{13}CO optical depth and 5.3 K is the energy level of the $J = 1 - 0$ transition. T is the line excitation temperature (assumed to be the same as the kinetic temperature of the gas) and here we use T = 10 K. The optical depth is calculated as:

$$\tau^{13} = -\ln\left[1 - \frac{T_{\mathrm{B}}^{13}}{5.3}\left\{\left[\exp\left(\frac{5.3}{T}\right) - 1\right]^{-1} - 0.16\right\}^{-1}\right],$$

where T_{B}^{13} is the line intensity measured by Mopra as a function of velocity.

To approximate a physically viable range for the abundance ratio we use the range of values from Pineda et al. (2008), based on a study of the Perseus Molecular Cloud Complex. Specifically, we estimate that

$$\left[\frac{\mathrm{H_2}}{\mathrm{CO13}}\right] \sim (2.7 - 7.5) \times 10^5. \tag{B.4}$$

.

Appendix C

Predicted γ-Ray Flux and Spectrum for the H_2 traced by Mopra

The CR propagation code GALPROP (v56) calculates the γ-ray sky maps for the H_2-related emission. Here, we employ the GALPROP input parameters described in Karwin et al. (2019) to determine the γ-ray emission related to the H_2 column density $N(H_2)_{CO12}$ traced by Mopra in the 17 radial bins.

Figure C.1 shows the γ-ray spectra corresponding to the input column density $N(H_2)_{CO12}$, summed over the entire sky. As a consistency check, we compare it to the spectra from Karwin et al. (2019), based on the CO maps from Dame et al. (2001). The Mopra data was found to be systematically higher than the CO observations in Dame et al. (2001) by a factor of 1.36 (as discussed in Braiding et al. (2018)), so we include this scaling factor in the comparison. As can be seen, the two spectra are in excellent agreement.

Appendix C. Predicted γ-Ray Flux and Spectrum for the H_2 traced by Mopra

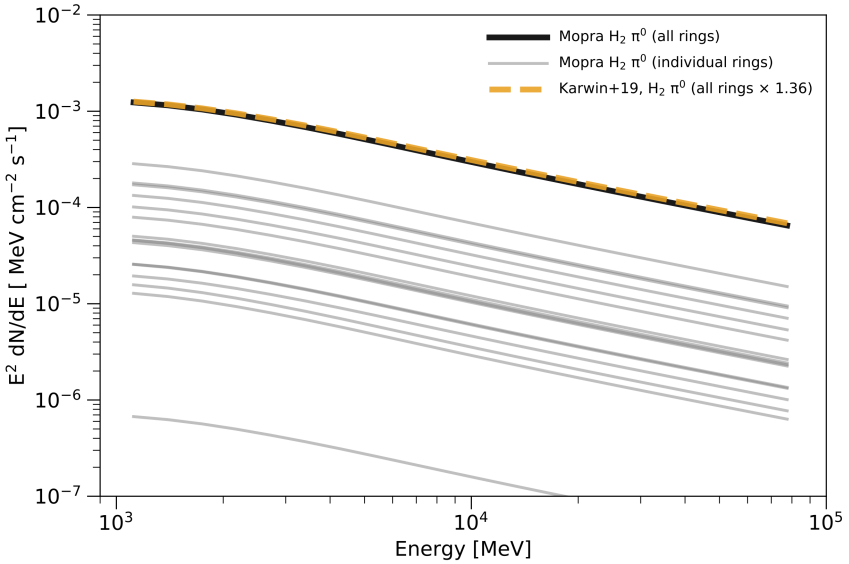


Figure C.1: γ-ray spectra for the H_2-related components summed over the Mopra region. The gray curves are for the individual radial bins, and the black curve is the total model. The dashed yellow curve is the total spectrum from Karwin et al. (2019), scaled by a factor of 1.36, which is the systematic offset of the Mopra data compared to the CO maps in Dame et al. (2001).

Appendix D

Instrument Signature Removal Configurations

In Table D.1 we present the configurations we used with the LSST Science Pipelines for ISR of the spot images (Jurić et al. 2017; Bosch et al. 2019). For spot identification and characterization, we use an LSST Science Pipeline wrapper called the Mixed Calibration Optics Analysis Test Library (MixCOATL). Our data processing also utilized code (with default configurations) for matching and labeling spots to specific holes in the lithographic mask of the optical projector (`GridFitTask`). This code is described in Esteves et al. (2023) and integrated into the LSST Science Pipelines. For each task, including the calibration and ISR tasks, we used the default configuration parameters.

Our selected ISR configurations included subtraction with median-per-row overscan subtraction, dark current subtraction, non-linearity correction, bad-pixel masking, and serial deferred charge correction. Additionally, we corrected the charge transfer inefficiency in the horizontal (serial) direction across the image using the same method as described in Snyder & Roodman (2020) and Astier et al. (2019). We also applied bad-pixel masks and interpolated over the locations of bright and dark pixels and any cosmic ray artifacts, though these

pixels were excluded from calculations of the flat field statistics. We intentionally turned off the saturated pixel repair to avoid any non-physical distortions introduced by the repair.

After this sequence of ISR, we modeled and subtracted the background light produced by the spot grid projector. We performed this background subtraction separately for both the corrected and uncorrected images before calculating and cataloging the final shape statistics even though we ultimately ignored these low signal points in measuring the slope of the brighter-fatter because the background fit was poor in the dimmest exposures. We performed this using a 65×65px median filter to match the artificial star spacing, masking over the fitted footprints of the artificial stars.

Appendix D. Instrument Signature Removal Configurations

Configurations for Instrument Signature Removal

Parameter	Value
config.doSaturation	False
config.growSaturationFootprintSize	0
config.doSuspect	False
config.edgeMaskLevel	'DETECTOR'
config.doOverscan	True
config.overscan.fitType	'MEDIAN'
config.overscan.order	1
config.overscan.numSigmaClip	3.0
config.overscan.maskPlanes	['BAD', 'SAT']
config.overscan.overscanIsInt	True
config.overscan.doParallelOverscan	False
config.overscan.parallelOverscanMaskThreshold	100000
config.overscan.parallelOverscanMaskGrowSize	7
config.doBias	True
config.doBiasBeforeOverscan	False
config.doDeferredCharge	True
config.deferredChargeCorrection.nPixelOffsetCorrection	15
config.deferredChargeCorrection.nPixelTrapCorrection	6
config.deferredChargeCorrection.useGains	False
config.deferredChargeCorrection.zeroUnusedPixels	False
config.doLinearize	True
config.doCrosstalk	False
config.doDefect	True
config.doNanMasking	True
config.doWidenSaturationTrails	True
config.doBrighterFatter	True

Table D.1: The configuration parameters for `lsst.ip.isr.isrTask.IsrTask` in the LSST Science Pipelines (Jurić et al. 2017; Bosch et al. 2019).

Configurations for Instrument Signature Removal

config.doFluxConservingBrighterFatterCorrection	True
config.brighterFatterLevel	'DETECTOR'
config.brighterFatterMaxIter	10
config.brighterFatterThreshold	10.0
config.brighterFatterApplyGain	True
config.brighterFatterMaskGrowSize	0
config.doDark	True
Parameter	Value
config.doStrayLight	False
config.doFlat	False
config.doApplyGains	False
config.usePtcGains	True
config.doFringe	False
config.doAmpOffset	False
config.doInterpolate	False
config.doSaturationInterpolation	False
config.doNanInterpolation	True
config.doVignette	False

Table D.1: Continued.

Appendix E

Image Moments

For shape fitting of each artificial star, we utilized the re-Gaussianization adaptive image moment technique defined by Hirata & Seljak (2003). This method "adaptively" estimates the best-fit elliptical Gaussian to the object by recursively computing the moments using a circular Gaussian weighting function. This elliptical Gaussian is computed recursively by initially guessing a circular Gaussian and using that as a weight function to re-compute the image moments, continuing this process until the image moments converge to the moments of the weight function.

The procedure is performed individually on each object, over a square stamp centered on each detected object's centroid, with a side length equal to the average grid spacing between each object (65px), as shown at the bottom of Figure 7.1.

$$w(\boldsymbol{x}) = \exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}_0)^T\mathbf{I}^{-1}(\boldsymbol{x}-\boldsymbol{x}_0)\right] \tag{E.1}$$

$$\boldsymbol{x}_0 = \frac{\int_{\mathbf{R}^2}\boldsymbol{x}w(\boldsymbol{x})I(\boldsymbol{x})\,\mathrm{d}^2\boldsymbol{x}}{\int_{\mathbf{R}^2}w(\boldsymbol{x})I(\boldsymbol{x})\,\mathrm{d}^2\boldsymbol{x}} \tag{E.2}$$

From these we can calculate the second moments of images for measuring PSF:

$$I_{\mu\nu} = 2 \times \frac{\int_{\mathbf{R}^2} (\boldsymbol{x} - \boldsymbol{x}_0)_\mu (\boldsymbol{x} - \boldsymbol{x}_0)_\nu w(\boldsymbol{x}) I(\boldsymbol{x}) \, \mathrm{d}^2\boldsymbol{x}}{\int_{\mathbf{R}^2} w(\boldsymbol{x}) I(\boldsymbol{x}) \, \mathrm{d}^2\boldsymbol{x}} \tag{E.3}$$

This algorithm is very sensitive to residual sky backgrounds, like those produced by the aberration in the projector lens (Hirata & Seljak 2003; Refregier et al. 2012; Okura & Futamase 2018). For this reason, great care was taken for sky background estimation and removal. The linear combinations of these components then define the unnormalized ellipticity:

$$\begin{aligned} e_0 &= I_{xx} + I_{yy} \\ e_1 &= I_{xx} - I_{yy} \\ e_2 &= 2I_{xy} \end{aligned} \tag{E.4}$$