Cosmology and Astrophysics with Intensity Mapping

Thesis by
Yun-Ting Cheng

In Partial Fulfillment of the Requirements for the
Degree of
Doctor of Philosophy

Caltech

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2022
Defended June 15, 2021

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor Jamie Bock. I have learned a lot from his great vision and most rigorous attitude toward science. He has given me numerous insightful thoughts on my work and provided careful examination of my writings. I also appreciate his great support of all my research and life decisions and for giving me freedom to explore topics I am interested in. I am fortunate to have him as my advisor.

I am sincerely grateful to Tzu-Ching Chang. I would not have gotten this far without her dedication. She has been my best mentor throughout this journey. She has always patiently listened to my struggles in research and life and gave me enormous support and advice in every stage of my graduate study. I also want to express my gratitude to Olivier Doré, who has continuously provided me with helpful research and career advice. It has been my greatest pleasure to work closely with them on many projects throughout the years, and I look forward to our collaboration in the future.

I appreciate all the collaborators in the CIBER and TIME projects. They have taught me valuable knowledge about different aspects of those experiments. I especially want to thank Phil Korngut, Mike Zemcov, Asantha Cooray, Abby Crites, and Matt Bradford for their guidance and advice on my projects.

I also want to thank my collaborators on theory projects, Roland de Putter and Benjamin Wandelt. They introduced me to the fascinating world of theoretical cosmology. I always learn inspiring insights from their deep knowledge in theory and statistics. The fruitful discussions with them have been the most satisfying times in my research life.

I am thankful to the Caltech ObsCos group. I have learned a tremendous amount of knowledge in different aspects of observational cosmology from experts in this group. I am especially grateful to fellow graduate students Alicia Lanz, Jonathan Hunacek, Jason Sun, Richard Feder, and Yen-Yung Chang for the stimulating conversation and feedback on my work.

I would like to thank Chuck Steidel and Phil Hopkins for serving on my committee and providing comments to my work. I also want to express my appreciation to Patrick Shopbell for maintaining the SPIRE cluster and to Sheri Stoll for handling administrative matters. I acknowledge the generous support by the Ministry of Education, Taiwan through the Taiwan-Caltech Scholarship.

I am blessed to be accompanied by many friends in my PhD life at Caltech. I owe special thanks to my awesome housemates Zi-Yu Huang and Chien-I Yang. I also thank Kung-Yi Su, Min-Feng Tu, Allen Shang, Yu-Li Ni, Chien-Han Lin, Hsiao-Yi Chen, Ying-Husan Lin, Yu-An Chen, Tai-Jung Kuo, Yu-Chen Hsu, Po-Hsuan Lin, Steve Lu, Meng-Jhang Fong, Han-Hsin Lin, and many others for their friendships. Last but not least, my greatest gratitude goes to my parents and sister for their unconditional love and support.

ABSTRACT

Intensity mapping (IM) has emerged as a promising technique to probe the large-scale structures and galaxy formation and evolution across cosmic history. As IM measures the aggregate emission from all sources, it can overcome the limitation of conventional detection-based observations, where the emission from diffuse populations and high-redshift faint galaxies cannot be resolved individually. As several IM experiments will come online in the next decade, demand for IM modeling and data analysis strategies has increased. In this thesis, we present a range of analysis techniques, theoretical modeling, and data analysis results related to IM.

In Chapter 2, we aim to answer the question: When should we use IM? We present a formalism to describe both IM and galaxy detection (GD) approaches, and use it to quantify their individual performance when measuring the large-scale structure (LSS). With this formalism, we can identify the scenarios where each approach is advantageous. We also develop a simple metric for determining the optimal strategy to map the LSS with future experiments.

In Chapters 3 and 4, we interrogate methods for improving the line intensity mapping (LIM) analysis. LIM traces the three-dimensional structure of the universe by probing the emission field from a spectral line. One particular challenge for LIM is to separate the target line signals from interloper lines along the line of sight in order to extract the desired cosmological and astrophysical information. Previously proposed methods of line de-blending, such as masking and cross-correlation, rely on the external galaxy tracers, but sometimes a galaxy catalog with sufficient depth and sky coverage does not exist. Therefore, we develop two new methods for performing line de-confusion that do not require any external information. The first method (Chapter 3) uses the distinct shape of large-scale two-dimensional power spectra of signals and interlopers to distinguish the line emission from different redshifts. The second method (Chapter 4) reconstructs the intensity maps of individual lines from LIM data in the phase space, using multiple lines from the same source to identify the source redshift. We show that both of our methods are able to effectively extract desired line signals from the upcoming LIM experiments.

In Chapter 5, we discuss the application of IM for studying the extragalactic background light (EBL), the integrated light from all sources of emission in the universe. Previous studies on the fluctuations of the EBL indicate that the intra-halo light

(IHL) has a significant contribution to the near-infrared EBL. Chapter 5 presents the results on probing the IHL using a stacking analysis of images from the Cosmic Infrared Background Experiment (CIBER). CIBER is a rocket-borne experiment designed to image and perform photometry of the near-infrared EBL. Our results suggest that at $z \sim 0.3$ the IHL comprises a large fraction of light associated with $\sim L_*$ galaxies, implying that the IHL accounts for a non-negligible fraction of the near-infrared cosmic radiation budget.

In Chapter 6, we present a forecast on the EBL constraints with the upcoming SPHEREx mission. We consider cross correlating SPHEREx intensity maps with galaxy catalogs from several current and future surveys. Our model predicts that the EBL spectrum as a function of redshift can be detected from the local universe to the epoch of reionization.

The analysis techniques developed in this thesis can help better extract the information from the IM data; the future IM experiments will extend our current works on investigating the EBL. Therefore, the research in this thesis provides important toolkits and foundations for upcoming IM experiments.

PUBLISHED CONTENT AND CONTRIBUTIONS

[1] Cheng, Y.-T., Chang, T.-C., Bock, J., Bradford, C. M., & Cooray, A. 2016, ApJ, 832, 165, doi: `10.3847/0004-637X/832/2/165`

[2] Cheng, Y.-T., Chang, T.-C., & Bock, J. J. 2020, ApJ, 901, 142, doi: `10.3847/1538-4357/abb023`

[3] Cheng, Y.-T., de Putter, R., Chang, T.-C., & Doré, O. 2019, ApJ, 877, 86, doi: `10.3847/1538-4357/ab1b2b`

[4] Cheng, Y.-T., Arai, T., Bangale, P., et al. 2021, arXiv e-prints, arXiv:2103.03882. `https://arxiv.org/abs/2103.03882`

In all of the papers above, Y.-T. C. participated in the conception of the project, conducted the science analysis and wrote the manuscript.

TABLE OF CONTENTS

Acknowledgements . . . iii
Abstract . . . v
Published Content and Contributions . . . vii
Bibliography . . . vii
Table of Contents . . . vii
List of Illustrations . . . x
List of Tables . . . xiii
Chapter I: Introduction . . . 1
Bibliography . . . 3
Chapter II: Optimally Mapping Large-Scale Structures With Luminous Sources 7
Bibliography . . . 7
  2.1 Introduction . . . 7
  2.2 Formalism . . . 9
  2.3 Toy Model . . . 16
  2.4 Schechter Luminosity Function Model . . . 23
  2.5 Optimal Strategy for IM Experiments . . . 38
  2.6 Example Application: Pixel Size Optimization . . . 44
  2.7 Conclusion . . . 46
  .1 Proving $F_{\delta\delta}^{\rm opt} = F_{\delta\delta}$ . . . 48
  .2 Comparing Linear and Quadratic Terms in the Toy Model $N \gg 1$ Optimal Observable . . . 49
  .3 Explaining $F_{\delta\delta}^{\rm IM} \propto N$ . . . 49
  .4 Different Choice of $\ell_{\rm min}$ . . . 50
  .5 Different Choice of α . . . 50
  .6 Unit Conversion of the Survey Parameters . . . 52
Bibliography . . . 55
Chapter III: Spectral Line De-Confusion in an Intensity Mapping Survey . . . 61
Bibliography . . . 61
  3.1 Introduction . . . 61
  3.2 Power Spectrum Modeling . . . 63
  3.3 MCMC-Based Parameter Inference . . . 68
  3.4 Results . . . 73
  3.5 Discussion . . . 78
  3.6 Conclusion . . . 84
  .1 Power Spectrum Modeling . . . 84
Bibliography . . . 87
Chapter IV: Phase-Space Spectral Line De-Confusion in Intensity Mapping . 92
Bibliography . . . 92
  4.1 Introduction . . . 92

4.2 Mock Light Cone Construction 95
4.3 Methods 98
4.4 Results 105
4.5 Discussion 109
4.6 Conclusion 128
.1 The Matching Pursuit (MP) Algorithm 131
.2 Proving var$(u_{t\gamma}) = \sigma_n^2$ 133
.3 SPHEREx Line Signal Model 133
Bibliography 134
Chapter V: Probing Intra-Halo Light with Galaxy Stacking in CIBER Images 139
Bibliography 139
5.1 Introduction 139
5.2 CIBER Experiment 141
5.3 Data Processing 143
5.4 External Catalogs 148
5.5 Simulation Catalog—MICECAT 149
5.6 Stacking 150
5.7 PSF Modeling 153
5.8 Galaxy Stacking 157
5.9 Modeling the Galaxy Profiles 164
5.10 Results 171
5.11 Discussion 173
5.12 Conclusions 182
.1 Extension and IHL Fraction 183
Bibliography 184
Chapter VI: Cosmic Near-Infrared Background Tomography with SPHEREx Using Galaxy Cross-Correlations 191
Bibliography 191
6.1 Introduction 191
6.2 Formalism 194
6.3 Emission Model 201
6.4 Cross-Correlation Forecast 203
6.5 SPHEREx Intensity Mapping 206
6.6 Redshift Surveys 208
6.7 Results 214
6.8 Discussion 219
6.9 Science Interpretation 225
6.10 Conclusion 227
Bibliography 228

LIST OF ILLUSTRATIONS

Number *Page*

2.1 Optimal observable for $N \ll 1$ toy model 20
2.2 Optimal observable for $N \sim 1$ toy model 21
2.3 Optimal observable for $N \gg 1$ toy model 22
2.4 Normal Schechter function with $\alpha = -1.5$. 25
2.5 $\sigma_{\mathrm{SN}}(\ell)$ with different source densities 27
2.6 $L_{\mathrm{SN}} - N_{\mathrm{eff}}$ relation 29
2.7 Schechter function without instrument noise 30
2.8 Ordering of $\{L_{\mathrm{SN}}, \sigma_L, \ell_*\}$ for each Schechter function regime 32
2.9 Case Ia example 33
2.10 Case Ib example 34
2.11 Case Ic example 36
2.12 Case II example 37
2.13 Fisher information as a function of pixel size 47
.14 Dependence of optimal observable on ℓ_{min} 51
.15 Dependence of optimal observable on α 52
3.1 Isotropic power spectra of [C II] at $z = 6$ and CO interlopers 65
3.2 Projected [C II] power spectra with different μ_{CII} 69
3.3 The “ave-prj” power spectra of [C II] and CO interlopers 70
3.4 [C II] and CO 2D power spectrum templates 71
3.5 MCMC posterior fiducial line deblending case 75
3.6 P_{CII} S/N dependence on P_{CO} 75
3.7 P_{CII} S/N dependence on P_{CII} 76
3.8 The ave-prj power spectrum constraints 77
3.9 The ave-prj P_{CII} S/N as a function of P_n and P_{CO} 78
3.10 MCMC posterior of multiple interloper lines 79
3.11 The ave-prj power spectrum constraints of fitting with S15 template . 80
3.12 MCMC posterior of fitting with S15 template 81
3.13 MCMC posterior for lowest $k^{\parallel}$ mode removed 83
.14 The SFRD of the Planck CIB model 87
4.1 Light cone construction steps 96
4.2 I_* of CO lines 99

4.3 **A** matrix . . . 100

4.4 The redshift cumulated $N_{\rm eff}$ . . . 104

4.5 Visualization of the phase-space reconstruction . . . 106

4.6 Correlation between the true and the reconstructed maps . . . 107

4.7 Average correlation in each broad band . . . 108

4.8 VID of the true and reconstructed maps . . . 110

4.9 Averaged r with SLED variation . . . 111

4.10 VID of reconstructed maps with SLED variation . . . 111

4.11 Averaged r with SLED bias . . . 113

4.12 VID of reconstructed maps with SLED bias . . . 113

4.13 Correlation for the SPHEREx case . . . 114

4.14 Correlation for the SPHEREx case with SLED variation . . . 114

4.15 Averaged r with background subtraction . . . 116

4.16 VID of reconstructed maps with background subtraction . . . 116

4.17 Averaged r with external catalog . . . 117

4.18 Correlation for the case without interlopers . . . 118

4.19 S/N on the cross-power spectrum . . . 121

4.20 Correlation between CO map and total observed map . . . 122

4.21 S/N on [C II] shot-noise power spectrum . . . 127

.22 Illustration of MP algorithm . . . 132

5.1 Images from the SWIRE field in the 1.1 μm band . . . 144

5.2 Images from the SWIRE field in the 1.8 μm band . . . 144

5.3 PSF construction and validation . . . 155

5.4 PSF validation with star stacking . . . 156

5.5 PSF reconstruction . . . 157

5.6 Properties of stacking galaxy samples . . . 160

5.7 An example stacked galaxy profile . . . 162

5.8 Stacking profile and the best fit model . . . 165

5.9 Excess profile and the best fit model . . . 165

5.10 Illustration of the components in stacking . . . 168

5.11 Marginalized constraints on galaxy profile and clustering . . . 171

5.12 Constraints on galaxy profile and clustering . . . 174

5.13 Fraction of flux between 10/20 and 100 kpc . . . 177

5.14 The fraction of EBL intensity from galaxy extension . . . 178

5.15 The IHL fraction $f_{\rm IHL}$ as a function of halo mass . . . 180

5.16 The IHL fraction $f_{\rm IHL}$ as a function of redshift . . . 180

5.17 The ratio of the total one-halo term and stacked galaxy profile term . 181
6.1 Redshift-dependent IGL spectra from our model 207
6.2 SPHEREx spectral resolution and the noise level 209
6.3 Number density of the spectroscopic and photometric galaxy catalogs 210
6.4 An examples of cross power spectra 214
6.5 Constraints on the amplitude of cross power spectra of SPHEREx and galaxy surveys on large scales at $z \leqslant 3$ 215
6.6 Constraints on the amplitude of cross power spectra of SPHEREx and galaxy surveys on large scales at $4 < z < 10$ 217
6.7 Constraints on the cross power spectrum shot noise amplitude at $z \leqslant 3$ 218
6.8 Constraints on the cross power spectrum Poisson noise amplitude for $4 < z < 10$ 219
6.9 Cross power spectra error dependence on redshift 220
6.10 Cross power spectra error dependence on tracer density 221
6.11 Cross power spectra error dependence on mask 222
6.12 Cross power spectrum with different masking thresholds 223
6.13 Average quasar's spectrum constraints from cross correlation 226

LIST OF TABLES

Number *Page*
2.1 Survey targets and their expected σ_L, ℓ_*, and $L_{\rm SN}$ relation. 43
4.1 Frequencies and redshifts of the six defined broad bands 99
5.1 CIBER observing fields 142
5.2 Summary of the properties on each stacked galaxy sub-sample 161
5.3 Parameter constraints from stacking 172
5.4 Fraction of flux in core component compared to flux captured in Petrosian and SDSS model flux 175
5.5 Color ($m_{1.1} - m_{1.8}$) of the galaxy inner and outer components 182
.6 Fraction of galaxy flux between 10/20 kpc and 100 kpc 184
.7 IHL fraction with $r_{\rm cut=}$ 10/20 kpc 184
6.1 Summary of the sky area and redshift coverage of the surveys 208

Chapter 1

INTRODUCTION

How did the universe begin? How did we get here? These are the key questions in cosmology, and the answers require a thorough study of the cosmological model that governs the formation and evolution of the cosmic large-scale structure (LSS). Observations of the cosmic microwave background (CMB), the relic radiation from $\sim 400,000$ years after the Big Bang, have laid a major cornerstone for our standard cosmological model, known as the ΛCDM framework [3, 4, 16, 22, 26, 27]. On the other hand, decades of observations across the whole electromagnetic spectrum [e.g., 7, 33, 34] enable us to study the properties of stars, galaxies, and gas across cosmic time, as well as the underlying LSS they trace. More advances in the near future are expected from next-generation telescopes (JWST [14], Rubin Observatory [23], Euclid [1], Roman Space Telescope [31], etc.). However, some key questions in cosmology and astrophysics still remain unresolved. For example, where are the missing baryons in our universe? What type of sources were driving the cosmic reionization at approximately one billion years after the Big Bang? The reason that these questions are unresolved is in part limited by the fact that most of the current instruments are designed for detecting individual sources, and hence makes the faint populations largely unexplored. However, the faint sources in our universe are essential for building a comprehensive picture of the LSS and galaxy evolution, especially for studying the high-redshift universe and the epoch of reionization.

Intensity mapping (IM) is a novel technique to overcome this limitation from conventional detection-based measurements. Rather than resolving individual galaxies, IM probes the aggregate light from all sources to statistically characterize both the source properties and the underlying LSS. Therefore, IM promises to open up a new window on studying the high-redshift universe and the diffuse radiation, where individual sources are difficult to resolve. IM has attracted great interest in the cosmology and astrophysics communities in recent years as a number of IM experiments covering a wide range of spectrum and redshifts have been funded, and will start their observing campaign within the next few years (e.g., SPHEREx [11, 12], HERA[8], TIME [6], COMAP [5]).

IM uses the aggregate emission as a proxy for the underlying density field, whereas the conventional galaxy detection (GD) resolves individual galaxies as tracers of the matter distribution. Understanding which approach to use in a given situation is important for optimizing the LSS measurements. In Chapter 2, we develop a formalism to quantify the performance of both approaches when measuring the density field. Formally, GD and IM can be described by different weighting of the observed data, and we extend this representation beyond the GD and IM dichotomy by introducing the quantity, "optimal observable," which can describe the best weighting function to extract the underlying density field. With this formalism, we identify regimes where the IM or GD approach is advantageous and discuss optimal strategies for different survey setup.

Line intensity mapping [LIM; 19] is an emerging type of IM. LIM maps the LSS in three dimensions by measuring the emission field from individual spectral lines (e.g., 21 cm, Lyα, Hα, [C II], CO), and infers the line of sight distance of the emitting sources from the frequency-redshift relation. However, as IM aims to measure the aggregate light, the observed data is a mixture of emission from all redshifts and foreground contaminations. Another pressing issue for LIM is line-blending, where spectral lines from sources at different redshifts can end up mixed in the same observed frequencies. Previously proposed methods for line de-blending, such as masking and cross-correlation, rely on the external galaxy tracers, but sometimes a galaxy catalog with sufficient depth and sky coverage does not exist for LIM surveys. Therefore, we develop two novel line de-blending techniques for LIM, which do not require any external dataset. In Chapter 3, we use the distinct shape of large-scale two-dimensional power spectra of each line to distinguish the line emission from different redshifts. By fitting power spectrum templates of each line to simulate LIM data cubes that contain multiple emission lines, we show that the large-scale power spectrum of each line can be successfully extracted without the help of external information. In Chapter 4, we present another method to reconstruct the intensity field of individual lines in map space from LIM data, using the fact that if multiple spectral lines of a source are observable in multiple frequencies, we can fit the LIM data to a set of spectral templates to identify the redshift of sources. We show that this method can reconstruct the intensity maps of individual line emission for the upcoming LIM experiments TIME [6] and SPHEREx [11, 12].

Another useful application of IM is to study the extragalactic background light (EBL), the integrated emission from all sources throughout cosmic history. In the near-infrared wavelengths, the EBL is mostly produced by redshifted ultraviolet and optical stellar emissions, and thus it probes the history of stellar synthesis processes in our universe. Previous EBL studies report potential excess emission [2, 20, 21, 24, 25, 29, 30, 32, 35] in the near-infrared EBL over the integrated galaxy light derived from deep galaxy counts [10, 13, 15, 17, 18, 28]. These studies indicate the existence of faint, diffuse populations in the EBL that have not yet been resolved by current galaxy surveys. Zemcov et al. [36] performed a fluctuation analysis using imaging data from the *Cosmic infrared background experiment* (CIBER), and their results suggest that the excess signal arises from the intra-halo light (IHL). The IHL is the emission from stars tidally stripped from their host galaxies and becomes redistributed to the dark matter halo during the galaxy merger. In Chapter 5, we further investigate the IHL using a stacking analysis with CIBER images. Stacking directly probes the relationship between individual galaxies and their stellar halos, which complements fluctuation measurements that only characterize the IHL in a statistical manner. By stacking $\sim 30,000$ galaxies with stellar masses spanning $M_* \sim 10^{10.5}$ to $10^{12} M_\odot$ at $z \sim 0.3$, we find $\sim 30\%/15\%$ of the total galaxy light budget from galaxies is at radius $r > 10/20$ kpc, respectively. These results imply that the IHL has a significant contribution to the EBL intensity and its fluctuations.

In the near future, NASA's next Medium-Class Explorer mission, SPHEREx [11, 12], will be launched in 2024 and will carry out an all-sky spectro-imaging survey at near infrared. One of the three main science goals for SPHEREx is to explore the origin and evolution of galaxies through IM. With its unparalleled spectral resolution and sensitivity, SPHEREx data will be exceptional for studying the EBL. In Chapter 6, we build a model for the expected EBL signal from SPHEREx, and forecast the constraints on the EBL emission across cosmic time by cross-correlating with the galaxy samples from several current and future galaxy surveys (SDSS [7], DESI [9], Rubin Observatory [23], Roman Space Telescope [31], Euclid [1]). According to our model, we predict that the EBL spectrum can be detected to a few σ levels from the present day to the epoch of reionization. This measurement will greatly enhance our understanding on those currently unresolved populations in our universe, including the diffuse IHL and the faint galaxies from the epoch of reionization.

References

[1] Amendola, L., Appleby, S., Avgoustidis, A., et al. 2018, Living Reviews in Relativity, 21, 2, doi: `10.1007/s41114-017-0010-3`

[2] Bernstein, R. A. 2007, ApJ, 666, 663, doi: `10.1086/519824`

[3] Bianchini, F., Wu, W. L. K., Ade, P. A. R., et al. 2020, ApJ, 888, 119, doi: 10.3847/1538-4357/ab6082

[4] BICEP2 Collaboration, Keck Array Collaboration, Ade, P. A. R., et al. 2018, Phys. Rev. Lett., 121, 221301, doi: 10.1103/PhysRevLett.121.221301

[5] Cleary, K., Bigot-Sazy, M.-A., Chung, D., et al. 2016, in American Astronomical Society Meeting Abstracts, Vol. 227, American Astronomical Society Meeting Abstracts #227, 426.06

[6] Crites, A. T., Bock, J. J., Bradford, C. M., et al. 2014, in Proc. SPIE, Vol. 9153, Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy VII, 91531W, doi: 10.1117/12.2057207

[7] Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, AJ, 145, 10, doi: 10.1088/0004-6256/145/1/10

[8] DeBoer, D. R., Parsons, A. R., Aguirre, J. E., et al. 2017, PASP, 129, 045001, doi: 10.1088/1538-3873/129/974/045001

[9] DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 2016, arXiv e-prints, arXiv:1611.00036. https://arxiv.org/abs/1611.00036

[10] Domínguez, A., Primack, J. R., Rosario, D. J., et al. 2011, MNRAS, 410, 2556, doi: 10.1111/j.1365-2966.2010.17631.x

[11] Doré, O., Bock, J., Ashby, M., et al. 2014, ArXiv e-prints. https://arxiv.org/abs/1412.4872

[12] Doré, O., Werner, M. W., Ashby, M. L. N., et al. 2018, arXiv e-prints, arXiv:1805.05489. https://arxiv.org/abs/1805.05489

[13] Driver, S. P., Andrews, S. K., Davies, L. J., et al. 2016, ApJ, 827, 108, doi: 10.3847/0004-637X/827/2/108

[14] Gardner, J. P., Mather, J. C., Clampin, M., et al. 2006, Space Sci. Rev, 123, 485, doi: 10.1007/s11214-006-8315-7

[15] Helgason, K., Ricotti, M., & Kashlinsky, A. 2012, ApJ, 752, 113, doi: 10.1088/0004-637X/752/2/113

[16] Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, ApJS, 208, 19, doi: 10.1088/0067-0049/208/2/19

[17] Keenan, R. C., Barger, A. J., Cowie, L. L., & Wang, W. H. 2010, ApJ, 723, 40, doi: 10.1088/0004-637X/723/1/40

[18] Koushan, S., Driver, S. P., Bellstedt, S., et al. 2021, arXiv e-prints, arXiv:2102.12323. https://arxiv.org/abs/2102.12323

[19] Kovetz, E. D., Viero, M. P., Lidz, A., et al. 2017, ArXiv e-prints. https://arxiv.org/abs/1709.09066

[20] Lauer, T. R., Postman, M., Weaver, H. A., et al. 2020, arXiv e-prints, arXiv:2011.03052. https://arxiv.org/abs/2011.03052

[21] Levenson, L. R., Wright, E. L., & Johnson, B. D. 2007, ApJ, 666, 34, doi: 10.1086/520112

[22] Louis, T., Grace, E., Hasselfield, M., et al. 2017, JCAP, 2017, 031, doi: 10.1088/1475-7516/2017/06/031

[23] LSST Science Collaboration, Abell, P. A., Allison, J., et al. 2009, arXiv e-prints, arXiv:0912.0201. https://arxiv.org/abs/0912.0201

[24] Matsumoto, T., Kim, M. G., Pyo, J., & Tsumura, K. 2015, ApJ, 807, 57, doi: 10.1088/0004-637X/807/1/57

[25] Matsuura, S., Arai, T., Bock, J. J., et al. 2017, ApJ, 839, 7, doi: 10.3847/1538-4357/aa6843

[26] Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, A&A, 641, A1, doi: 10.1051/0004-6361/201833880

[27] POLARBEAR Collaboration, Ade, P. A. R., Aguilar, M., et al. 2017, ApJ, 848, 121, doi: 10.3847/1538-4357/aa8e9f

[28] Saldana-Lopez, A., Domínguez, A., Pérez-González, P. G., et al. 2020, arXiv e-prints, arXiv:2012.03035. https://arxiv.org/abs/2012.03035

[29] Sano, K., Kawara, K., Matsuura, S., et al. 2015, ApJ, 811, 77, doi: 10.1088/0004-637X/811/2/77

[30] Sano, K., Matsuura, S., Yomo, K., & Takahashi, A. 2020, ApJ, 901, 112, doi: 10.3847/1538-4357/abad3d

[31] Spergel, D., Gehrels, N., Baltay, C., et al. 2015, arXiv e-prints, arXiv:1503.03757. https://arxiv.org/abs/1503.03757

[32] Tsumura, K., Matsumoto, T., Matsuura, S., Sakon, I., & Wada, T. 2013, PASJ, 65, 121, doi: 10.1093/pasj/65.6.121

[33] Williams, R. E., Blacker, B., Dickinson, M., et al. 1996, AJ, 112, 1335, doi: 10.1086/118105

[34] Wootten, A., & Thompson, A. R. 2009, IEEE Proceedings, 97, 1463, doi: 10.1109/JPROC.2009.2020572

[35] Zemcov, M., Immel, P., Nguyen, C., et al. 2017, Nature Communications, 8, 15003, doi: 10.1038/ncomms15003

[36] Zemcov, M., Smidt, J., Arai, T., et al. 2014, Science, 346, 732, doi: 10.1126/science.1258168

Chapter 2

OPTIMALLY MAPPING LARGE-SCALE STRUCTURES WITH LUMINOUS SOURCES

Cheng, Y.-T., de Putter, R., Chang, T.-C., & Doré, O. 2019, ApJ, 877, 86, doi: 10.3847/1538-4357/ab1b2b

Intensity mapping has emerged as a promising tool to probe the three-dimensional structure of the universe. The traditional approach of galaxy redshift surveys is based on individual galaxy detection, typically performed by thresholding and digitizing large-scale intensity maps. By contrast, intensity mapping uses the integrated emission from all sources in a 3D pixel (or voxel) as an analog tracer of large-scale structure. In this work, we develop a formalism to quantify the performance of both approaches when measuring large-scale structures. We compute the Fisher information of an arbitrary observable, derive the optimal estimator, and study its performance as a function of source luminosity function, survey resolution, instrument sensitivity, and other survey parameters. We identify regimes where each approach is advantageous and discuss optimal strategies for different scenarios. To determine the best strategy for any given survey, we develop a metric that is easy to compute from the source luminosity function and the survey sensitivity, and we demonstrate the application with several planned intensity mapping surveys.

2.1 Introduction

Studying the large-scale structure (LSS) of the universe is a major focus in cosmology. The initial conditions of the LSS have been well characterized from the cosmic microwave background (CMB) measurements [e.g., 60, 61], and powerful constraints on the cosmological parameters have been inferred from its measurement. Nevertheless, to map LSS at late time is an essential cosmological probe, in particular regarding the properties of dark matter and dark energy. By detecting a large number of individual galaxies as tracers of the underlying density field, one can map out the large-scale matter distribution and infer powerful cosmological constraints from its power spectrum, for example. This galaxy detection (GD) approach has been successfully demonstrated by several major observational programs such as 2dF [19], 6dF [38], WiggleZ [58], VIMOS [35], SDSS [84], and BOSS [25].

Upcoming galaxy surveys are expected to provide further unparalleled cosmological insights, e.g., eBOSS [26], DESI [27], PFS [76], Euclid [43], LSST [51], WFIRST [73], and SPHEREx [28].

At higher redshift, GD becomes difficult, as galaxies at earlier times are on average fainter, and the increased distance reduces the observed flux. As a result, to detect a given number of galaxies at high redshift requires a longer integration time. This has in part motivated the development of intensity mapping (IM) as an alternative technique to probe LSS. Without thresholding to identify individual sources, IM traces the underlying density field using the integrated light emission from all the sources, including unresolved faint galaxies (see Kovetz et al. [42] for a recent review). In addition, line intensity mapping probes the three-dimensional structure by mapping the emission of a particular spectral line and uses the frequency-redshift relation to infer the matter distribution along the line of sight. The 21cm hyperfine emission from neutral hydrogen [14, 52, 68, 83], the CO rotational lines [6–8, 10, 17, 30, 33, 39, 40, 46, 47, 54, 63, 66, 81], the [C II] 157.7 μm fine-structure line [30, 32, 70, 79, 85], and the Lyman-α emission line [20, 23, 24, 30, 34, 64, 71] are amongst the most studied lines in the IM regime.

Although the measurement can be challenged by the presence of continuum foregrounds [e.g., 4, 15, 31, 50, 56, 59, 74] and line interlopers [16, 48], it is still anticipated that line intensity mapping can provide an efficient path to access the faint, high-redshift Universe owing to its relatively low requirement on spatial resolution and sensitivity, which enables the use of small apertures to efficiently scan a large comoving volume.

The first measurement of IM signals from LSS used the 21 cm line. The detection was made in cross-correlation with spectroscopic galaxy catalog [1, 13, 55], and auto-power spectrum constraints have been reported in Switzer et al. [75]. Pullen et al. [63] made the first attempt at measuring CO IM signal in cross-correlation but detected no signal. The COPSS II experiment measured the CO auto-power spectrum at z$\sim$3 [39]. A tentative [C II] measurement has been made by Pullen et al. [65] in cross-correlation. While Croft et al. [24] reported a first detection of Lyα emission in the IM regime by cross-correlating SDSS spectra with a quasar sample, a new analysis in Croft et al. [23] using cross-correlation with both quasars and Lyα forest showed no detection of diffuse Lyα emission.

Formally, the main difference between GD and IM resides in the "weighting" of the observed data. In GD, the universe is digitized into a binary map where detected galaxies have a weight of one, and zero elsewhere. This is essentially giving a uniform weight to all the detected sources, regardless of their flux. On the contrary, IM is a linear mapping between the universe and the data, weighted by the observed intensity. These two different options are suitable for gleaning more information from the data in two extreme regimes: GD is ideal in the high spatial/spectral resolution and deep integration limit, where detected sources are less susceptible to the effects of noise and confusion; IM is ideal if the individual voxel intensity is composed of highly confused sources with a non-negligible noise component.

In this work, we formally explore this dichotomy by introducing an "observable," $\hat{O}$, and quantify the information that can be extracted using this observable for a given survey using the Fisher information formalism. The GD and IM approaches represent two special cases of $\hat{O}$. We define an "optimal observable" that optimizes the information extraction, not necessarily limited to the usual GD or IM approaches. We further develop a simple diagnostic to evaluate the two strategies (e.g. GD or IM) for a survey. We then apply this method to optimize survey design for future experiments and, as an example, optimize the pixelization of intensity maps considering two different noise levels.

This paper is organized as follows. We first introduce our mathematical formalism in Sec. 2.2 before discussing two toy models within this formalism in Sec. 2.3. Scenarios with a more realistic model based on the Schechter luminosity function model are presented in Sec. 2.4. We then follow with two applications of our framework: we determine the optimal observable for several future planned surveys in Sec. 2.5, and we optimize the survey pixel size in Sec. 2.6. The conclusions are given in Sec. 2.7.

2.2 Formalism

A major goal of large-scale galaxy or intensity mapping surveys is to use emission from luminous sources to trace the underlying density field. In particular, we are interested in the *matter* overdensity field $\delta(x) \equiv (\rho(x)-\bar{\rho})/\bar{\rho}$, where $\rho(x)$ is the local matter density and $\bar{\rho}$ its mean on large scales, from which cosmological information can be extracted (e.g., using the power spectrum statistics). We can use luminous sources to learn about δ because, on large scales, the overdensity of a sample of galaxies is a linearly biased tracer of the underlying matter density. In other words, neglecting stochastic noise, on large scales we have

$$\delta_g(x) \equiv (n_g(x) - \bar{n}_g)/\bar{n}_g = b\,\delta(x), \tag{2.1}$$

where $n_g(x)$ is the number density of a sample of galaxies at position x, $\bar{n}_g$ its global mean, and b the galaxy bias.

However, we do not observe n_g directly, but the light emitted by galaxies. For a wide range of survey scenarios, we simply have access to the observed fluxes L in many pixels or voxels, typically small in comparison to the large-scale overdensity modes of interest. These fluxes may include contributions from multiple luminous sources. The question we will tackle is how to optimally extract δ from this "data cube" composed of these small pixels/voxels.

The terms "pixel" and "voxel" above respectively refer to a spatial 2D resolution element or a spatial-spectral 3D resolution volume element. Voxels are the data element in 3D line intensity mapping. A voxel volume can be written as $V_{\rm vox} \propto \Omega_{\rm pix}\,\Delta\nu$, where $\Omega_{\rm pix}$ is the solid angle of the angular size of a voxel and $\Delta\nu$ is the wavelength or frequency width. $\Omega_{\rm pix}$ and $\Delta\nu$ are usually chosen to be of the order of the survey point spread function (PSF; or beam size) and spectral resolution, respectively. However, the analysis in this work is not necessarily limited to the original voxel configuration of a given survey, as voxel size can always be increased by rebinning.

For simplicity, we assume that every source in the surveys fills in at most a single voxel, i.e., all the flux from a given source is measured in only one voxel, so that the correlation between voxels only arises from the underlying cosmological signal, i.e., source clustering. This assumption requires that the voxel size be at least a few times larger than the PSF (beam) size and the size of the sources themselves. Likewise, in the spectral dimension, we require the voxel size to be larger than a few times of the spectral resolution and the target line width. Alternatively, the analysis in this work also applies to 2D imaging of a single frequency band. In this case, a 3D voxel reduces to a 2D pixel, and we also require the pixel size to be a few times larger than the beam size.

Observables

To extract information about the underlying cosmological matter overdensity, we consider a general "observable function," $\mathcal{O}(L)$, serving as a weight function turning the observed map of voxel fluxes[1] $\hat{L}$ into a transformed "observable map" with

---

[1] The unit of flux L in each voxel is power per area, in [W m^{-2}] (or [photons s^{-1} m^{-2}]). L is an "extensive" quantity under this definition, i.e. its value is scaled with the voxel size. Furthermore, later in the paper we will directly compare L with the intrinsic luminosity (in units of W or $L_\odot$) of the sources ℓ. In this case, we implicitly assume that ℓ has been converted to the flux $\ell/4\pi D_L^2$ such

values $\hat{O} \equiv O(\hat{L})$ in each voxel[2]. The power spectrum of this new $O(\hat{L})$ map is then computed as a proxy for the underlying overdensity field matter density power spectrum.

As an alternative way of thinking about how the voxel map can be used to constrain the large-scale matter overdensity, we consider a region that is small compared to the matter overdensity long-wavelength modes of interest so that, in this region, the δ of the long-wavelength modes is nearly uniform (i.e., it can be treated as a "DC mode"). We can further assume the voxel scale to be much smaller than the scale of the long-wavelegth cosmological modes of interest, so we may choose our local region such that it still contains a large number of voxels. In this picture, the way the local overdensity δ is constrained is using the sum (or average) of the values of $\hat{O}$ in the voxels in the local region.

In our context, the quantity of interest is the "large-scale" rather than "total" density field. In principle, each voxel traces the "total" underlying density field, $\delta_{\rm tot}$, which is composed of both large- and small-scale fluctuations: $\delta_{\rm tot} = \delta_L + \delta_S$. Here we only constrain δ_L through the average value of the observable $\hat{O}$ over a large number of voxels living in approximately the same local δ_L, i.e. we use an ensemble average of $\langle\hat{O}\rangle$, not individual voxel measurement $\hat{O}$, to trace the large-scale fluctuation δ_L. Since δ_L does not refer to a specific scale of fluctuation, this argument applies to any modes that have a fluctuation scale much greater than the voxel size. We will thus write δ instead of δ_L from now on, but the readers should keep in mind that the δ we discuss in this work does not include small-scale fluctuations δ_S.

GD and IM represent two special cases of such a mapping $O(L)$. For GD, a voxel is labeled as a "detection" if it is brighter than a threshold luminosity $L_{\rm th}$ (say, five times the noise rms for a 5σ detection). A power spectrum can then be calculated with this "digital map" that consists of ones (detection) and zeros (nondetection) with a proper normalization. Therefore, $O(L)$ in this case is a step function at $L_{\rm th}$,

$$O^{\rm GD}(L) = \begin{cases} 1 & \text{if } L > L_{\rm th} \\ 0 & \text{if } L \leq L_{\rm th}. \end{cases} \tag{2.2}$$

On the contrary, IM directly calculates a power spectrum of the measured intensity (or luminosity) map, so the observable is a linear function of L (the trivial, identity

---

that the two quantities are in the same units.

[2] Throughout the paper, we use the hat notation as a specific realization of the quantity. Thus, L is a variable, while $\hat{L}$ refers to a specific realization of L. Likewise, $O(L)$ refers to function O with variable L, and $\hat{O}$ is the function value at $L = \hat{L}$.

map),

$$O^{\mathrm{IM}}(L) = L. \tag{2.3}$$

While the observed fluxes contain a wealth of additional information (for instance, on galaxy evolution and small-scale clustering), we focus our study on how to optimally extract the underlying cosmological matter overdensity δ. Let's consider a fixed realization of the overdensity δ in some region containing many voxels. A given choice of observable $O(L)$ leads to a noisy estimate of the local value of δ, where the noise is due to the shot noise in the source population used as density tracers and to the instrumental noise. In practice, we will aim at minimizing the combined noise. Our final goal is to measure the large-scale power spectrum of the observable map $\hat{O}$. Uncertainties in the power spectrum contain a cosmic variance component (signal), due to the variance in the underlying matter overdensity δ, and a stochastic/shot-noise component, which is given by how well the observed fluxes from luminous tracers measure the underlying cosmological clustering. By minimizing the noise in the local determination of δ, we minimize the stochastic noise power spectrum relative to the cosmic variance part of the power spectrum, which is the signal of interest.

We will quantify the maximum information content of δ by its Fisher information. We will show that there exists an "optimal observable" O^{opt} such that this observed map contains the same amount of information as the Fisher information. The functional form of this optimal observable depends on the voxel luminosity probability density function (pdf), and we detail its derivation in Sec. 2.2 before describing in Sec. 2.2 the Fisher information and optimal observable.

Voxel Luminosity pdf

The voxel luminosity pdf $P(L, \delta)$ is defined as the probability of a voxel residing in an overdensity field δ with a luminosity between $[L, L + dL]$. This can be computed by the $P(D)$ analysis presented in Lee et al. [45]. First, we define $P_k(L, \delta)$ to be the probability of the voxel with luminosity between $[L, L + dL]$ given that there are k sources in that voxel. The $P(L, \delta)$ is the summation of all the $P_k(L, \delta)$ weighted by the probability of occurrence of each k. If the sources are uncorrelated, the weight function is a Poisson distribution, and thus

$$P(L, \delta) = \sum_{k=0}^{\infty} \frac{e^{-N(\delta)}\, N^k(\delta)}{k!}\, P_k(L, \delta), \tag{2.4}$$

where $N(\delta)$ is the expectation value of the number of sources in a voxel with overdensity δ. The clustering effects can be accounted for by modifying the Poisson term in Eq. 2.4, for example, the approaches presented in Breysse et al. [5]. For simplicity in this work, we only adopt the Poisson distribution in the $P(L)$ function, and we leave the consideration of clustering to future work.

$N(\delta)$ and $P_k(L,\delta)$ can be derived for any given luminosity function[3] $\Phi(\ell,\delta)$ and voxel volume $V_{\rm vox}$,

$$N(\delta) = V_{\rm vox}\int \Phi(\ell,\delta)\,d\ell, \tag{2.5}$$

$$P_0(L,\delta) = \delta^D(L), \tag{2.6}$$

$$P_1(L,\delta) = \Phi(L,\delta)\Big/\int \Phi(\ell,\delta)\,d\ell, \tag{2.7}$$

$$P_k(L,\delta) = \int P_1(L',\delta)\,P_{k-1}(L-L',\delta)\,dL'. \tag{2.8}$$

The effect of instrumental noise can be easily included by convolving $P(L,\delta)$ with the noise pdf. In this work, we only consider Gaussian noise with a constant rms σ_L that does not depend on the intrinsic luminosity, so the noisy $P(L,\delta,\sigma_L)$ is given by[4],

$$\begin{aligned} P(L,\delta,\sigma_L) &= P(L,\delta) * G(\sigma_L) \\ &\equiv \int dL'\,P(L',\delta)\,\frac{1}{\sqrt{2\pi}\,\sigma_L}\,e^{-(L-L')^2/2\sigma_L^2}. \end{aligned} \tag{2.9}$$

Throughout this paper we consider multiple values of $N \equiv N(\delta=0)$, the mean number of sources per voxel, given in Eq. 2.5. We note that variations in N can be interpreted in two useful ways. First, a change in N can represent a change in the number of objects for a fixed voxel size, i.e., a change in the amplitude of the luminosity function $\Phi(\ell)$ describing the source population. Alternatively, it is often instructive to consider a change in N as a change in the voxel volume, $V_{\rm vox}$, for a fixed physical source population. This allows us to study information content vs. voxel size. In the latter case, the noise per voxel, σ_L, may of course also vary as voxel size or N is varied.

---

[3] Throughout this paper, L refers to the total luminosity in a voxel, and ℓ denotes the luminosity of a single source.

[4] To simplify the notation, we will drop the σ_L notation in $P(L,\delta,\sigma_L)$ in the following chapter unless it is helpful to clarify in certain situations.

Fisher Information

Assuming that voxels are independent tracers of the large-scale density field δ, the likelihood of the whole measurement is the product of the likelihood over all voxels i, $P(\hat{L}_i, \delta)$, (Eq. 2.4),

$$\mathcal{L}(\{\hat{L}_i\}; \delta) = \prod_i P(\hat{L}_i, \delta). \tag{2.10}$$

The full Fisher information content on δ of this whole measurement is defined as [77]

$$F^{\mathrm{full}}_{\delta\delta} = -\langle \partial^2_\delta \ln \mathcal{L}(\{\hat{L}_i\}; \delta)^2 \rangle = \langle (\partial_\delta \ln \mathcal{L}(\{\hat{L}_i\}; \delta))^2 \rangle, \tag{2.11}$$

where $\langle f \rangle = \int dL\, P(L, \delta)\, f(L)$ is the expectation value of function f. The Cramér-Rao inequality states that $\sigma^2_\delta \geq 1/F^{\mathrm{full}}_{\delta\delta}$, thus placing a lower bound on the variance of parameter δ that one can attain with the data [78]. Using Eq. 2.10, we get

$$F^{\mathrm{full}}_{\delta\delta} = \langle (\partial_\delta \ln[\prod_i P(\hat{L}_i, \delta)])^2 \rangle = \sum_i \langle (\partial_\delta \ln P(\hat{L}_i, \delta))^2 \rangle, \tag{2.12}$$

and thus $F_{\delta\delta} \equiv \langle (\partial_\delta \ln P(\hat{L}_i, \delta))^2 \rangle$ is the *total Fisher information content per voxel*. Below we will quantify the Fisher information of this per-voxel basis.

In the context of this work, the parameter δ is estimated from the mean value of observable map $\hat{O}$ over a large amount of voxel data. In this case, the Fisher information per voxel for this observable is [11]

$$F^{O}_{\delta\delta} = \frac{\left(\partial_\delta \left\langle \hat{O} \right\rangle\right)^2}{\left\langle \hat{O}^2 \right\rangle - \left\langle \hat{O} \right\rangle^2} = \frac{\left(\partial_\delta \left\langle \hat{O} \right\rangle\right)^2}{\sigma^2(\hat{O})}, \tag{2.13}$$

where the denominator $\sigma^2(\hat{O})$ is the variance in map $\hat{O}$ per voxel and $\langle \cdot \rangle$ is the expectation value defined above. The condition $F^{O}_{\delta\delta} \leq F_{\delta\delta}$ holds, as the Fisher information extracted with any given observable cannot exceed the total Fisher information content. The lower bound constraint on estimating δ from the observable is $\sigma^2_\delta \geq 1/F^{O}_{\delta\delta}$; the equals sign occurs if the error on O is Gaussian.[5]

Observing LSSs with an Observable

To quantify how well an observable measures LSSs, we consider a two-point statistic, the power spectrum of observable map $\hat{O}$. Since we only consider the power spectrum on large scales, this is equivalent to smoothing fluctuation on the large

---

[5] Note that $F^{O}_{\delta\delta}$ is unchanged under rescaling of $O(L)$, i.e. for any arbitrary constant (A, C), $O(L)$ and $AO(L) + C$, are equivalent in this context. All the plots of $O(L)$ shown in the following sections are rescaled arbitrarily for better presentation.

scale of interest, or the map of $\langle\hat{O}\rangle$, where $\langle\cdot\rangle$ is the average over many voxels living in the same large-scale δ value. Since on large scales $\delta \ll 1$, we can linearize $\langle\hat{O}\rangle$ in δ, and get

$$\begin{aligned}\langle\hat{O}\rangle(\delta,\mathbf{x}) &= \langle\hat{O}\rangle(\delta) + \Delta\hat{O}(\delta,\mathbf{x}) \\ &= \langle\hat{O}\rangle(\delta=0) + \delta\,\partial_\delta\langle\hat{O}\rangle + \Delta\hat{O}(\delta,\mathbf{x}). \qquad (2.14)\end{aligned}$$

Here $\mathbf{x}$ refers to the position of the large patch of volume over which the average $\langle\cdot\rangle$ is taken. In this second line, the first term is the fiducial value of $\hat{O}$, which is a constant across the whole observing volume, so it only contributes to the $k = 0$ mode. The second term linearly traces the large-scale overdensity field δ, so it encodes the cosmological clustering information. The last term accounts for the fluctuations due to the Poisson and instrument noise, has no spatial correlation, and thus contributes to the shot noise in the power spectrum. Therefore, the power spectrum consists of the cosmological clustering and shot-noise terms:

$$P_O(k) = \left(\partial_\delta\langle\hat{O}\rangle\right)^2 P(k) + P_{O,\mathrm{shot}}, \qquad (2.15)$$

where $P(k)$ is the underlying matter power spectrum and

$$P_{O,\mathrm{shot}} = V_{\mathrm{vox}}\,\sigma^2(\hat{O}) \qquad (2.16)$$

is the shot noise, where $\sigma^2(\hat{O})$ is the variance on $\hat{O}$ due to the Poisson and instrument noise. The ratio of the cosmic signal and stochastic noise contributions to the power spectrum can be expressed in terms of the Fisher information $F^O_{\delta\delta}$,

$$\frac{\left(\partial_\delta\langle\hat{O}\rangle\right)^2 P(k)}{P_{O,\mathrm{shot}}} = \frac{\left(\partial_\delta\langle\hat{O}\rangle\right)^2 P(k)}{\sigma^2(\hat{O})\,V_{\mathrm{vox}}} = \frac{F^O_{\delta\delta}}{V_{\mathrm{vox}}}\,P(k). \qquad (2.17)$$

This equation illustrates that it is sufficient to optimize the function $O(L)$, i.e. to maximize $F^O_{\delta\delta}/V_{\mathrm{vox}}$, to minimize the statistical errors in the power spectrum.

In this thesis, our goal is to maximize $F^O_{\delta\delta}$ in order to maximize the extracted information from the large-scale density field δ from a given image (voxel intensity map). This gives the maximum signal-to-noise ratio on the power spectrum of a given image by minimizing the shot noise, and one can use this derived power spectrum to extract the cosmological information. In practice, the optimal observable to constrain δ might not be the optimal choice for a given specific type of cosmological information. For example, to measure the redshift space distortion, one might prefer an observable that can pick out low-biased tracers to boost the redshift space distortion signals. This practical consideration is beyond the scope of this thesis, so we will leave it to the future works.

Optimal Observable

According to Carron & Szapudi [11], there exists an optimal observable for δ such that the equality in $F^{O}_{\delta\delta} \leq F_{\delta\delta}$ holds; this observable can extract all the information and give the minimum variance of parameter δ. The optimal observable $O^{\rm opt}(L)$ is given by the "score function" of parameter δ evaluated at its fiducial value ($\delta = 0$):

$$O^{\rm opt}(L) = \partial_\delta \ln P(L, \delta)|_{\delta=0}. \tag{2.18}$$

This is optimal because its Fisher information is equal to the total Fisher information content per voxel, $F_{\delta\delta}$,

$$F^{\rm opt}_{\delta\delta} = F_{\delta\delta} = \partial_\delta \langle \hat{O}^{\rm opt} \rangle = \langle (\hat{O}^{\rm opt})^2 \rangle. \tag{2.19}$$

See Appendix .1 for the proof. We further define the cumulative optimal Fisher information:

$$F^{\rm opt}_{\delta\delta}(L) = \int_{-\infty}^{L'} dL\, P(L')\, (O^{\rm opt})^2(L'). \tag{2.20}$$

The limit of $L' \to \infty$ gives the optimal Fisher information $F^{\rm opt}_{\delta\delta}$. The gradient of $F^{\rm opt}_{\delta\delta}(L)$ is the amount of information gained from each L scale.

In this work, we are purely concerned with quantifying the (formal) information content. In order to demonstrate the essence of the formalism in the simple and clear context, we will assume some fixed source luminosity function and its response to density field δ, as well as the instrument noise, and quantify the information content under the particular scenario. Therefore, we do not take into account the uncertainties in the modeling of the luminosity function and the relation of the galaxy emission and the underlying density field.

2.3 Toy Model

We first start with a toy model to illustrate the concepts introduced above. In this toy model, we assume all the targeted sources have the same luminosity ℓ, and the luminosity function linearly traces the density field:

$$\Phi(\ell', \delta) = (1 + b\,\delta)\,\frac{N}{V_{\rm vox}}\,\delta^D(\ell' - \ell), \tag{2.21}$$

where δ^D is the Dirac delta function, $N = N(\delta = 0)$ is the mean number of sources per voxels, and b is the bias of the source. Here we set $\ell = 1$ for convenience.

We further consider a Gaussian noise in the measurement with rms σ_L, and thus the voxel luminosity pdf reads

$$P(L, \delta, \sigma_L) = \sum_{k=0}^{\infty} \frac{e^{-N(\delta)}\, N^k(\delta)}{k!}\, G(L, k\ell, \sigma_L), \tag{2.22}$$

where $N(\delta) = (1 + b\,\delta)\, N$ is the expectation value of the number of sources for a voxel residing in density field δ, and

$$G(x, \bar{x}, \sigma) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(x-\bar{x})^2/2\sigma^2} \tag{2.23}$$

is the Gaussian function of x with rms σ centered at $\bar{x}$.

The GD observable, described by $O^{\rm GD}(L)$, is a natural choice if $N \ll 1$, so that if a detection is made it is likely coming from a single source, and if $\sigma_L \ll \ell$, so that false detections are unlikely. In this limit, the signal is

$$\partial_\delta \langle \hat{O}^{\rm GD} \rangle = b\, N, \tag{2.24}$$

the (Poisson) variance in $\hat{O}^{\rm GD}$ reads

$$\sigma^2(\hat{O}^{\rm GD}) = N, \tag{2.25}$$

and the Fisher information on the overdensity δ is

$$F^{\rm GD}_{\delta\delta} = b^2 N. \tag{2.26}$$

This is the information on δ that one obtains from a direct measurement of the number of sources in each voxel, which is only limited by the Poisson noise owing to the finite number of sources, and is thus the maximum attainable information content for a given value of N and b. The limit $F = b^2 N$ is referred to as the “Poisson limit” hereafter. For this reason, below we will compare the ratio, $F/(b^2 N)$, of the Fisher information obtained in a given scenario, F, to the maximum Fisher information $F = b^2 N$.

For IM, the signal is

$$\partial_\delta \langle \hat{O}^{\rm IM} \rangle = \partial_\delta \langle \hat{L} \rangle = \partial_\delta L = b\, N\, \ell, \tag{2.27}$$

with variance

$$\sigma^2(\hat{O}^{\rm IM}) = \sigma^2(\hat{L}) = \sigma^2_{\rm SN} + \sigma^2_L, \tag{2.28}$$

where $\sigma_{\mathrm{SN}}^2 \equiv N\,\ell^2$ is the shot noise due to the finite number of sources contributing to the intensity signal. This gives the Fisher information,

$$F_{\delta\delta}^{\mathrm{IM}} = \frac{b^2\,N^2\,\ell^2}{N\,\ell^2 + \sigma_L^2} = b^2 N\,\frac{\sigma_{\mathrm{SN}}^2}{\sigma_{\mathrm{SN}}^2 + \sigma_L^2}. \tag{2.29}$$

In the limit where the noise in the intensity is dominated by the Poisson noise, $\sigma_L \ll \sigma_{\mathrm{SN}}$, this gives the optimal result, $F = b^2 N$ (Poisson limit). However, in general, the Fisher information may be suppressed by the instrument noise. If we model variations in voxel volume by changing N, Eq. 2.29 shows that the performance of IM as quantified by $F/(b^2 N)$ is independent of voxel size as long as either (1) we are in the Poisson-noise-dominated regime $\sigma_L \ll \sigma_{\mathrm{SN}}$ or (2) the instrument noise scales with voxel size as $\sigma_L^2 \propto N \propto V_{\mathrm{vox}}$. The noise scaling in case (2) is what one would expect if the instrument noise is photon noise dominated.

Below we discuss the optimal observable $O^{\mathrm{opt}}(L)$, and compare its Fisher information with $O^{\mathrm{GD}}(L)$ and $O^{\mathrm{IM}}(L)$ in three different regimes: $N \ll 1$, $N \sim 1$, and $N \gg 1$.

$N \ll 1$

In the $N \ll 1$ limit, the voxel luminosity probability distribution can be simplified by Taylor-expanding Eq. 2.22 and keeping terms only up to first order in $N(\delta)$:

$$P(L,\delta) \simeq (1 - N(\delta))\,G(L,0,\sigma_L) + N(\delta)\,G(L,\ell,\sigma_L). \tag{2.30}$$

The optimal observable can then be calculated from Eq. 2.18,

$$O^{\mathrm{opt}}(L) \simeq \frac{b\,N(G(L,\ell,\sigma_L) - G(L,0,\sigma_L))}{(1-N)\,G(L,0,\sigma_L) + N\,G(L,\ell,\sigma_L)}. \tag{2.31}$$

In Fig. 2.1, the top panels show $P(L)$ and $O^{\mathrm{opt}}(L)$, for $N = 0.01$ ($\sigma_{\mathrm{SN}} = 0.1$) with various instrument noise σ_L levels[6] , and the bottom panels show the Fisher information (See Eq. 2.13) of the optimal observable (cumulated Fisher information; see Eq. 2.20), the IM observable, and the GD observable for a range of threshold L_{th}.

---

[6] The true optimal observable of this case is indeed a stair-like function like the one shown in Fig. 2.2, rather than a single step we get from approximation with only $k = 0, 1$ terms. However, this approximation gives almost the same Fisher information as the optimal observable derived from including more k terms. This is due to the fact that the probability of higher k terms is too small to have a significant contribution to Fisher information. Therefore, for the purpose of demonstrating the idea, we ignore the higher-order terms for the optimal observable.

Considering first the low-noise regime, $\sigma_L \ll \ell$ (left panels), we find as expected that thresholded GD is optimal. This is clearly seen from the fact that the optimal observable $O^{\rm opt}(L)$ (red curve) is close to a step function. In addition, the Fisher information of $O^{\rm GD}(L)$ as a function of $L_{\rm th}$ attains approximately the same total information as the optimal observable, for a wide range of values of $L_{\rm th}$. Any threshold from a few times σ_L to ℓ minus a few times σ_L perfectly "counts" sources. As a result, the information content is optimal, in the sense that $F/(b^2N) = 1$.

In the very low noise regime, $\sigma_L \ll \sigma_{\rm SN}$ (where $\sigma_{\rm SN}$ is the Poisson noise in luminosity L), IM is *also* optimal, as can be seen by the horizontal blue line in the bottom panel. This is because in the $N \ll 1$ and low-noise ($\sigma_L^2 \ll N\ell^2$) limit, most voxels have either $L \approx 0$ or $L \approx \ell$, as shown by the $P(L)$ function, and thus the information content must be concentrated at these two L scales as well. As long as an observable is able to discriminate these two classes of voxels, i.e. having distinct values at $L = 0$ and $L = \ell$, it is able to capture the signals (quantified by $\partial_\delta \langle \hat{O} \rangle$) in the map, regardless of the O(L) function values at other L values, as almost no voxel falls in this regime. However, in the intermediate regime ($\sigma_L = 0.2$ case), $\sigma_{\rm SN} < \sigma_L \ll \ell$, IM suffers from instrument noise suppression (see Equation 2.29), while source detection is still optimal.

Moving on from the low-noise regime toward cases where $\sigma_L \ll \ell$ no longer holds ($\sigma_L = 1, 3$), the Gaussian noise profiles of the $P(L)$ function centered at 0 and ℓ start to overlap, so a GD threshold function is no longer optimal, as it cannot effectively count the sources. Indeed, the optimal observable $O^{\rm opt}(L)$ is now a more gradually increasing function of L. As for the Fisher information, we can see from Fig. 2.1 that even for the optimal choice of $L_{\rm th}$, the information contained in the GD observable is lower than the information in the optimal observable. At the same time, the IM information content becomes larger relative to the optimal information content. In the largest noise regime ($\sigma_L = 3$), IM is very close to optimal.

We note, however, that as the noise increases, the *absolute* information content strongly decreases, i.e., $F/(b^2N) \ll 1$. This is of course to be expected: instrument noise makes it difficult to measure cosmological signals.

$N \sim 1$

Next, we consider the $N \sim 1$ regime. In this scenario, the $k \geq 2$ terms in Eq. 2.22 must be taken into account. We take $N = 1$ in this example and consider different σ_L values as before. The results are shown in Fig. 2.2. The $P(L)$ function is the linear

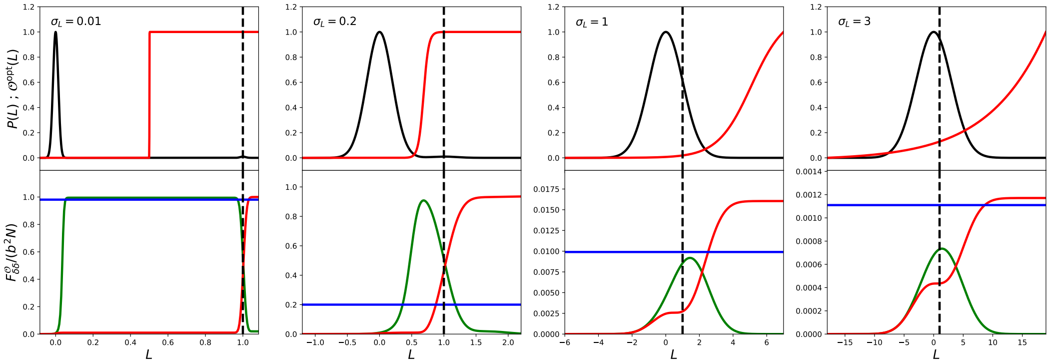

Figure 2.1: **Top:** $P(L)$ (black) and $O^{\rm opt}$(L) (red) of the toy model with a single type of source with luminosity $\ell = 1$ and mean number of sources per voxel $N = 0.01$, for different Gaussian noise σ_L. **Bottom:** $F^O_{\delta\delta}$ of IM observable (linear function; blue), GD observables (step function) as a function of step L (green), and the cumulative optimal Fisher information (red). The black dashed lines mark $L = \ell$ for reference.

combination of the Gaussian profile with variance σ_L^2 centered at $L = 0,\ \ell,\ 2\ell, ...,$ with their amplitude following a Poisson distribution. We can see that the optimal observable is a stair-like function, which gradually smoothed out with increasing noise.

The linear observable is better than the step function in all cases in terms of their Fisher information. The reason is the same as in the $N \ll 1$ situation: in the low-noise regime, where most voxel luminosity L has values around $L = 0,\ \ell,\ 2\ell, ...,$ the only observable value that matters is where L is near these values. The linear observable gives exactly the same value at these points as the optimal one. On the other hand, the step function is not a good observable in this case. The step function gives the same weights for all the voxels above the step, so it ignores the fact that higher-luminosity voxels likely have more sources and are more likely to reside in high-δ regions. Note that this is not an issue for the $N \ll 1$ case, as there are very few voxels containing multiple sources; the total information content in these voxels is also negligible. Whereas here we have $N \sim 1$, the multiple-source voxels contribute to a significant portion of the total information content, and a proper weighting for them in the observable is essential for capturing the information from the map.

In the high instrument noise regime, the linear observable is also superior to the step function, which follows the same argument as in the $N \ll 1$ case.

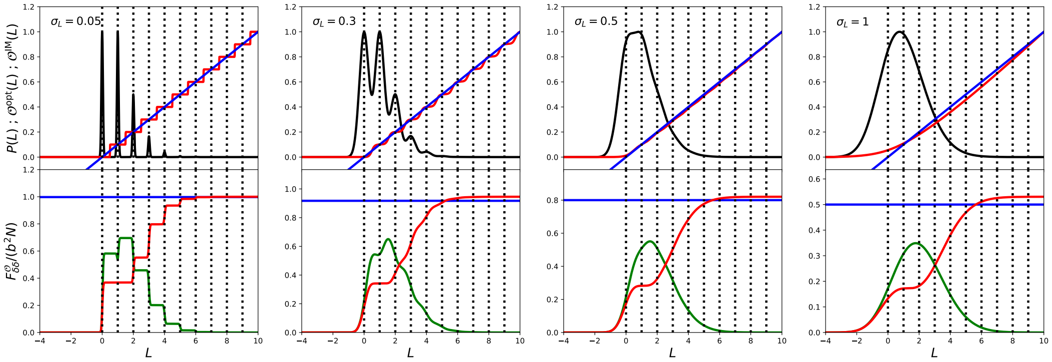

Figure 2.2: **Top:** $P(L)$ (black) and $O^{\rm opt}$(L) (red) of toy model with a single type of source with luminosity $\ell = 1$ and mean number of sources per voxel $N = 1$, for different Gaussian noise σ_L. **Bottom:** $F^{O}_{\delta\delta}$ of IM observable (linear function; blue), GD observables (step function) as a function of threshold L (green), and the cumulative optimal Fisher information (red). The black dotted lines mark the integer of ℓ, the possible intrinsic voxel luminosity.

$N \gg 1$

In the $N \gg 1$ limit, the Poisson function converges to a Gaussian,

$$\frac{e^{-N(\delta)}\,N^k(\delta)}{k!} \simeq \frac{1}{\sqrt{2\pi\,N(\delta)}}\,e^{-\frac{(N(\delta)-k)^2}{2\,t\,N(\delta)}} \tag{2.32}$$

and the summation over k in the $P(D)$ formalism can be approximated by an integral, so Eq. 2.22 becomes the convolution of two Gaussian functions, which gives another Gaussian,

$$\begin{aligned} P(L,\delta) &= \sum_{k=0}^{\infty} \frac{e^{-N(\delta)}\,N^k(\delta)}{k!}\,G(L,k\ell,\sigma_L) \\ &\simeq \int_0^{\infty} dk \left[\frac{1}{\sqrt{2\pi\,N(\delta)}}e^{-\frac{(k-N(\delta))^2}{2\,N(\delta)}}\right]\left[\frac{1}{\sqrt{2\pi}\,\sigma_L}e^{-\frac{(L-k\ell)^2}{2\,\sigma_L^2}}\right] \\ &= \frac{1}{\sqrt{2\pi}\,\bar{\sigma}}e^{-\frac{L'^2}{2\,\bar{\sigma}^2}}, \end{aligned} \tag{2.33}$$

where $L'(\delta) \equiv L - N(\delta)\,\ell$, and $\bar{\sigma}^2(\delta) \equiv \sigma_L^2 + N(\delta)\,\ell^2$. Note that $\bar{\sigma}^2$ is the total variance from both instrument noise and Poisson noise. In the absence of instrument noise, we still have a nonzero voxel pdf P(L) owing to the Poisson variance of the sources themselves. We then derive the optimal observable from Eq. 2.18, with some rescaling to get rid of all irrelevant constants,[7]

$$O^{\rm opt}(L) = L' + \frac{\ell}{2\,\bar{\sigma}^2}\,L'^2. \tag{2.34}$$

---

[7] $L' \equiv L'(\delta = 0) = L - N\,\ell$; $\bar{\sigma}^2 \equiv \bar{\sigma}^2(\delta = 0) = \sigma_L^2 + N\,\ell^2$.

Hence, the optimal observable is a linear combination of a linear and a quadratic term, and the contribution from the latter gets smaller as the noise increases.

The top row of Fig. 2.3 shows the $P(L)$ and $O^{\mathrm{opt}}(L)$ for different σ_L levels, while fixing $N = 100$. We can see that as σ_L increases, the $P(L)$ profile is broadened, and $O^{\mathrm{opt}}(L)$ becomes closer to the linear function. The bottom row shows the Fisher information for the different observables. In all cases the step function is not the preferable observable. The linear function performs as well as the optimal observable, even in the $\sigma_L = 0$ limit, where the optimal observable deviates from the linear function significantly. This is because the quadratic term in the optimal observable has negligible contribution to the optimal Fisher information (see Appendix. .2 for explanation).

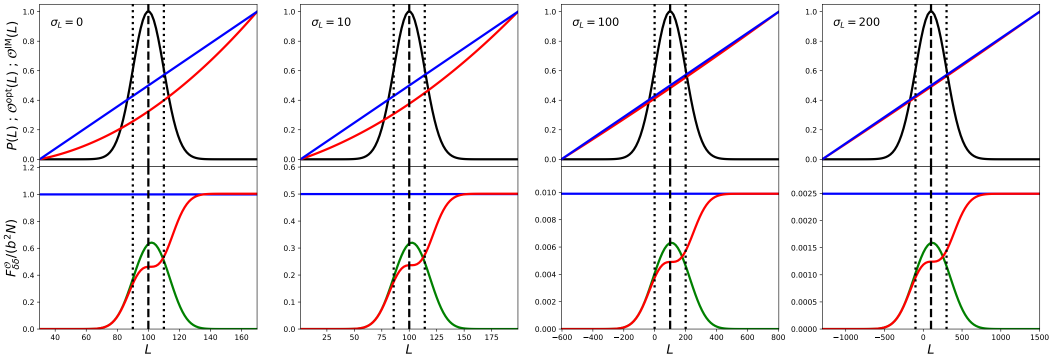

Figure 2.3: **Top:** $P(L)$ (black) and O^{opt}(L) (red) of the toy model with a single type of source with luminosity $\ell = 1$ and mean number of sources per voxel $N = 100$, for different Gaussian noise σ_L. **Bottom:** $F^{O}_{\delta\delta}$ of IM observable (linear function; blue), GD observables (step function) as a function of step L (green), and the cumulative optimal Fisher information (red). The black dashed and dotted lines mark the mean and $\pm\bar{\sigma}$ of the $P(L)$ profile, i.e. $N\ell$ and $N\ell \pm \bar{\sigma}$, respectively.

Toy Model Summary

In conclusion, for our toy model with a luminosity function describing sources with a single luminosity ℓ, we find the following limiting behaviors:

- For a low number of sources per voxel, $N \ll 1$, *and* low noise compared to the source luminosity, $\sigma_L \ll \ell$, it is optimal to detect individual sources by applying the threshold observable $O^{\mathrm{GD}}(L)$. In this scenario, the voxels below the detection threshold contain only noise and make up the majority of voxels. The GD observable assigns them zero weight, and therefore they do not contribute to the noise in the map. On the other hand, voxels with luminosity above the threshold all contain a (single) source (as the probability of a noise fluctuation exceeding the threshold is infinitesimally small in the limit $\sigma_L \ll \ell$). This leads to a measurement of the source number density only limited by the shot noise owing to the finite number of sources N.

- In the same low-N but high-noise regime where $\sigma_L > \ell$, the signal from sources cannot be unambiguously distinguished from noise fluctuations, so that the GD approach is suboptimal and instead the IM observable is close to optimal. The measurement is limited by instrument noise (as opposed to by shot noise owing to the finite number of sources), so that our ability to constrain δ (as quantified by the Fisher information) is unsurprisingly much weaker than the one in the $\sigma_L \ll \ell$ regime.
- In the opposite regime of a large number of sources per voxel, $N \gg 1$, we find that IM is (nearly) optimal independently of the instrument noise.

The above results are intuitive and serve as useful benchmarks to refer to in the following sections. Intermediate cases can be understood as interpolations between the above limiting scenarios.

2.4 Schechter Luminosity Function Model

For a more realistic description, we consider that the galaxy populations follow a Schechter luminosity functional form: $\Phi(\ell) = \phi_* \, (\ell/\ell_*)^\alpha \, e^{-\ell/\ell_*}$[67][8]. To simplify the notation, below all the ℓ represent ℓ/ℓ_*; in other words, we use ℓ_* as the unit for luminosity. This can be easily scaled to any desired unit in real experiments.

One requirement for applying the $P(D)$ formalism is to have a finite N, the mean number of sources per voxel. To ensure that the integration in Eq. 2.5 converges, we use a modified Schechter function introduced by Breysse et al. [5]:

$$\Phi(\ell) = \phi_* \, \ell^\alpha \, e^{-\ell} \, e^{-\ell_{\min}/\ell}. \tag{2.35}$$

We assume that the luminosity function linearly traces the density field,

$$\Phi(\ell, \delta) = (1 + b\,\delta)\,\Phi(\ell). \tag{2.36}$$

---

[8] To simplify the notations, $\Phi(\ell)$ refers to $\Phi(\ell, \delta = 0)$, the average luminosity function across the universe.

The optimal observable, $P(L)$, and $F^O_{\delta\delta}$ can be derived from equations in Sec. 2.2. Note that Eq. 2.36 assumes a luminosity-independent clustering bias. In a more realistic description, we would describe the response to the underlying *matter* overdensity δ in terms of a luminosity-dependent bias $b(\ell)$. This is a straightforward modification to our formalism, but for simplicity we will not pursue it here.

Applying the low-ℓ suppression for $\ell \lesssim \ell_{\text{min}}$ has a physical motivation: galaxies cannot be infinitely faint. The value of ℓ_{min} is not easily constrained observationally; however, it is not an issue for our calculation. In Appendix .4, we show that the choice of ℓ_{min} does not affect our results as long as ℓ_{min} is much smaller than σ_L, the instrumental noise in the observation. In this work, we adopt the fiducial $\ell_{\text{min}} = 10^{-3}$.

The faint-end slope α usually has the value $-2 < \alpha < -1$ from observations. We take $\alpha = -1.5$ as our fiducial value in this work, and we discuss the effects of choosing different α values in Appendix .5.

Quantifying the Confusion

Fig. 2.4 shows the normal Schechter function (without ℓ_{min} cutoff) with fiducial α. We also plot the first three moments of the Schechter function that give the quantity of particular interest:

$$N = V_{\text{vox}} \int d\ell \, \Phi(\ell); \tag{2.37}$$

$$\langle \hat{L} \rangle = V_{\text{vox}} \int d\ell \, \Phi(\ell) \, \ell; \tag{2.38}$$

$$\sigma^2_{\text{SN}} = V_{\text{vox}} \int d\ell \, \Phi(\ell) \, \ell^2. \tag{2.39}$$

As shown in the plot, the total number of sources N diverges as we take ℓ_{min} to zero, corresponding to an infinite number of (mostly faint) sources per voxel in the absence of a cutoff. As a result, the value of N in the modified Schechter function depends on the choice of ℓ_{min}, while for $\langle \hat{L} \rangle$ and σ^2_{SN}, the integration is converged at the faint end, so its value is not susceptible to the artificial ℓ_{min} cutoff (these convergence properties are true for all $-2 < \alpha < -1$).

For the above reasons, N is not a well-defined quantity in the Schechter function case and is ill-suited to quantify the level of confusion as used in the toy model. We therefore introduce an effective number of sources per voxel, N_{eff}, defined with the cutoff-independent quantities $\langle \hat{L} \rangle$ and σ^2_{SN}.

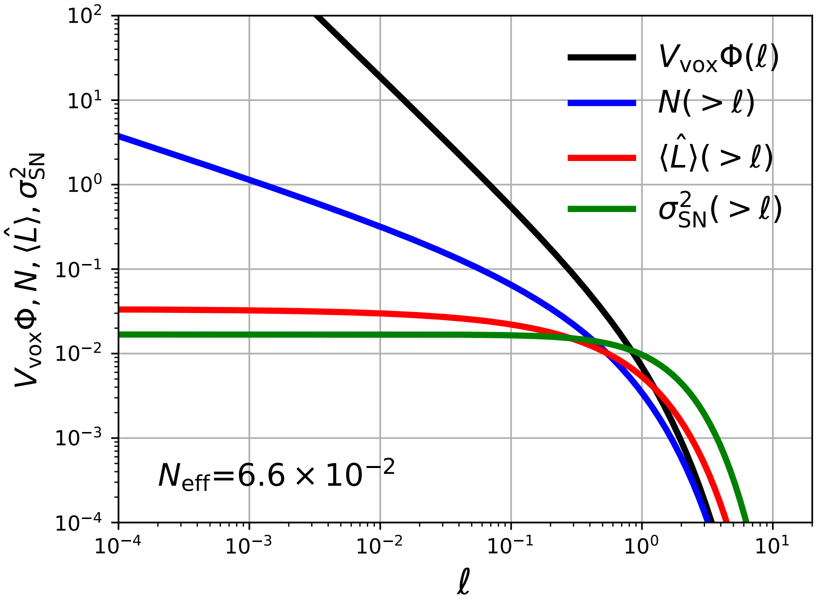

Figure 2.4: Normal Schechter luminosity function (without $\ell_{\min}$ cutoff) using fiducial $\alpha = -1.5$ (black), and its cumulative N (blue), $\langle\hat{L}\rangle$ (red), and σ^2_{SN} (green).

N_{eff}

The IM signal in the Schechter model is given by

$$\partial_\delta \langle \hat{O}^{\mathrm{IM}} \rangle = \partial_\delta \langle \hat{L} \rangle = b\, V_{\mathrm{vox}} \int d\ell\, \Phi(\ell)\, \ell, \tag{2.40}$$

with variance

$$\sigma^2(\hat{O}^{\mathrm{IM}}) = \sigma^2(\hat{L}) = \sigma^2_{\mathrm{SN}} + \sigma^2_L. \tag{2.41}$$

The Fisher information is therefore

$$F^{\mathrm{IM}}_{\delta\delta} = \frac{b^2 \left(V_{\mathrm{vox}} \int d\ell\, \Phi(\ell)\, \ell\right)^2}{V_{\mathrm{vox}} \int d\ell\, \Phi(\ell)\, \ell^2 + \sigma^2_L} = \frac{b^2 \langle \hat{L} \rangle^2}{\sigma^2_{\mathrm{SN}} + \sigma^2_L}. \tag{2.42}$$

We now define the effective number of sources per voxel as the IM Fisher information in the Poisson-limited case, $\sigma_L \ll \sigma_{\mathrm{SN}}$,

$$N_{\mathrm{eff}} \equiv \frac{\left(V_{\mathrm{vox}} \int d\ell\, \Phi(\ell)\, \ell\right)^2}{V_{\mathrm{vox}} \int d\ell\, \Phi(\ell)\, \ell^2} = \frac{\langle \hat{L} \rangle^2}{\sigma^2_{\mathrm{SN}}}. \tag{2.43}$$

This can be interpreted as the reciprocal of the effective shot noise in the IM regime, which is an analogy to the $1/N$ shot noise in GD.

The total Fisher information from IM (Eq. (2.42)) can be rewritten as

$$F_{\delta\delta}^{\mathrm{IM}} = b^2 N_{\mathrm{eff}} \frac{\sigma_{\mathrm{SN}}^2}{\sigma_{\mathrm{SN}}^2 + \sigma_L^2}. \tag{2.44}$$

The effective number of sources per voxel thus tells us how well the IM observable can possibly perform given a source population, while the performance is weakened when $\sigma_L \gtrsim \sigma_{\mathrm{SN}}$. As is the case for the toy model, the IM performance is independent of V_{vox} if the instrument noise scales like $\sigma_L^2 \propto V_{\mathrm{vox}}$ or if the instrument noise is negligible, $\sigma_L \ll \sigma_{\mathrm{SN}}$.

L_{SN}

Aside from N_{eff}, we further introduce the luminosity scale where the voxels are highly susceptible to shot noise, L_{SN}, to be another quantity related to confusion.

We first define the cumulative intensity shot noise,

$$\sigma_{\mathrm{SN}}^2(\ell) \equiv V_{\mathrm{vox}} \int_0^{\ell} d\ell' \, \Phi(\ell') \, \ell'^2. \tag{2.45}$$

This includes the shot-noise variance from all the sources fainter than ℓ. A useful quantity is then the "crossover luminosity," L_{SN}, where the intensity shot noise equals the source luminosity, $\sigma_{\mathrm{SN}}(L_{\mathrm{SN}}) = L_{\mathrm{SN}}$.

When $\ell < L_{\mathrm{SN}}$, $\sigma_{\mathrm{SN}}(\ell) > \ell$, which means that the confusion noise from the fainter source is comparable to ℓ; when $\ell > L_{\mathrm{SN}}$, $\sigma_{\mathrm{SN}}(\ell) < \ell$, which means that the confusion noise from faint sources becomes negligible. Fig. 2.5 shows the $\sigma_{\mathrm{SN}}(\ell)$ with four different source densities and their L_{SN} marked by the dotted vertical lines.

Relation between N_{eff} and L_{SN}

The modified Schechter luminosity function we adopted in this work is composed of a power law with slope α and exponential cutoffs at both low- and high-ℓ ends, which guarantee convergence of integration for all moments. Of particular interest are the first three moments that give N (zeroth), $\langle \hat{L} \rangle$ (first), σ_{SN}^2 (second) respectively.

If the luminosity function is only a power law (i.e. $\Phi \propto \ell^{\alpha}$) with $-2 < \alpha < -1$, the zeroth moment converges at the high-ℓ end and diverges at the low$-\ell$ end, while the convergence of higher moments is reversed. Applying the exponential cutoff suppresses contribution from scales beyond the cutoff scale, and thus the integration

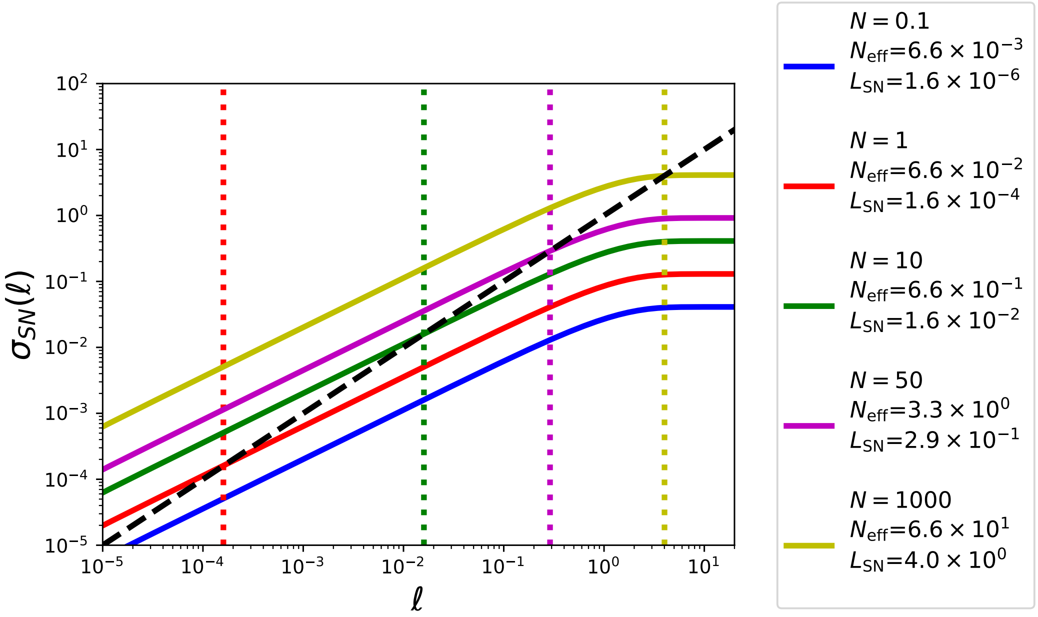

Figure 2.5: $\sigma_{\rm SN}(\ell)$ with different source densities (solid lines). The black dashed line is $\sigma_{\rm SN} = \ell$, and its intersection with $\sigma_{\rm SN}(\ell)$ is $L_{\rm SN}$.

is dominated by the sources with luminosity around the cutoff. Therefore,

$$N = V_{\rm vox} \int \Phi(\ell)\, d\ell \sim V_{\rm vox}\, \Phi(\ell_{\rm min})\, \ell_{\rm min}; \tag{2.46}$$

$$\langle \hat{L} \rangle = V_{\rm vox} \int \Phi(\ell)\, \ell\, d\ell \sim V_{\rm vox}\, \Phi(\ell_*)\, \ell_*^2; \tag{2.47}$$

$$\sigma_{\rm SN}^2 = V_{\rm vox} \int \Phi(\ell)\, \ell^2\, d\ell \sim V_{\rm vox}\, \Phi(\ell_*)\, \ell_*^3. \tag{2.48}$$

Note that the quantity $\ell\, \Phi(\ell)$ is the count per $\log \ell$, so the above approximations imply that N is dominated by sources with luminosity around $\ell_{\rm min}$, whereas $\langle \hat{L} \rangle$ and $\sigma_{\rm SN}^2$ are dominated by $\ell \sim \ell_*$ sources.

From these relations we can also derive

$$N_{\rm eff} = \frac{\langle \hat{L} \rangle^2}{\sigma_{\rm SN}^2} \sim V_{\rm vox}\, \Phi(\ell_*)\, \ell_*, \tag{2.49}$$

so $N_{\rm eff}$ is approximately the number of sources per $\log(\ell)$ at ℓ_*.

Based on the above, we can roughly infer the relation between $L_{\rm SN}$ and $N_{\rm eff}$. Since

$$\sigma_{\rm SN}^2(L_{\rm SN}) \equiv L_{\rm SN}^2 \sim V_{\rm vox}\, \Phi(L_{\rm SN})\, L_{\rm SN}^3, \tag{2.50}$$

if $L_{\rm SN} < \ell_*$, we get

$$V_{\rm vox}\,\Phi(L_{\rm SN})\,L_{\rm SN} \sim 1 > V_{\rm vox}\,\Phi(\ell_*)\,\ell_* = N_{\rm eff}. \tag{2.51}$$

On the contrary, if $L_{\rm SN} > \ell_*$, then

$$V_{\rm vox}\,\Phi(L_{\rm SN})\,L_{\rm SN} \sim 1 < V_{\rm vox}\,\Phi(\ell_*)\,\ell_* = N_{\rm eff}. \tag{2.52}$$

Hence, we conclude that

$$\begin{aligned} L_{\rm SN} < \ell_* \;\Leftrightarrow\; N_{\rm eff} < 1; \\ L_{\rm SN} > \ell_* \;\Leftrightarrow\; N_{\rm eff} > 1. \end{aligned} \tag{2.53}$$

The argument above is only an order-of-magnitude estimation. The $L_{\rm SN} - N_{\rm eff}$ relation with our fiducial Schechter parameters is shown in Fig. 2.6. The actual scales where $N_{\rm eff} = 1$ and $L_{\rm SN} = \ell_*(= 1)$ happen are off by around an order of magnitude. Later we will focus on the limiting scenarios where $L_{\rm SN} \ll \ell_*$ and $L_{\rm SN} \gg \ell_*$ respectively. In the situation where $L_{\rm SN} \sim \ell_*$ within roughly an order of magnitude, one should keep in mind the caveat that the cases of interest might be closer to either of the limiting regimes, or some intermediate situation, so the arguments for the limiting cases cannot be applied naively.

Noiseless Scenario

We first consider an idealized scenario without instrument noise σ_L. This example will allow us to derive some useful insights before we move on to the more realistic scenario including instrument noise σ_L.

The major difference between the toy model and the Schechter function case is that in the toy model with zero instrumental noise, even in the highly confused scenario ($N \gg 1$), the Fisher information of the optimal observable (and of $\mathcal{O}^{\rm IM}(L)$) still reaches the Poisson limit, since we can unambiguously count the number of sources for any given voxel luminosity L in the toy model. In the Schechter function case, on the other hand, we are not able to distinguish the exact composition of sources in the voxels, and thus the information content will be suppressed by the confusion.

Fig. 2.7 shows the $P(L)$, $\mathcal{O}^{\rm opt}(L)$, and the Fisher information relative to the total information from directly counting sources, $F/(b^2N)$, for three different N levels. Below we describe the important observations from these results.

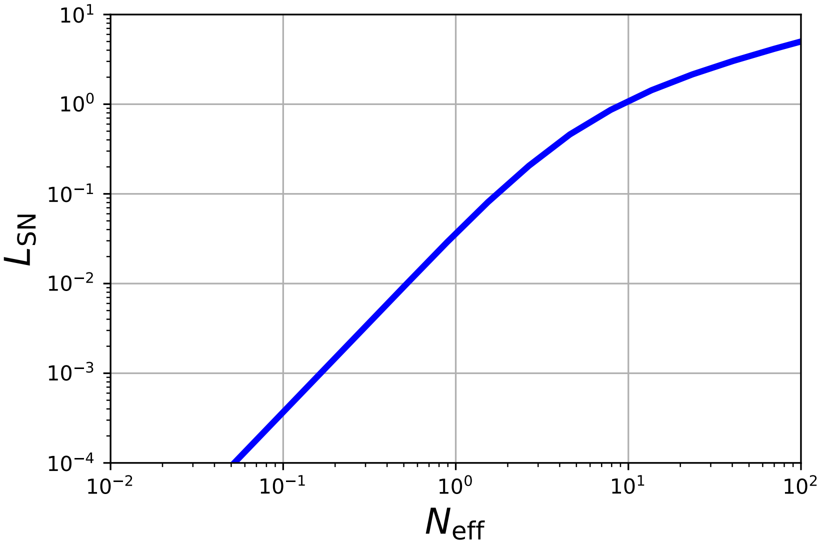

Figure 2.6: $L_{\rm SN} - N_{\rm eff}$ relation with fiducial Schechter function faint-end slope $\alpha = -1.5$. Note that the actual scales where $N_{\rm eff} = 1$ and $L_{\rm SN} = \ell_*(= 1)$ happen are off by around an order of magnitude.

- The probability distribution of the total voxel luminosity, $P(L)$, shifts to higher L as N increases.

- The optimal observable has a smoothed step-function-like shape. The transition L scale is around $L_{\rm SN}$, except for the $N = 0.1$ case, where $L_{\rm SN} \ll \ell_{\rm min}$, and the transition is strongly affected by the cutoff $\ell_{\rm min}$. The interpretation is as follows: when $L \lesssim L_{\rm SN}$, $\sigma_{\rm SN} \gtrsim L$ (and the effective number of sources below L is not small), and thus the possibility that a given L voxel is composed of multiple faint sources is non-negligible. In this regime, the optimal observable prefers giving brighter voxels more weight since they are more likely to hold more sources, and this explains the rising part of the $O^{\rm opt}$ function. On the bright end, where $L > L_{\rm SN}$, most of the voxels with these L values are dominated by the single $\ell \sim L$ source, and thus this is in the GD regime, and the optimal observable is a uniform weighting.

- The $N = 0.1$ case reaches the Poisson limit. This is because a threshold $L_{\rm th}$ below $\ell_{\rm min}$ has the property that whenever a voxel luminosity exceeds $L_{\rm th}$, that voxel is likely to contain only a single source. Thus, (only) this scenario allows us to directly count galaxies and thus to optimally trace the overdensity δ. For larger N, only sources with $\ell > L_{\rm SN} > \ell_{\rm min}$ can be "counted."

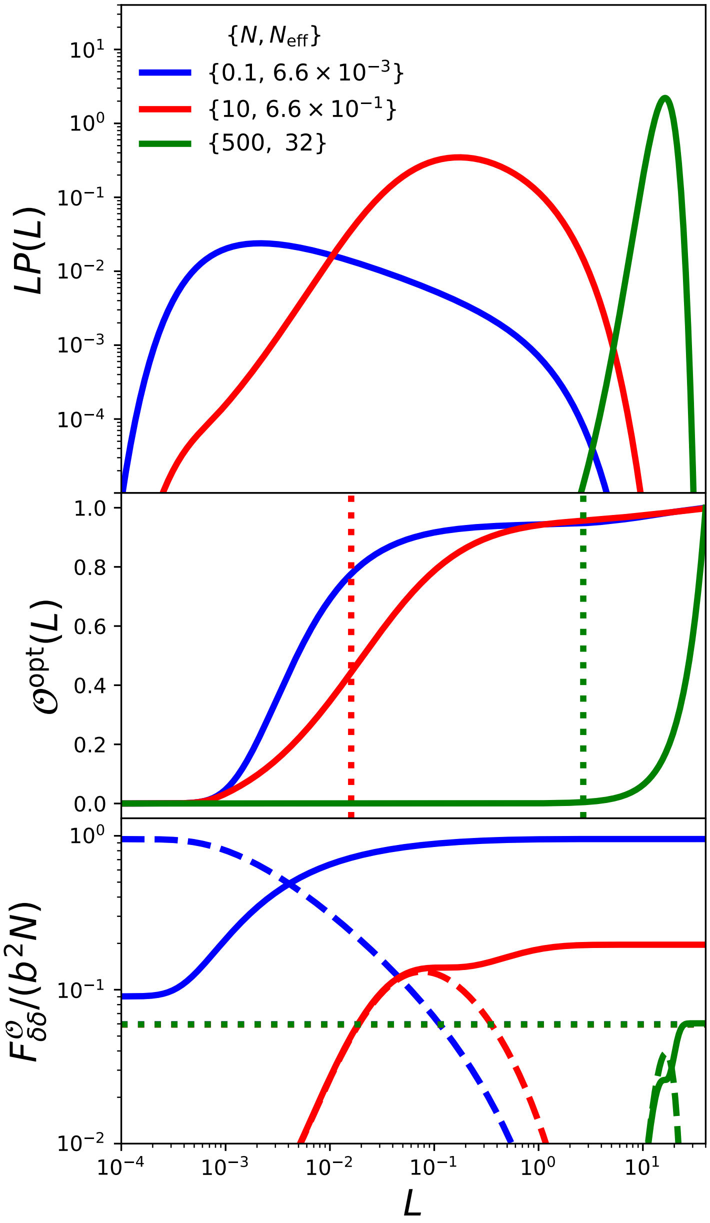

Figure 2.7: Fiducial Schechter function without instrument noise. **Top:** $LP(L)$ with different N levels. Note that the area underneath the $LP(L)$ curve gives the probability per log L. **Middle:** The optimal observables for each case. The dotted lines mark the $L_{\rm SN}$ (in $N = 0.1$ case, $L_{\rm SN} \ll 10^{-4}$, so the blue dot line is outside the x-axis range). **Bottom:** $F^{O}_{\delta\delta}$ of IM observable (dotted; note that the three dotted lines overlap), GD observable as a function of step L (dashed), and the cumulative optimal Fisher information (solid).

- In the $N = 0.1$ case, the step function with threshold $L_{\rm th} < \ell_{\rm min}$ is approximately optimal as discussed above.
- In the two larger-N scenarios, the confusion has a significant impact on fainter voxels ($L \lesssim L_{\rm SN}$) that degrades the information content, and thus the optimal Fisher information is less than the Poisson limit.
- In the two larger-N scenarios, the optimal Fisher information is built up at two stages that correspond to the IM part at $L \lesssim L_{\rm SN}$, where the observable is weighted by luminosity, and the GD part at $L \gtrsim L_{\rm SN}$, where the bright sources can be counted individually.
- In the absence of instrument noise, $F^{\rm IM}_{\delta\delta}/(b^2N)$ is independent of N (and thus the voxel size). This can be understood in the following way: the IM observable measures a luminosity-weighted "count" of the number of sources. Because of the properties of the Schechter function discussed in Sec.2.4, this weighted count is dominated by sources with luminosity near ℓ_* $(= 1)$, and the information content is given by $N_{\rm eff} \ll N$. See also Appendix. .3 for further discussion of this point.

In summary, when N is not small, confusion, in combination with a range of source luminosities, implies that we cannot reach the Poisson limit even without instrument noise. The IM observable never reaches the Poisson limit, regardless of N, while GD reaches $F/(b^2N) = 1$ only if $N \ll 1$.

General Case with Instrumental Noise σ_L

In reality, the instrumental noise σ_L has to be taken into account. Just as $L_{\rm SN}$ sets the approximate luminosity where a source rises above the confusion noise due to fainter objects, σ_L determines the luminosity where objects rise above the instrument noise. Another characteristic scale is the ℓ_* of the Schechter function, which is set to unity in this chapter as we scale luminosities in units of ℓ_*. The shape of the optimal observable and the Fisher information are determined by the relative value of these three luminosity scales $\{L_{\rm SN}, \sigma_L, \ell_*\}$. In this section, we will classify different scenarios by the relative ordering of these scales and discuss each case in detail.

We split the scenarios into two categories depending on the $L_{\rm SN}$ and ℓ_* relation. Case I is the low-confusion regime where $L_{\rm SN} < \ell_*$, corresponding to the $N_{\rm eff} < 1$, and we further discuss three subcases in this category depending on values of σ_L. Case II is the highly confused regime defined by $L_{\rm SN} > \ell_*$, corresponding to $N_{\rm eff} > 1$.

Fig. 2.8 summarizes the schematic ordering of these categories, and the shaded regions mark the optimal observing strategy for each case discussed below.

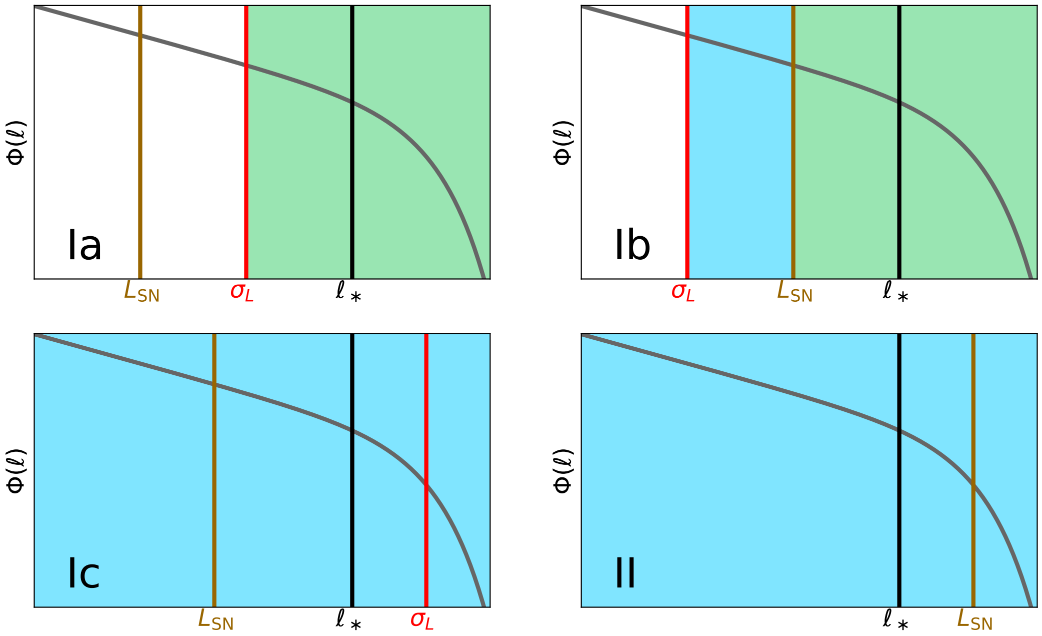

Figure 2.8: Ordering of $\{L_{\rm SN}, \sigma_L, \ell_*\}$ in each case discussed in Sec. 2.4. Case I is defined by $L_{\rm SN} < \ell_*$, corresponding to the $N_{\rm eff} < 1$ low-confusion regime, and its three subcases in this category as determined by the position of σ_L. Whereas the Case II is the highly confused regime defined by $L_{\rm SN} > \ell_*$, corresponding to $N_{\rm eff} > 1$. The blue shaded regions are where IM is the optimal strategy, and the green shaded regions mark the scales above the optimal threshold when the GD observable is the optimal strategy.

Case I: $L_{\rm SN} < \ell_*$

Here we have a relatively low number density, with $L_{\rm SN} < \ell_*$, approximately corresponding to the $N_{\rm eff} < 1$ regime. We will thus apply the $P(D)$ calculation to derive the $P(L)$ and the optimal observable.

Case Ia: $\mathbf{L_{SN} < \sigma_L < \ell_*}$ We first consider the case of intermediate instrument noise, i.e., between $L_{\rm SN}$ and ℓ_*. Fig. 2.9 shows two examples in this case with different σ_L. This is the regime where GD works well: the instrument noise is much smaller than ℓ_*, and the voxels with $L \gtrsim \sigma_L$ do not suffer from confusion noise.

Therefore, as expected, the optimal observable here is close to a step function with a transition at a few times σ_L (Fig. 2.9, two middle panels). The optimal step function has a threshold at $\sim 3\sigma_L$ (dashed vertical lines in the two middle panels), and this optimal step function observable indeed captures nearly the optimal information, as shown in the right panel of Fig. 2.9. This indicates that GD using a threshold at a few σ is the optimal strategy.

We also note from the solid curves in the right panel of Figure 2.9 that the information content is dominated by voxels with total luminosity within an order of magnitude of the optimal threshold value at $\sim 3\sigma_L$.

The total optimal Fisher information $F_{\delta\delta}^{\rm opt}/b^2$ in this case should be of the order $N(\ell > \sigma_L)$, the number of sources per voxel above σ_L, since we can count sources brighter than the noise level without confusion. This is consistent with the results in the right panel of Fig. 2.9, though $F_{\delta\delta}^{\rm opt}/b^2$ is slightly lower than $N(\ell > \sigma_L)$ owing to instrumental noise σ_L.

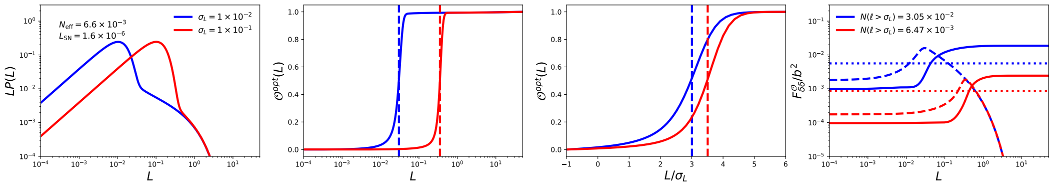

Figure 2.9: Two examples of case Ia. **Left:** $P(L)$ after convolving with $\sigma_L = 0.01$ (blue) and 0.1 (red). **Middle left:** optimal observables (solid lines). The dashed lines are the optimal threshold for the step function observable, i.e. the peak of the dashed curve in the right panel. **Middle right:** same as the middle left panel, but with L/σ_L on the x-axis on a linear scale. **Right**: the integrated Fisher information for the optimal observable (solid), Fisher information of the step function observable as a function of step position (dashed), and the Fisher information of the linear observable (dotted).

Case Ib: $\sigma_{\rm L} < {\rm L}_{\rm SN} < \ell_*$ We now consider the low-noise regime, $\sigma_L < L_{\rm SN}$. Here the optimal observable is an intermediate between the IM and GD observables. Fig. 2.10 shows one scenario in this regime. As in case Ia, one might naively apply a GD threshold at a few times σ_L. In the case Ia scenario, the voxel fluxes above the threshold are indeed "detected" since they rise above the instrumental noise and confusion. However, in case Ib, voxels above this threshold typically contain multiple sources with ℓ above the threshold, and the confusion noise from sources below the threshold is larger than the the sources at or just above the threshold. The regime of voxel fluxes $\sigma_L < L < L_{\rm SN}$ is thus more amenable to the IM technique. Individual sources can be detected with a threshold $L_{\rm th} \gtrsim L_{\rm SN}$ because only those sources rise above the confusion noise.

The resulting optimal observable can thus be understood as a hybrid between the two methods, detecting individual sources in the brightest voxels ($L > L_{\rm SN}$), and benefiting from IM in the fainter voxels that still rise above the instrumental noise ($\sigma_L < L < L_{\rm SN}$).

Fig. 2.10 indeed shows that neither the pure IM (linear) nor the pure GD (step function) observables capture the optimal information. The Fisher information for the optimal observable gains information in two stages, corresponding to the IM and GD parts, respectively. The total optimal Fisher information falls between $N(\ell > \sigma_L)$ and $N(\ell > L_{\rm SN})$, captured by GD and IM observables, respectively.

The detailed shape of the optimal observable depends on the luminosity function. In practice, we usually do not have sufficient knowledge of the source luminosity function, and it might be difficult to derive the optimal observable within our formalism. From our analysis, we know that the optimal observable in case Ib is GD above a threshold around $L_{\rm SN}$ and IM between that and another threshold around σ_L. Therefore, in practice, the optimal observable in the case Ib regime could be designed by choosing these two threshold scales, and by considering a linear function in between and a constant plateau about the upper threshold. By trying a range of values for both thresholds, the optimal threshold can be determined as the one giving the minimum shot-noise level in the power spectrum.

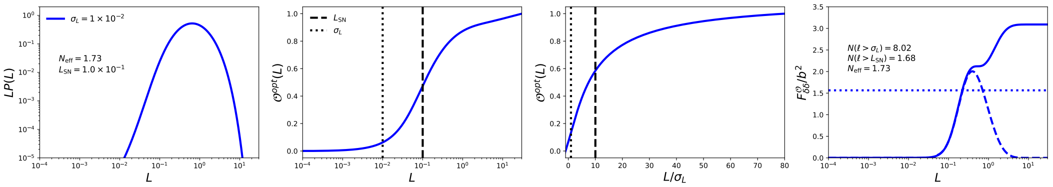

Figure 2.10: Case Ib example. **Left:** $P(L)$ after convolving with $\sigma_L = 0.01$. **Middle left:** optimal observables (solid line). The dashed and dotted lines are $L_{\rm SN}$ and σ_L, respectively. **Middle right:** same as the middle left panel but plot with L/σ_L in the x-axis. **Right**: integrated Fisher information for the optimal observable (solid), Fisher information of the step function observable as a function of step position (dashed), and the Fisher information of the linear observable (dotted).

Case Ic: $\mathbf{L}_{\mathrm{SN}} < \ell_* < \sigma_\mathbf{L}$ The final scenario in the $L_{\mathrm{SN}} < \ell_*$ ($N_{\mathrm{eff}} < 1$) regime is that of a very large instrument noise, $\sigma_L > \ell_*$. This is the case of noisy surveys, where only sources in the bright exponential tail of the Schecter function rise above the instrument noise.

Fig. 2.11 shows an example of case Ic. At first sight, the middle left panel appears to suggest that the optimal observable is close to a GD step function with a threshold at $\sim 6\sigma_L$. However, when we consider the actual step function, we see first that the optimal threshold lies at $\sim 1\sigma_L$, and second (from the right panel) that its information content is far from optimal. Inspecting the optimal observable in more detail, we see from the right panel that its information content is dominated by voxel luminosities up to $L \lesssim 3\sigma_L$. In this regime, as shown by the middle right panel, the optimal observable is close to linear (and voxel luminosities are noisy). Thus, the optimal observable is closer to the IM observable. This interpretation is confirmed by considering in the right panel the information contained in the IM observable, which is indeed close to optimal.

Since sources brighter than the noise are not confused ($L_{\mathrm{SN}} < \sigma_L$), one might a priori expect GD to be the optimal strategy, just like in case Ia. The reason the present case is different is that sources brighter than the instrument noise are in the exponential tail of the Schechter function. A detection threshold at a few times σ_L that unambiguously distinguishes sources above the threshold from noise fluctuations would detect only a very small number of sources and throw away information in almost all voxels. A slightly better approach is GD with a low threshold at $L \sim \sigma_L$. In this case, there are many false detections owing to the high instrumental noise, but a larger number of sources are probed. As discussed above, the approximately optimal approach is the IM observable, which gives an information content determined by the effective number of sources and the instrument noise suppression, $F^{\mathrm{IM}}_{\delta\delta}/b^2 = N_{\mathrm{eff}}\,\sigma^2_{\mathrm{SN}}/(\sigma^2_{\mathrm{SN}} + \sigma^2_L)$, larger than the information content given by the number of objects that can be detected, $(F^{\mathrm{GD}}_{\delta\delta}/b^2) \sim N(\ell > \sigma_L) \ll N_{\mathrm{eff}}$.

Case II: $\ell_* < L_{\mathrm{SN}}$

The defining criterion of case II, $\ell_* < L_{\mathrm{SN}}$, approximately corresponds to a large effective number of sources per voxel, $N_{\mathrm{eff}} > 1$. The $P(L)$ function here (at least in the $N_{\mathrm{eff}} \gg 1$ limit) can be approximated by a Gaussian with mean μ and variance $\bar{\sigma}^2$ given by

$$\mu = \int d\ell\, V_{\mathrm{vox}}\, \Phi(\ell)\, \ell = \langle \hat{L} \rangle, \qquad (2.54)$$

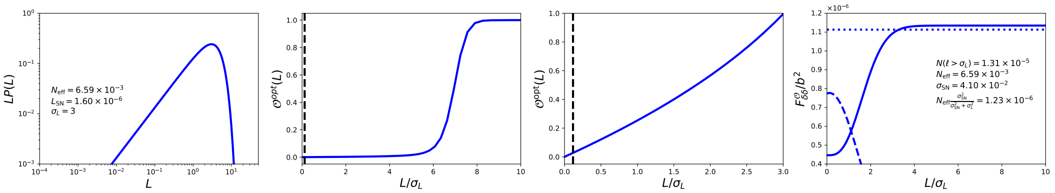

Figure 2.11: Case Ic example. **Left:** $P(L)$ after convolving with $\sigma_L = 3$. **Middle Left:** optimal observables (solid line). The dashed line is the optimal threshold for the step function observable, i.e. the peak of the dashed curve in the right panel. **Middle right:** optimal observable zoomed in to around $\sigma_L = 1.5$. **Right**: the integrated Fisher information for the optimal observable (solid), Fisher information of the step function observable as a function of step position (dashed), and the Fisher information of the linear observable (dotted).

and

$$\bar{\sigma}^2 = \int d\ell\, V_{\rm vox}\, \Phi(\ell)\, \ell^2 + \sigma_L^2 = \sigma_{\rm SN}^2 + \sigma_L^2. \tag{2.55}$$

Fig. 2.12 shows results for three different noise levels, corresponding to the three subclasses of case II: $\sigma_L < \ell_* < L_{\rm SN}$ (blue), $\ell_* < \sigma_L < L_{\rm SN}$ (red), and $\ell_* < L_{\rm SN} < \sigma_L$ (green).

As in the $N \gg 1$ case in the toy model (Sec. 2.3), we derive the optimal observable to be the sum of a linear and a quadratic term,

$$\mathcal{O}^{\rm opt}(L) = L' + \frac{\langle \hat{L}^2 \rangle}{2\,\mu\,\bar{\sigma}^2} L'^2, \tag{2.56}$$

where $L' = L - \mu$. The quadratic term has a negligible contribution to the optimal Fisher information, similarly to the toy model, so IM (the linear function observable) is the optimal strategy, and the optimal Fisher information $F^{\rm opt}_{\delta\delta} \sim F^{\rm IM}_{\delta\delta}/b^2$ has the upper bound $N_{\rm eff}$ (see Eq. 2.44) and drops as the noise goes up.

Schechter Luminosity Function Model Summary

In this section, we explore four different scenarios defined by different ordering of ℓ_*, $L_{\rm SN}$, and σ_L. Our formalism is not restricted to the IM or GD observable, but we found that in most cases either IM or GD is indeed the optimal strategy for mapping LSSs. Only in case Ib will an alternative strategy defined as the hybrid of the two will outperform a pure IM or pure GD observable, but case Ib is a very rare situation. None of the future surveys discussed in Sec. 2.5 are in the case Ib regime. Therefore, we conclude that the GD / IM dichotomy captures most of the optimal strategy in reality.

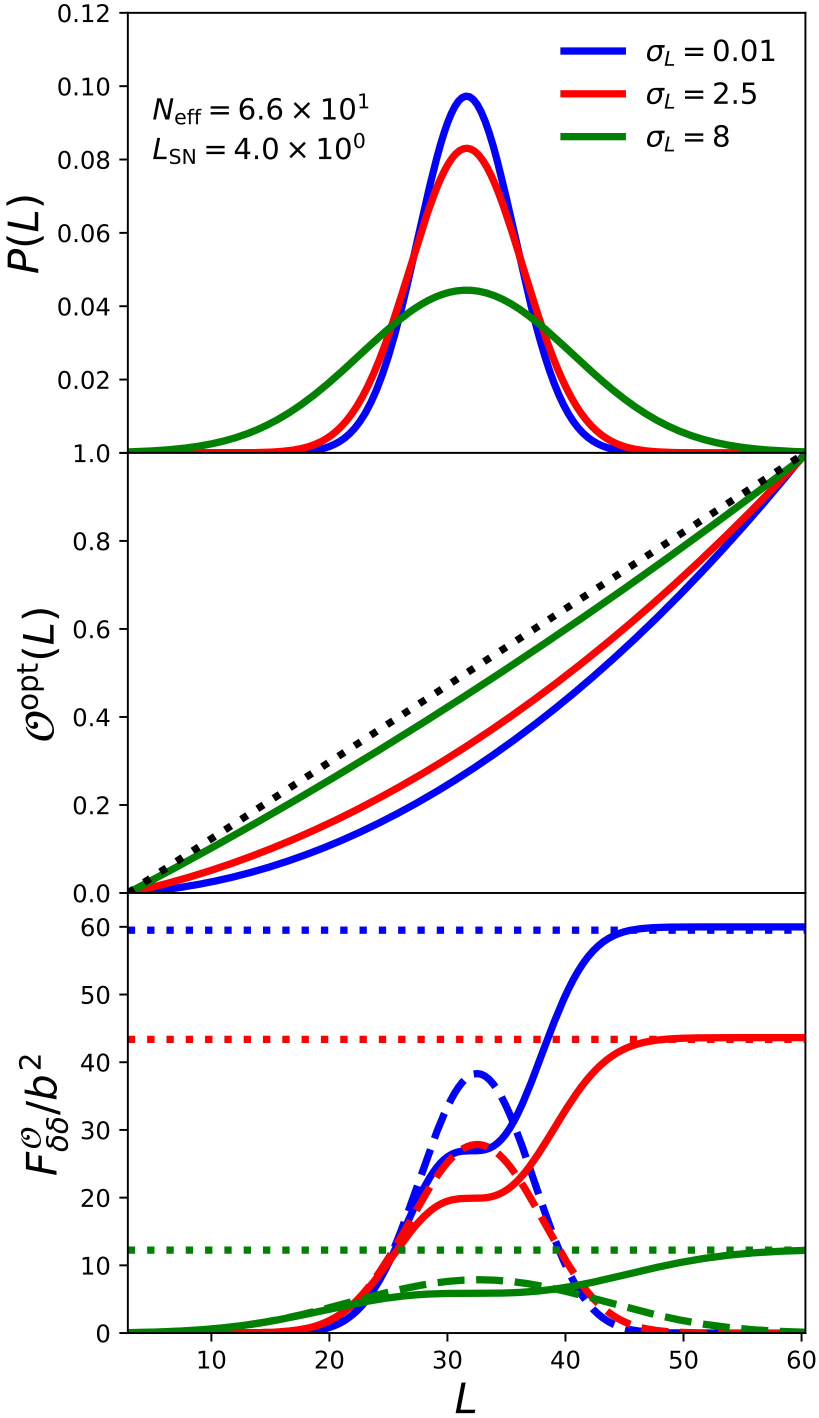

Figure 2.12: Case II with three different σ_L levels. **Top:** $P(L)$ after convolving with σ_L. **Middle:** optimal observables (solid lines). The dashed line is the linear observable for reference. **Bottom**: the integrated Fisher information for the optimal observable (solid), Fisher information of the step function observable as a function of step position (dashed), and the Fisher information of the linear observable (dotted).

2.5 Optimal Strategy for IM Experiments

We now apply the formalism we have developed to proposed and ongoing IM experiments. By simply calculating $L_{\rm SN}$, ℓ_*, and σ_L from experimental parameters and empirical line luminosity functions, we can categorize a survey into one of the cases in Sec. 2.4, and identify its optimal observable.

As discussed in Sec. 2.4, there exist ambiguous regimes where the cases will be classified as case I ($L_{\rm SN} < \ell_*$), but the confusion is significant ($N_{\rm eff} > 1$). Therefore, we also calculate $N_{\rm eff}$ for each experiment, and we label these cases I/II as they are intermediate, instead of classifying them into either one of the cases.

Below we consider several experiments targeting different spectral lines across redshift. The results for all the surveys and lines we discuss below are summarized in Table 2.1. We present the relevant parameters of each survey and leave the details in Appendix .6.

An important potential caveat to the discussion here is that we only include the instrumental noise as the noise term σ_L. In reality, astrophysical foreground contaminations, for example, are another source of noise, and their fluctuations could be much higher than the instrumental noise without any foreground mitigation procedure. These foregrounds may include both local contributions from the Milky Way galaxy and emissions from extragalactic sources. Fortunately, these foregrounds are in principle distinguishable from the line signal of interest because of their distinct spectral and spatial signatures, often being much smoother spectrally than the signal that enables us to remove them with the strategies advocated for foreground cleaning in 21 cm IM measurements [49, 59, 74]. Quantifying the effect of residual foregrounds requires a more sophisticated model, which is outside the scope of this work.

SPHEREx

SPHEREx is a planned space mission for an all-sky near-infrared spectro-imaging survey [28][9]. SPHEREx would carry out the first all-sky spectral survey at wavelengths between 0.75 and 2.42 μm (with spectral resolution $R = 41$), between 2.42 and 3.82 μm (with $R = 35$), between 3.82 and 4.42 μm (with $R = 110$), and between 4.42 and 5.00 μm (with $R = 130$), with a pixel size of 6.2″. We take the 5σ sensitivity to be $m_{AB} = 19.5$ and 22 per spectral channel, which is approximately the expected sensitivity in the all-sky and the deep regions ($2 \times \sim 100$ deg^2), re-

---

[9] `http://spherex.caltech.edu`.

spectively. SPHEREx is able to detect multiple lines, including Hα, Hβ, [O III], and Lyα, at different redshifts. Here we discuss the cases of Hα and Lyα.

Hα SPHEREx can detect the Hα line at $0.1 < z < 5$. We adopt the Hα luminosity function at $z = 2.23$ from Sobral et al. [72]: a Schechter function with $\log_{10}\phi^* = -2.78$ Mpc^{-3}, $\log_{10}\ell_* = 42.87$ erg/s, and $\alpha = -1.59$. We then derive from the luminosity function and instrument parameters that $L_{\rm SN}/\ell_* = 5.8 \times 10^{-5}$, $N_{\rm eff} = 2.2 \times 10^{-2}$, and $\sigma_L/\ell_* = 0.19$ (deep regions) and $\sigma_L/\ell_* = 1.9$ (all-sky). The all-sky survey is clearly in the case Ic regime, where IM is optimal. As for the deep regions, at first sight, it is in the case Ia regime ($L_{\rm SN} < \sigma_L < L_*$), where GD is the optimal strategy. However, since σ_L is close to ℓ_*, we are really at the boundary between the case Ia and the case Ic scenario, the latter suggesting that IM is preferred. Since we are in this gray area between the two regimes, an explicit calculation is required to check which approach is optimal. We thus computed the Fisher information for the linear and step function observables and found that the two approaches have the similar performance. Therefore, we label it with IM/GD as there is no preferred approach in this case.

Lyα The Lyα line from high redshifts ($5.2 < z < 8$) also falls within the SPHEREx bands. Here we use the Lyα luminosity function at $z = 5.56$ from Cassata et al. [12]: a Schechter function with $\phi^* = 9.2 \times 10^{-4}$ Mpc^{-3}, $\log_{10}\ell_* = 42.72$ erg/s, $\alpha = -1.69$, and from this we get $L_{\rm SN}/\ell_* = 2.1 \times 10^{-4}$, $N_{\rm eff} = 3.2 \times 10^{-2}$, and $\sigma_L/\ell_* = 6.4$ (deep regions) and $\sigma_L/\ell_* = 64$ (all-sky). Both are in the case Ic regime, so IM is again the optimal strategy.

CDIM

The Cosmic Dawn Intensity Mapper [CDIM, 21] is a NASA Probe Study designed for Cosmic Dawn and Epoch of Reionization studies, probing Lyα, Hα, and other spectral lines through cosmic history as part of its science goals. It plans to cover the wavelength range of $0.75 - 7.5\,\mu$m, with a spectral resolution of $R = 300$ and 1 arcsec^2 pixel size. The planned ~ 30 deg^2 deep surveys would reach a 5σ point-source sensitivity of $\mathrm{m_{AB}} = 22.5$. We calculate the Hα and Lyα line signals using the same luminosity functions described in the SPHEREx analysis above.

Hα For Hα at $z = 2.23$, we found $L_{\rm SN}/\ell_* = 1.8 \times 10^{-9}$, $N_{\rm eff} = 4.9 \times 10^{-5}$, and $\sigma_L/\ell_* = 9.8 \times 10^{-3}$. This is clearly inside the case Ia regime ($L_{\rm SN} < \sigma_L < \ell_*$), where the sources above the instrumental noise can be detected without confusion, so GD is the optimal strategy and the Fisher information is $\sim N(>\sigma_L)$.

Lyα For Lyα at $z = 5.56$, we have $L_{\rm SN}/\ell_* = 8.0 \times 10^{-8}$, $N_{\rm eff} = 1.4 \times 10^{-4}$, and $\sigma_L/\ell_* = 0.68$. This is at the boundary between the Ia and Ic scenarios, as with the SPHEREx Hα (deep regions) case, where IM and GD observables have the similar performance, so we label it with IM/GD.

We remind the reader that, to reach the conclusion that thresholded detection of individual lines is optimal for this survey, we have assumed that residual foregrounds can be ignored so that only the instrumental noise (and the shot noise in the line-emitting galaxies) enters the problem. Incorporating foregrounds (including continuum emission from extragalactic sources) in a realistic way may alter the conclusion on the optimal observable.

HETDEX

The Hobby-Eberly Telescope Dark Energy Experiment [HETDEX, 36][10] is a wide-field survey covering 300 deg^2 at the north Galactic cap. Its main science goal is to detect 0.8 million Lyα-emitting (LAE) galaxies within $1.9 < z < 3.5$ to provide a direct probe of dark energy at $z \sim 3$. The survey will have a $3'' \times 3''$ pixel size, and the spectral resolution is $R = 800$. The quoted sensitivity for 1200 s exposures per field is approximately $6 \times 10^{-17}{\rm erg/s/cm^2}$ (5σ), so we set $\sigma_{\rm L} = 1 \times 10^{-17}{\rm erg/s/cm^2}$ in our calculation.

Lyα Here we consider the Lyα measurement at $z = 2.5$ using the luminosity function from Cassata et al. [12] in their $1.95 < z < 3$ redshift bin (a Schechter function with $\phi^* = 7.1 \times 10^{-4}$ Mpc^{-3}, $\log_{10}\ell_* = 42.70\,{\rm erg/s}$, $\alpha = -1.6$). Then, we derive $L_{\rm SN}/\ell_* = 1.2 \times 10^{-8}$, $N_{\rm eff} = 1.3 \times 10^{-4}$, and $\sigma_L/\ell_* = 9.3 \times 10^{-2}$, which is also the in the Ia regime, so that line detection is the optimal strategy.

Although our calculations for CDIM and HETDEX for detecting Lyα indicate that galaxy/line detection is a better option than IM, we have assumed that the Lyα emission comes from point sources. However, Lyα photons are very often rescattered with nearby neutral hydrogen before they escape from galaxies, and thus the Lyα emission is extended. According to radiative transfer simulations, the extended Lyα halos have a size of tens or even hundreds of kpc [9, 41, 44, 86], which

---

[10] `www.hetdex.org`.

is comparable to the pixel size we consider here (the comoving voxel dimension in our Lyα calculation is $8.4 \times 0.027 \times 0.027\, Mpc/h$ and $3.5 \times 0.059 \times 0.059\, Mpc/h$ for CDIM and HEDEX, respectively). As a result, it is possible that IM is a better way to capture the extended Lyα emission; a more detailed investigation is needed to quantify the best observable for the Lyα line.

Another potential caveat is that the "GD" we discuss in this work is only based on the targeting line emission, while no external information is used for source detection. In reality, however, sources might be *detected* based on their full spectrum, and the line is then used to get its redshift. This is closer to the observing strategy for HETDEX. Since our model is not applicable for this type of survey strategy, a more sophisticated formalism is needed in order to quantify its ability to extract the LSS information.

TIME

TIME is a grating spectrometer dedicated to probe the [C II] line at $5.3 < z < 8.5$ [22]. The instrument has a spectral resolution of $R = 150$ and a pixel size of $0'.45$. The noise-equivalent intensity (NEI) is around $10^6 - 10^7 \mathrm{Jy}\sqrt{\mathrm{sec}}/\mathrm{sr}$, and we adopt $\mathrm{NEI} = 4 \times 10^6 \mathrm{Jy}\sqrt{\mathrm{sec}}/\mathrm{sr}$ for the calculation. The proposed 1000-hr survey gives an integration time per pixel of $\mathrm{t_{pix}} = 100$ hr, leading to $\sigma_{\mathrm{L}} = \mathrm{NEI}/\sqrt{2\,\mathrm{t_{pix}}} = 4.71 \times 10^3$ Jy/sr.

[C II] We now calculate the performance of TIME probing [C II] at $z = 6$. For the luminosity function, we adopt the semianalytic model from Popping et al. [62] (a Schechter function with $\phi^* = (\ln 10) 10^{-2.95}\,\mathrm{Mpc}^{-3}$, $\log_{10} \ell_* = 7.80\,\mathrm{L}_\odot$, $\alpha = -1.77$). From these we get $L_{\mathrm{SN}}/\ell_* = 1.9 \times 10^{-2}$, $N_{\mathrm{eff}} = 0.75$, and $\sigma_L/\ell_* = 2.17$. This is in the case Ic regime, where IM is the optimal strategy.

COMAP

The CO Mapping Array Pathfinder [COMAP, 18] aims at tracing star formation through cosmic time with the CO rotational transition lines. COMAP will observe in the 30-34 GHz window with a 40 MHz spectral resolution, corresponding to CO(1-0) at $2.4 < z < 2.8$ and CO(2-1) at $5.8 < z < 6.7$. Following the formalism and the instrument parameters of the Pathfinder in Li et al. [46], we obtain a pixel size of $2'.55$ and a system noise of $23\ \mu K$.

CO(1-0) We now consider the CO(1-0) line at $z = 3$. For the luminosity function at $z = 3$, we take the averaged value of each of the three Schechter function parameters for $z = 2$ and $z = 4$ in Popping et al. [62]: $\phi^* = (\ln 10)10^{-2.79}$ Mpc^{-3}, $\log_{10} \ell_* = 7.28$ Jy km s^{-1} Mpc^2, $\alpha = -1.62$. From these we get $L_{\rm SN}/\ell_* = 1.4 \times 10^{-1}$, $N_{\rm eff} = 2.5$, and $\sigma_L/\ell_* = 13$, so this is near the borderline of the Ic ($L_{\rm SN} < \ell_* < \sigma_L$) and II regimes ($N_{\rm eff} > 1$), where IM is the optimal strategy in both cases.

CHIME

The Canadian Hydrogen Intensity Mapping Experiment [CHIME, 2] is a cylindrical interferometer designed to measure the neutral hydrogen HI power spectrum at $0.8 < z < 2.5$. We consider the HI signal at $z = 1$. The instrument has a $15 - 25$ arcmin angular resolution, and we adopt 15 arcmin as the pixel size. The frequency resolution is 390 kHz [2], and the noise level at $z = 1$ is $\sigma_T = 2.9 \times 10^{-4}$K for 1.4 yr of integration, calculated from the survey parameters given in Bandura et al. [2] (see Appendix. .6 for the derivation).

For the HI luminosity function, we use the local ($z < 0.06$) HI observations from Martin et al. [53], in which the HI mass function is fitted with a Schechter function with $\phi_* = 4.8\ \mathrm{h}_{70}^3$ Mpc^{-3} dex^{-1}, $\log(\mathrm{M}_*/\mathrm{M}_\odot) + 2\log \mathrm{h}_{70} = 9.96$, and $\alpha = -1.33$, and we ignore redshift evolution from $z = 1$ to the present day. See Appendix .6 for converting the HI mass function to the luminosity function.

With this information in hand, we get $L_{\rm SN}/\ell_* = 0.63$, $N_{\rm eff} = 4.2$, and $\sigma_L/\ell_* = 3.4$, which is again near the borderline of Ic and II regimes, where IM is optimal for both cases. We stress again that this is a calculation for an idealized situation that ignores foreground effects.

The above analysis focuses on the 3D line IM experiments. Two-dimensional continuum surveys such as the cosmic infrared background (CIB) experiments are also worth discussing in this context, given that they usually suffer from confusion [3, 80, 82], which induces errors in measuring the properties of bright sources (e.g. the position and flux error from confusion noise described in Hogg [37]). Another common issue in the CIB experiments is the correlated confusion noise, which refers to the fact that the fluctuations from the faint, unresolved sources are spatially correlated with the bright sources. Our $P(D)$ formalism intrinsically captures the dependency of the density of all the sources and their underlying overdensity field δ, regardless of the detection limit, and thus it is a suitable way to quantify the confusion in CIB. However, according to the observations, the CIB source luminosity function

Table 2.1: Summary of the survey targets and their expected σ_L, ℓ_*, and $L_{\rm SN}$ relation.

survey	Line	redshift	σ_L/ℓ_*	$L_{\rm SN}/\ell_*$	Ls' Relation	$N_{\rm eff}$	Case	Optimal Strategy
SPHEREx (deep regions)	Hα	2.23	0.19	5.8×10^{-5}	$L_{\rm SN} < \sigma_L \lesssim \ell_*$	2.2×10^{-2}	Ia/Ic	GD/IM[a]
	Lyα	5.56	6.4	2.1×10^{-4}	$L_{\rm SN} < \ell_* < \sigma_L$	3.2×10^{-2}	Ic	IM
SPHEREx (all-sky)	Hα	2.23	1.9	5.8×10^{-5}	$L_{\rm SN} < \ell_* < \sigma_L$	2.2×10^{-2}	Ic	IM
	Lyα	5.56	64	2.1×10^{-4}	$L_{\rm SN} < \ell_* < \sigma_L$	3.2×10^{-2}	Ic	IM
CDIM	Hα	2.23	9.8×10^{-3}	1.8×10^{-9}	$L_{\rm SN} < \sigma_L < \ell_*$	4.9×10^{-5}	Ia	GD
	Lyα	5.56	0.68	8.0×10^{-8}	$L_{\rm SN} < \sigma_L \lesssim \ell_*$	1.4×10^{-4}	Ia/Ic	GD/IM[a]
HETDEX	Lyα	2.5	9.3×10^{-2}	1.2×10^{-8}	$L_{\rm SN} < \sigma_L < \ell_*$	1.3×10^{-4}	Ia	GD
TIME	[C II]	6	2.17	1.9×10^{-2}	$L_{\rm SN} < \ell_* < \sigma_L$	7.5×10^{-1}	Ic	IM
COMAP	CO(1-0)	3	13	1.4×10^{-1}	$L_{\rm SN} < \ell_* < \sigma_L$	2.5	Ic/II	IM
CHIME	HI	1	3.4	0.63	$L_{\rm SN} < \ell_* < \sigma_L$	4.2	Ic/II	IM

---

[a] These cases are at the boundary of Ia and Ic, so we contfirm that the IM is better than GD by numerically calculating their $P(L)$ and their Fisher information of the GD, IM, and optimal observable.

is close to a simple power law without an exponential cutoff at the bright end [80]. Therefore unlike the Schechter function, there is no characteristic ℓ_* we can use to compare with σ_L and $L_{\rm SN}$ to classify the regimes. A detailed $P(D)$ analysis is needed to study this different kind of luminosity function, and we leave it to the future works.

2.6 Example Application: Pixel Size Optimization

In this section, we use our framework to calculate the information content as a function of pixel (or beam) size. The choice of pixel size in a survey is a trade-off between confusion and instrumental noise, which are quantified by $L_{\rm SN}$ (or $N_{\rm eff}$) and σ_L, respectively. A smaller pixel size gives less confusion, but the instrumental noise σ_L/ℓ_* also changes according to the properties of the dominant noise source and how the integration time and collecting area scaled with the pixel size. The two effects cannot be treated independently if our observable is not a linear function, and thus it requires a full $P(D)$ analysis to construct the $P(L)$ distribution and then to derive the Fisher information.

We consider changing the pixel size from $\Omega_{\rm pix}$ to $a\,\Omega_{\rm pix}$, while fixing the spectral bandwidth per voxel. Here a is a rescaling parameter that quantifies the change in pixel size relative to a fiducial survey configuration, and we would thus like to compute $N_{\rm eff}$, σ_L, and ultimately the Fisher information in the new pixel, as a function of a. The voxel volume and $N_{\rm eff}$ trivially scale linearly with a. The exact effect on the instrumental noise per voxel depends on the details of the experiment and on how its specifications are varied as the pixel size is changed, as we will discuss in more detail below. With the variation in voxel size and σ_L, we can calculate the Fisher information in the new $a\,\Omega_{\rm pix}$ voxel. However, it is not sufficient to simply consider the variation (with a) in the Fisher information per voxel. A smaller pixel size gives a larger number of pixels to constrain the underlying δ for a fixed survey region. Therefore, the meaningful quantity for the performance of different voxel sizes is $F(a)/a$, where $F(a)$ is the Fisher information of a single voxel with size $a\,\Omega_{\rm pix}$. The quantity $F(a)/a$ gives the information content on δ for a fixed survey region.

The scaling of σ_L/ℓ_* is derived from comparing the number of photons from a ℓ_* source and the rms of the number of photons from noise for a given integration time.

The number of photons $N_{\rm src}$ from a ℓ_* source per voxel per integration $t_{\rm int}$ is given by

$$N_{\mathrm{src}} = \frac{\ell_*}{4\pi D_L(z)^2} A_{\mathrm{coll}} t_{\mathrm{int}}. \tag{2.57}$$

We assume that the instrument's collecting area A_{coll}, is fixed by the aperture size, and we assume a fixed total integration time/survey duration and a fixed total sky coverage for the survey. If we change the angular size of pixel from Ω_{pix} to $a\,\Omega_{\mathrm{pix}}$ by moving the focal length of the telescope, while fixing the physical configuration of the detector (the physical pixel size and number of pixels on the detector stay the same), the instantaneous field of view also scaled with a, and thus the integration time per pixel t_{int} becomes $a\,t_{\mathrm{int}}$ in order to preserve the total integration time of the survey. Therefore, we get $N_{\mathrm{src}} \propto a$.

As for the noise, below we will focus on two simple scenarios for the instrumental noise scaling with pixel size: a read-noise-dominated case and a photon-noise-dominated case. We will apply these two scalings relative to a fiducial experiment given by the SPHEREx $H\alpha$ case, presented in Sec. 2.5.

Photon-noise-dominated scenario For the photon noise, we assume that the dominant photon source from the sky is a uniform bright foreground, e.g. the zodiacal light in the optical/near infrared. Say this foreground has surface brightness I, which has units Jy sr^{-1}. The number of photons N_I from I per voxel per integration is thus

$$N_I = I\, A_{\mathrm{coll}}\, \Omega_{\mathrm{pix}}\, \delta\nu\, t_{\mathrm{int}} \propto a^2, \tag{2.58}$$

where $\delta\nu$ is the bandwidth, and we take it unchanged while varying the pixel size. The photon noise is the Poisson noise of N_I, and thus the rms of photon noise σ_{ph} is

$$\sigma_{\mathrm{ph}} = \sqrt{N_I} \propto a. \tag{2.59}$$

Therefore, the scaling of σ_L/ℓ_* with a is proportional to $\sigma_{\mathrm{ph}}/N_{\mathrm{src}}$, which is a constant independent of voxel size.

Read-noise-dominated scenario For the read noise, assuming we only read at the beginning and the end of the integration, and each read has rms σ_{read} electrons, the expected rms number of photon of read noise σ_{RN} thus does not scale with a. As a result, $\sigma_L/\ell_* = \sigma_{\mathrm{RN}}/N_{\mathrm{src}}$ scales with $1/a$.

Fig. 2.13 shows the Fisher information ($F(a)/a$) for varying pixel/voxel size in the SPHEREx Hα case, normalized by the Fisher information for the fiducial 6.2 arcsec pixel size. As shown in the plot, if the noise is dominated by read noise, increasing the voxel size will have a dramatic improvement on information gain, since this crosses the transition from Ic (IM) to Ia (GD) (see the bottom panel), and we expect a lot more information gain from individual detection.

Here we only demonstrate a simple and idealized example of using this framework to quantify the information with different pixel sizes. We remind the reader that the scaling relation with pixel size we adopted here is not a unique behavior in the photon-noise- and read-noise-dominated cases. In reality, the pixel size can be changed in different ways (e.g. change the physical configuration of the pixels on the detector itself) and results in different scaling relation.

In addition, the discussion above assumes the fixed total survey volume. In reality, we can optimize the experiments by varying the survey volume as well. There is another trade-off between the survey volume and the depth (σ_L in our context) for the given observing time. Increasing the total survey volume reduces the cosmic variance in the power spectrum. In this work, our formalism only accounts for the variance on the voxel-by-voxel basis, which corresponds to the shot noise in the power spectrum. In reality, cosmic variance is another noise source in the power spectrum that plays a significant role in the large-scale (low-k) mode uncertainty. To optimize the survey for probing the large-scale power spectrum, an analysis taking into account both the shot noise and cosmic variance is needed. We leave the consideration to future works.

2.7 Conclusion

We use a general “observable” as a weight function to turn the observed voxel flux map into the observable map that traces the LSS. The two well-studied approaches, GD and IM, are two special observable cases. The performance of observables is quantified by the Fisher information, and from it we derive the optimal observable, which is able to extract the full information content in the data.

We first work on a toy model assuming that all the targeting sources have the same flux ℓ. By considering a range of source density N (number of sources per voxel) and instrument noise level σ_L, we derive the optimal observable and its Fisher information for each case and compare it with the Fisher information of the GD and IM observables. In the toy model, we found that IM is preferred when the sources are either confused ($N > 1$) or suppressed by the noise ($\sigma_L > \ell$).

Next we move on to a more general model with the source population follows

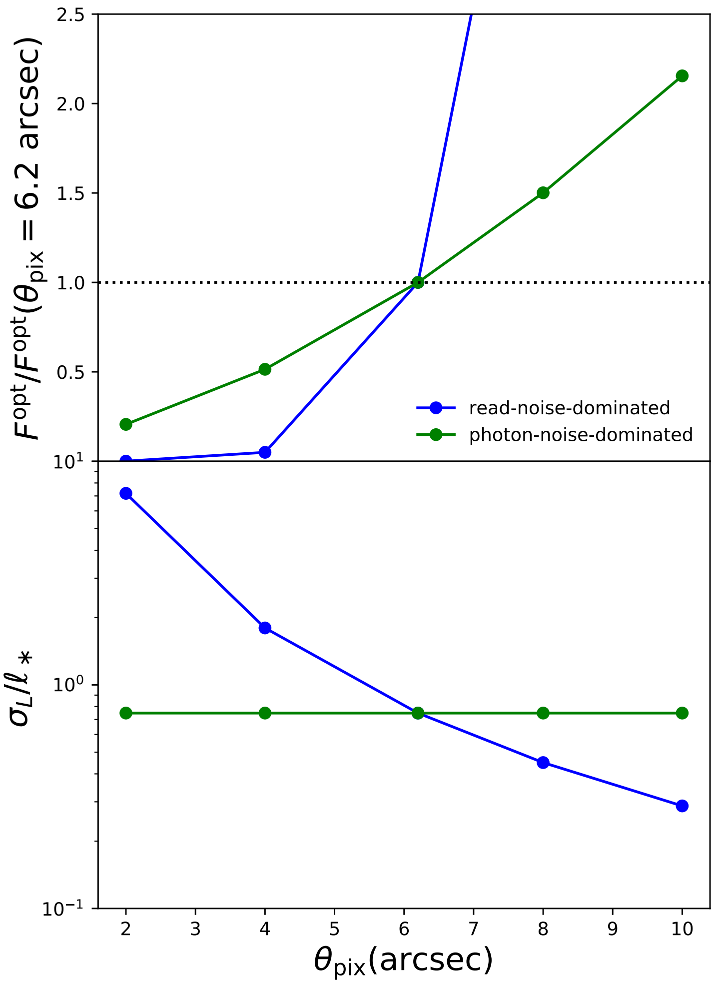

Figure 2.13: **Top**: The Fisher information of SPHEREx Hα case with different pixel size. The Fisher information is normalized by the Fiducial 6.2 arcsec pixel size case. **Bottom**: σ_L/ℓ_* ratio in each scenarios.

Schechter function form. Then, we identify four limiting regimes depending on the relative value of the three scales: $\{L_{\rm SN}, \sigma_L, \ell_*\}$. Again, we found that in the high-noise ($\sigma_L > \ell_*$, case Ic) or high-confusion ($N_{\rm eff} > 1$ or $L_{\rm SN} > \ell_*$, case II) regime, the IM observable is preferred, as it reaches the performance of the optimal observable. Whereas on the opposite situation ($N_{\rm eff} < 1$ and $\sigma_L < \ell_*$), we can further identify two distinct scenarios. The first one is where $L_{\rm SN} < \sigma_L < \ell_*$ (case Ia), such that all the voxels above the noise are not confused, so the detection with a threshold around σ_L is the preferred strategy. The other scenario is where $\sigma_L < L_{\rm SN} < \ell_*$ (case Ib). In this case, the optimal strategy is the hybrid of the IM and GD observables. The IM observable is suitable for the voxels above noise but highly confused ($\sigma_L < L < L_{\mathrm{SN}}$), whereas for voxels above L_{SN}, the voxel flux is dominated by a single bright source, and thus the GD is the favored choice for them.

Finally, we demonstrate the usage of this formalism with two applications. The first application is to identify the optimal strategy for the proposed (and ongoing) IM experiments (e.g. SPHEREx, TIME, COMAP). The second application is to calculate the information content for different pixel sizes in a survey. Although we have made some simplified assumptions in these two demonstrations, the formalism we developed here can be easily applied to optimizing the experiment parameters of interest with their own specification of noise and confusion level.

.1 Proving $F_{\delta\delta}^{\mathrm{opt}} = F_{\delta\delta}$

Here we prove that the Fisher information per voxel of optimal observable $F_{\delta\delta}^{\mathrm{opt}}$ is equal to $F_{\delta\delta}$, the maximum Fisher information per voxel that any observable can possibly attain. Writing out each element in Eq. 2.13 explicitly, we get

$$\begin{aligned}\langle \hat{O}^{\mathrm{opt}} \rangle &= \int dL\, P(L,\delta)\, O^{\mathrm{opt}}(L) \\ &= \int dL\, P(L,\delta)\, \partial_\delta \ln P(L,\delta) \\ &= \partial_\delta \int dL\, P(L,\delta) = 0\end{aligned} \tag{60}$$

$$\begin{aligned}\langle (\hat{O}^{\mathrm{opt}})^2 \rangle &= \int dL\, P(L,\delta)\, (O^{\mathrm{opt}})^2(L) \\ &= \int dL\, P(L,\delta)\, (\partial_\delta \ln P(L,\delta))^2 \\ &= \langle (\partial_\delta \ln P(L,\delta))^2 \rangle = F_{\delta\delta}\end{aligned} \tag{61}$$

$$\begin{aligned}\partial_\delta \langle \hat{O}^{\mathrm{opt}} \rangle &= \int dL\, \partial_\delta P(L,\delta)\, O^{\mathrm{opt}}(L) \\ &= \int dL\, P(L,\delta)\, (\partial_\delta \ln P(L,\delta))\, O^{\mathrm{opt}}(L) \\ &= \langle (\hat{O}^{\mathrm{opt}})^2 \rangle = F_{\delta\delta},\end{aligned} \tag{62}$$

and thus

$$F_{\delta\delta}^{\mathrm{opt}} = \frac{\left(\partial_\delta \langle \hat{O}^{\mathrm{opt}} \rangle\right)^2}{\langle (\hat{O}^{\mathrm{opt}})^2 \rangle - \langle \hat{O}^{\mathrm{opt}} \rangle^2} = F_{\delta\delta}. \tag{63}$$

.2 Comparing Linear and Quadratic Terms in the Toy Model $N \gg 1$ Optimal Observable

To explain why the quadratic term has a negligible contribution to the optimal Fisher information in the toy model $N \gg 1$ case (Sec. 2.3), below we explicitly calculate the components of Fisher information in Eq. 2.13 for the linear ($O^{\mathrm{lin}}(L) = L'$) and quadratic ($O^{\mathrm{quad}}(L) \equiv \frac{\ell}{2\bar{\sigma}^2} L'^2$) terms in Eq. 2.34 respectively (note that $L' \equiv L - N\ell$, which is also the peak of the Gaussian $P(L)$ profile). The signals on these two components are

$$\begin{aligned} \partial_\delta \langle \hat{O}^{\mathrm{lin}} \rangle &= bN\ell \\ \partial_\delta \langle \hat{O}^{\mathrm{quad}} \rangle &= \frac{bN\ell}{2} \left(\frac{\ell^2}{\sigma_L^2 + N\ell^2} \right). \end{aligned} \tag{64}$$

Since this is in the $N \gg 1$ regime, the signal from the quadratic term is always much smaller than from the linear term, regardless of the instrument noise σ_L. The variance terms of the two observables are

$$\begin{aligned} \langle (\hat{O}^{\mathrm{lin}})^2 \rangle - \langle \hat{O}^{\mathrm{lin}} \rangle^2 &= \bar{\sigma}^2 - 0 = \sigma_L^2 + N\ell^2 \\ \langle (\hat{O}^{\mathrm{quad}})^2 \rangle - \langle \hat{O}^{\mathrm{quad}} \rangle^2 &= \left(\frac{\ell}{2\,\bar{\sigma}^2} \right)^2 \left[3\,\bar{\sigma}^4 - \left(\bar{\sigma}^2 \right)^2 \right] = \ell^2/2. \end{aligned} \tag{65}$$

Again, with the $N \gg 1$ condition, the contribution from the quadratic term is also negligible[11]. Hence, the contribution of the quadratic term to the Fisher information is negligible, which implies a purely linear (IM) observable can reach the optimal performance.

.3 Explaining $F_{\delta\delta}^{\mathrm{IM}} \propto N$

The Fisher information of the IM observable is given by

$$F_{\delta\delta}^{\mathrm{IM}} = \frac{\left(\partial_\delta \langle \hat{L} \rangle \right)^2}{\langle \hat{L}^2 \rangle - \langle \hat{L} \rangle^2}, \tag{66}$$

where

$$\langle \hat{L} \rangle \quad = V_{\mathrm{vox}} \int d\ell \, \Phi(\ell) \, \ell \; \propto N \tag{67}$$

$$\langle \hat{L}^2 \rangle \quad = V_{\mathrm{vox}} \int d\ell \, \Phi(\ell) \, \ell^2 \; \propto N. \tag{68}$$

---

[11] To compare the Fisher information of purely linear observable with the full optimal observable (linear + quadratic), one also has to take into account the covariance term of these two observables $\langle \hat{O}^{\mathrm{lin}} \hat{O}^{\mathrm{quad}} \rangle$. Fortunately, this term vanished since it is an odd function with respect to the Gaussian $P(L)$ profile.

Below we will prove that the the numerator of $F^{\mathrm{IM}}_{\delta\delta}$ is proportional to N^2, and the denominator is proportional to N, and thus $F^{\mathrm{IM}}_{\delta\delta}$ is proportional to N.

The "signal" term is proportional to N since $\partial_\delta\langle\hat{L}\rangle \propto \partial_\delta N = bN$. As for the variance $\sigma^2(\hat{L}) = \langle\hat{L}^2\rangle - \langle\hat{L}\rangle^2$, we note the fact that we can divide each voxel into N_{sub} subvoxels, where the subvoxel fluxes $\hat{L}^{\mathrm{sub}}_i$ are independent of each other, so the total $\hat{L}$ is simply the sum of the subvoxel flux $\hat{L}^{\mathrm{sub}}_i$, and the variance $\sigma^2(\hat{L})$ is also the sum of the subvoxel variance $\sigma^2(\hat{L}^{\mathrm{sub}}_i)$, $\sigma^2(\hat{L}) = N_{\mathrm{sub}}\sigma^2(\hat{L}^{\mathrm{sub}}_i)$, as the subvoxels are independent. The subvoxel variance is given by

$$\sigma^2(\hat{L}^{\mathrm{sub}}_i) = \frac{V_{\mathrm{vox}}}{N_{\mathrm{sub}}}\int d\ell\,\Phi(\ell)\,\ell^2 - \left(\frac{V_{\mathrm{vox}}}{N_{\mathrm{sub}}}\int d\ell\,\Phi(\ell)\,\ell\right)^2. \tag{69}$$

We have the freedom to choose N_{sub} large enough such that the second term is much smaller than the first term, so $\sigma^2(\hat{L}^{\mathrm{sub}}_i) \propto V_{\mathrm{vox}}$ (and N), and the total voxel variance $\sigma^2(\hat{L}) = N_{\mathrm{sub}}\sigma^2(\hat{L}^{\mathrm{sub}}_i)$ is also proportional to V_{vox}(and N).

.4 Different Choice of ℓ_{min}

Here we will justify that the choice of ℓ_{min} does not affect the optimal observable and its information content. We compare the difference between fiducial $\ell_{\mathrm{min}} = 10^{-3}$ and $\ell_{\mathrm{min}} = 5\times10^{-4}$ cases, while keeping other parameters the same. The results are shown in Fig. .14. The optimal observable is different in the absence of noise. However, if the instrumental noise is much higher than ℓ_{min} (e.g. $\sigma_L = 10^{-2}$ in this example), the effect of the artificial cutoff ℓ_{min} is totally obscured by the noise, and thus both O^{opt} and $F^{\mathrm{opt}}_{\delta\delta}$ are nearly identical in the two cases here. Therefore, we justify that the arbitrary choice of the ℓ_{min} does not affect the optimal observable and Fisher information as long as the cutoff ℓ_{min} is much lower than the instrument noise σ_L.

.5 Different Choice of α

Here we show how the different faint-end slope α affects the optimal observable and the Fisher information. Fig. .15 compares the cases of fiducial $\alpha = -1.5$ with steeper faint-end slope $\alpha = -2$, while keeping other parameters the fiducial values. In the noiseless scenario, the optimal observable of the $\alpha = -2$ case has the step at lower L compared to $\alpha = -1.5$ case. This naturally reflects the fact that there are more faint sources in the $\alpha = -2$ case. When a $\sigma_L = 10^{-2}$ instrumental noise is applied, the difference is washed out by the noise. Another interesting feature is the peak in the $O^{\mathrm{opt}}(L)$ function for the $\alpha = -2$ case, which can be explained by the fact that the voxels with luminosity around the peak are more likely to have multiple

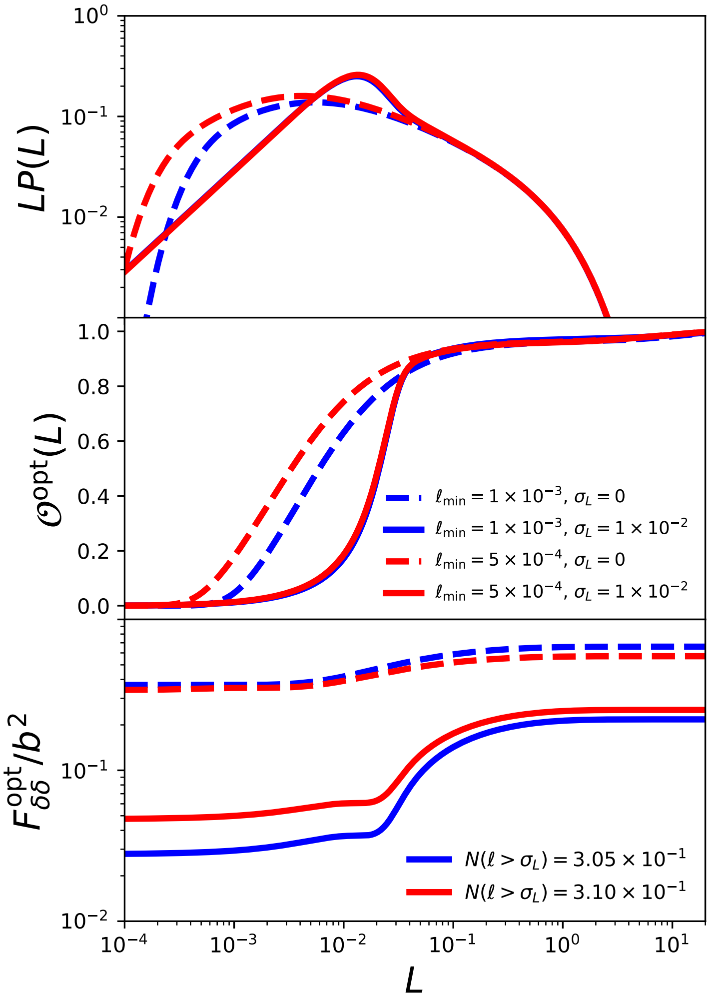

Figure .14: Fiducial Schechter function faint-end slope $\alpha = -1.5$ with and without instrumental noise $\sigma_L = 0.01$ and using two different $\ell_{\min}$. **Top:** $P(L)$ with (dashed lines) and without (solid lines) instrumental noise. **Middle:** optimal observables for each case. **Bottom:** the integrated Fisher information for the optimal observable.

sources, whereas higher-L voxels are mostly contributed by a single bright source. Because we assume a luminosity-independent bias, the source number density traces the underlying δ linearly, and thus the voxels around the peak are likely tracing the higher density field than the even brighter voxels. This does not happen in the $\alpha = -1.5$ case because of its lack of faint sources to reach this special regime.

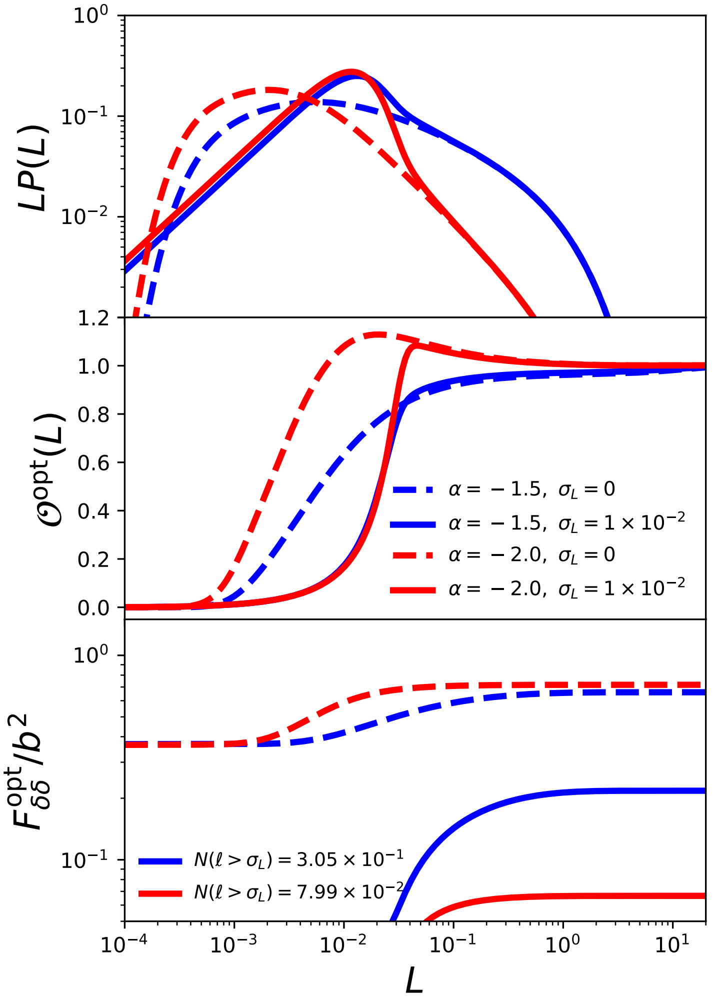

Figure .15: Fiducial case with two different α. **Top:** $P(L)$ of two different α with (solid lines) and without (dashed lines) instrumental noise noise $\sigma_L = 0.01$. **Middle:** the optimal observables for each case. **Bottom:** the integrated Fisher information for the optimal observable.

.6 Unit Conversion of the Survey Parameters

In Sec. 2.5, we derive the ℓ_*, $L_{\rm SN}$, and σ_L from the targeting source Schechter function parameters and the survey parameters (angular/spectral resolution and sensitivity). Here we provide the implementation details of the conversion from the observed quantities, which come with different units in the literature, to the final source luminosity, in $L_\odot$ or erg s^{-1}.

- Comoving voxel size $V_{\rm vox}$
  Consider that the targeting spectral line has the rest frequency $\nu_{\rm rest}$ at redshift z. The survey has the angular pixel size $\Omega_{\rm pix}$ (we use the beam size instead if the survey does not specify their pixelization) and the spectral resolution $R = \nu_{\rm obs}/\delta\nu_{\rm obs}$, where $\nu_{\rm obs} = (1+z)\nu_{\rm rest}$ is the observed frequency. Then, the comoving voxel size is

  $$V_{\rm vox} = \Omega_{\rm pix} \left[D_A^{CM}(z)\right]^2 \frac{c\,(1+z)}{H(z)\,R} \tag{70}$$

  where c is the speed of light, $H(z)$ is the Hubble parameter, and $D_A^{CM}(z)$ is the comoving angular diameter distance, which equals to the comoving distance in the flat ($\Omega_k = 0$) universe.

- Deriving $L_{\rm SN}$ from the Schechter parameters
  With the comoving voxel size and the luminosity function, we can calculate the $\sigma_{\rm SN}(\ell)$ following Eq. 2.45,

  $$\sigma^2_{\rm SN}(\ell) = V_{\rm vox}\,\phi_* \int_0^\ell d\ell'\,\ell'^{\alpha+2}\,e^{-\ell'}, \tag{71}$$

  and we find out $L_{\rm SN}$ numerically with the definition $\sigma_{\rm SN}(L_{\rm SN}) = L_{\rm SN}$.

- Deriving σ_L from the experiment sensitivity
  The conversion of the instrumental noise to σ_L is derived by matching the rms of noise flux F^n to the source emission line flux F^s. Below we will work with flux defined by power per area (in the units of W m^2). The flux F^s from a line luminosity ℓ source is given by

  $$F^s = \frac{\ell}{4\pi D_L^2(z)}, \tag{72}$$

  where $D_L(z)$ is the luminosity distance. As for the noise, if it is quoted as the "flux density" $F^n_\nu[erg/s/cm^2/Hz]$, the noise flux F^n is given by

  $$F^n = F^n_\nu\,\delta\nu_{\rm obs} = F_\nu\,(\nu_{\rm obs}/R). \tag{73}$$

  The σ_L is then defined by the ℓ scale where $F_s = F_n$, and thus

  $$\sigma_L = 4\pi\,D_L^2(z)\,F^n_\nu\,\nu_{\rm obs}/R. \tag{74}$$

If the sensitivity is quoted in m_{AB} instead, then the flux density F_ν^n is given by $F_\nu^n = 3631 \times 10^{-m_{AB}/2.5}[Jy]$. If this is the 5σ sensitivity, then we use $F_\nu^n/5$ in the σ_L calculation in Eq. 74.

If the noise level is quoted in intensity $I_\nu^n[Jy/sr]$, then the conversion to the noise flux density per voxel is $F_\nu^n = I_\nu^n \Omega_{\rm pix}$. Finally, when noise is in the units of brightness temperature T, the intensity I_ν^n can be derived using $I_\nu^n = 2\nu_{\rm obs} k_B T/c^2$, and then we can get σ_L with the equations listed above.

- Velocity-integrated luminosity
  Popping et al. [62] quote their CO luminosity function in the "velocity-integrated luminosity" L^V (Jy km s^{-1} Mpc^2), which is the "luminosity density" (in units proportional to W Hz^{-1}) per observed velocity. To convert it to the intrinsic luminosity unit $[L_\odot]$, we use the formalism in Obreschkow et al. [57] Appendix A:

$$\frac{L}{L_\odot} = 1.040 \times 10^{-3} \left(\frac{\nu_{\rm obs}}{\rm GHz}\right) \left(\frac{1+z}{4\pi}\right) \frac{L^V}{\rm Jy\ km\ s^{-1}\ Mpc^2}. \tag{75}$$

- HI mass-to-light ratio
  To convert the HI mass function to the luminosity function, we follow the equation in Draine [29] in the optically thin limit,

$$M_{\rm HI} = 4.945 \times 10^7 M_\odot \left(\frac{D_L}{Mpc}\right)^2 \left(\frac{F^s}{\rm Jy\ MHz}\right). \tag{76}$$

  Combining with Eq. 72, we obtain the mass-to-light ratio

$$\frac{M_{\rm HI}}{M_\odot} = 1.56 \times 10^8 \frac{L_{\rm HI}}{L_\odot}. \tag{77}$$

- CHIME instrument noise
  We calculate the CHIME instrument noise using the parameters in Seo et al. [69]. The noise rms per voxel is (in the temperature unit)

$$\sigma_T = \frac{g T_{\rm sky} + T_a}{\sqrt{t_{\rm int}\, \Delta f}} \tag{78}$$

  where g is the gain and $T_{\rm sky}$ and $T_{\rm a}$ are the sky and antenna temperature, respectively. Δf is the bandwidth, and $t_{\rm int}$ is the integration time per pixel:

$$t_{\rm int} = N_{\rm year}\, D_f\, \frac{1}{2\pi} \frac{\lambda_{\rm obs}}{W_{\rm cyl}} \tag{79}$$

where $N_{\rm year}$ is the total integration time, D_f is the duty factor, $\lambda_{\rm obs}$ is the observed wavelength (42 cm at $z = 1$), and $W_{\rm cyl}$ is the width of the cylinder. We use the parameter values listed in Seo et al. [69]: $N_{\rm year} = 1.4$ yr, $D_f = 0.5$, $W_{\rm cyl} = 14.3$ m, which gives $t_{\rm int} = 3.3 \times 10^{-3}$ yr. Then, we take $T_{\rm sky} = 50$ K, $T_{\rm a} = 10$ K, $g = 0.8$, $\Delta f = 390$ kHz, and we get $\sigma_T = 2.9 \times 10^{-4}$ K.

References

[1] Anderson, C. J., Luciw, N. J., Li, Y.-C., et al. 2018, MNRAS, 476, 3382, doi: 10.1093/mnras/sty346

[2] Bandura, K., Addison, G. E., Amiri, M., et al. 2014, in Proc. SPIE, Vol. 9145, Ground-based and Airborne Telescopes V, 914522, doi: 10.1117/12.2054950

[3] Béthermin, M., Wu, H.-Y., Lagache, G., et al. 2017, A&A, 607, A89, doi: 10.1051/0004-6361/201730866

[4] Bowman, J. D., Morales, M. F., & Hewitt, J. N. 2009, ApJ, 695, 183, doi: 10.1088/0004-637X/695/1/183

[5] Breysse, P. C., Kovetz, E. D., Behroozi, P. S., Dai, L., & Kamionkowski, M. 2017, MNRAS, 467, 2996, doi: 10.1093/mnras/stx203

[6] Breysse, P. C., Kovetz, E. D., & Kamionkowski, M. 2014, MNRAS, 443, 3506, doi: 10.1093/mnras/stu1312

[7] —. 2016, MNRAS, 457, L127, doi: 10.1093/mnrasl/slw005

[8] Breysse, P. C., & Rahman, M. 2017, MNRAS, 468, 741, doi: 10.1093/mnras/stx451

[9] Cantalupo, S., Porciani, C., Lilly, S. J., & Miniati, F. 2005, ApJ, 628, 61, doi: 10.1086/430758

[10] Carilli, C. L. 2011, ApJL, 730, L30, doi: 10.1088/2041-8205/730/2/L30

[11] Carron, J., & Szapudi, I. 2013, MNRAS, 434, 2961, doi: 10.1093/mnras/stt1215

[12] Cassata, P., Le Fèvre, O., Garilli, B., et al. 2011, A&A, 525, A143, doi: 10.1051/0004-6361/201014410

[13] Chang, T.-C., Pen, U.-L., Bandura, K., & Peterson, J. B. 2010, Nature, 466, 463, doi: 10.1038/nature09187

[14] Chang, T.-C., Pen, U.-L., Peterson, J. B., & McDonald, P. 2008, Physical Review Letters, 100, 091303, doi: 10.1103/PhysRevLett.100.091303

[15] Chapman, E., Abdalla, F. B., Harker, G., et al. 2012, MNRAS, 423, 2518, doi: 10.1111/j.1365-2966.2012.21065.x

[16] Cheng, Y.-T., Chang, T.-C., Bock, J., Bradford, C. M., & Cooray, A. 2016, ApJ, 832, 165, doi: 10.3847/0004-637X/832/2/165

[17] Chung, D. T., Viero, M. P., Church, S. E., et al. 2019, ApJ, 872, 186, doi: 10.3847/1538-4357/ab0027

[18] Cleary, K., Bigot-Sazy, M.-A., Chung, D., et al. 2016, in American Astronomical Society Meeting Abstracts, Vol. 227, American Astronomical Society Meeting Abstracts, 426.06

[19] Colless, M., Peterson, B. A., Jackson, C., et al. 2003, ArXiv Astrophysics e-prints

[20] Comaschi, P., & Ferrara, A. 2016, MNRAS, 455, 725, doi: 10.1093/mnras/stv2339

[21] Cooray, A., Bock, J., Burgarella, D., et al. 2016, ArXiv e-prints. https://arxiv.org/abs/1602.05178

[22] Crites, A. T., Bock, J. J., Bradford, C. M., et al. 2014, in Proc. SPIE, Vol. 9153, Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy VII, 91531W, doi: 10.1117/12.2057207

[23] Croft, R. A. C., Miralda-Escudé, J., Zheng, Z., Blomqvist, M., & Pieri, M. 2018, MNRAS, 481, 1320, doi: 10.1093/mnras/sty2302

[24] Croft, R. A. C., Miralda-Escudé, J., Zheng, Z., et al. 2016, MNRAS, 457, 3541, doi: 10.1093/mnras/stw204

[25] Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, AJ, 145, 10, doi: 10.1088/0004-6256/145/1/10

[26] Dawson, K. S., Kneib, J.-P., Percival, W. J., et al. 2016, AJ, 151, 44, doi: 10.3847/0004-6256/151/2/44

[27] DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 2016, ArXiv e-prints. https://arxiv.org/abs/1611.00036

[28] Doré, O., Bock, J., Ashby, M., et al. 2014, ArXiv e-prints. https://arxiv.org/abs/1412.4872

[29] Draine, B. T. 2011, Physics of the Interstellar and Intergalactic Medium

[30] Fonseca, J., Silva, M. B., Santos, M. G., & Cooray, A. 2017, MNRAS, 464, 1948, doi: 10.1093/mnras/stw2470

[31] Furlanetto, S. R., Oh, S. P., & Briggs, F. H. 2006, PhysRep, 433, 181, doi: 10.1016/j.physrep.2006.08.002

[32] Gong, Y., Cooray, A., Silva, M., et al. 2012, ApJ, 745, 49, doi: 10.1088/0004-637X/745/1/49

[33] Gong, Y., Cooray, A., Silva, M. B., Santos, M. G., & Lubin, P. 2011, ApJL, 728, L46, doi: 10.1088/2041-8205/728/2/L46

[34] Gong, Y., Silva, M., Cooray, A., & Santos, M. G. 2014, ApJ, 785, 72, doi: 10.1088/0004-637X/785/1/72

[35] Guzzo, L., Scodeggio, M., Garilli, B., et al. 2014, A&A, 566, A108, doi: 10.1051/0004-6361/201321489

[36] Hill, G. J., Gebhardt, K., Komatsu, E., et al. 2008, in Astronomical Society of the Pacific Conference Series, Vol. 399, Panoramic Views of Galaxy Formation and Evolution, ed. T. Kodama, T. Yamada, & K. Aoki, 115. https://arxiv.org/abs/0806.0183

[37] Hogg, D. W. 2001, AJ, 121, 1207, doi: 10.1086/318736

[38] Jones, D. H., Read, M. A., Saunders, W., et al. 2009, MNRAS, 399, 683, doi: 10.1111/j.1365-2966.2009.15338.x

[39] Keating, G. K., Marrone, D. P., Bower, G. C., et al. 2016, ApJ, 830, 34, doi: 10.3847/0004-637X/830/1/34

[40] Keating, G. K., Bower, G. C., Marrone, D. P., et al. 2015, ApJ, 814, 140, doi: 10.1088/0004-637X/814/2/140

[41] Kollmeier, J. A., Zheng, Z., Davé, R., et al. 2010, ApJ, 708, 1048, doi: 10.1088/0004-637X/708/2/1048

[42] Kovetz, E. D., Viero, M. P., Lidz, A., et al. 2017, ArXiv e-prints. https://arxiv.org/abs/1709.09066

[43] Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, ArXiv e-prints. https://arxiv.org/abs/1110.3193

[44] Laursen, P., & Sommer-Larsen, J. 2007, ApJL, 657, L69, doi: 10.1086/513191

[45] Lee, S. K., Ando, S., & Kamionkowski, M. 2009, JCAP, 7, 007, doi: 10.1088/1475-7516/2009/07/007

[46] Li, T. Y., Wechsler, R. H., Devaraj, K., & Church, S. E. 2016, ApJ, 817, 169, doi: 10.3847/0004-637X/817/2/169

[47] Lidz, A., Furlanetto, S. R., Oh, S. P., et al. 2011, ApJ, 741, 70, doi: 10.1088/0004-637X/741/2/70

[48] Lidz, A., & Taylor, J. 2016, ApJ, 825, 143, doi: 10.3847/0004-637X/825/2/143

[49] Liu, A., & Tegmark, M. 2011, PhRvD, 83, 103006, doi: 10.1103/PhysRevD.83.103006

[50] —. 2012, MNRAS, 419, 3491, doi: 10.1111/j.1365-2966.2011.19989.x

[51] LSST Science Collaboration, Abell, P. A., Allison, J., et al. 2009, ArXiv e-prints. https://arxiv.org/abs/0912.0201

[52] Madau, P., Meiksin, A., & Rees, M. J. 1997, ApJ, 475, 429

[53] Martin, A. M., Papastergis, E., Giovanelli, R., et al. 2010, ApJ, 723, 1359, doi: 10.1088/0004-637X/723/2/1359

[54] Mashian, N., Sternberg, A., & Loeb, A. 2015, JCAP, 11, 028, doi: 10.1088/1475-7516/2015/11/028

[55] Masui, K. W., Switzer, E. R., Banavar, N., et al. 2013, ApJL, 763, L20, doi: 10.1088/2041-8205/763/1/L20

[56] Morales, M. F., Bowman, J. D., & Hewitt, J. N. 2006, ApJ, 648, 767, doi: 10.1086/506135

[57] Obreschkow, D., Heywood, I., Klöckner, H.-R., & Rawlings, S. 2009, ApJ, 702, 1321, doi: 10.1088/0004-637X/702/2/1321

[58] Parkinson, D., Riemer-Sørensen, S., Blake, C., et al. 2012, PhRvD, 86, 103518, doi: 10.1103/PhysRevD.86.103518

[59] Parsons, A. R., Pober, J. C., Aguirre, J. E., et al. 2012, ApJ, 756, 165, doi: 10.1088/0004-637X/756/2/165

[60] Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 594, A13, doi: 10.1051/0004-6361/201525830

[61] Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2018, ArXiv e-prints. https://arxiv.org/abs/1807.06209

[62] Popping, G., van Kampen, E., Decarli, R., et al. 2016, MNRAS, 461, 93, doi: 10.1093/mnras/stw1323

[63] Pullen, A. R., Chang, T.-C., Doré, O., & Lidz, A. 2013, ApJ, 768, 15, doi: 10.1088/0004-637X/768/1/15

[64] Pullen, A. R., Doré, O., & Bock, J. 2014, ApJ, 786, 111, doi: 10.1088/0004-637X/786/2/111

[65] Pullen, A. R., Serra, P., Chang, T.-C., Doré, O., & Ho, S. 2018, MNRAS, 478, 1911, doi: 10.1093/mnras/sty1243

[66] Righi, M., Hernández-Monteagudo, C., & Sunyaev, R. A. 2008, A&A, 489, 489, doi: 10.1051/0004-6361:200810199

[67] Schechter, P. 1976, ApJ, 203, 297, doi: 10.1086/154079

[68] Scott, D., & Rees, M. J. 1990, MNRAS, 247, 510

[69] Seo, H.-J., Dodelson, S., Marriner, J., et al. 2010, ApJ, 721, 164, doi: 10.1088/0004-637X/721/1/164

[70] Silva, M., Santos, M. G., Cooray, A., & Gong, Y. 2015, ApJ, 806, 209, doi: 10.1088/0004-637X/806/2/209

[71] Silva, M. B., Santos, M. G., Gong, Y., Cooray, A., & Bock, J. 2013, ApJ, 763, 132, doi: 10.1088/0004-637X/763/2/132

[72] Sobral, D., Smail, I., Best, P. N., et al. 2013, MNRAS, 428, 1128, doi: 10.1093/mnras/sts096

[73] Spergel, D., Gehrels, N., Baltay, C., et al. 2015, ArXiv e-prints. https://arxiv.org/abs/1503.03757

[74] Switzer, E. R., Chang, T.-C., Masui, K. W., Pen, U.-L., & Voytek, T. C. 2015, ApJ, 815, 51, doi: 10.1088/0004-637X/815/1/51

[75] Switzer, E. R., Masui, K. W., Bandura, K., et al. 2013, MNRAS, 434, L46, doi: 10.1093/mnrasl/slt074

[76] Takada, M., Ellis, R. S., Chiba, M., et al. 2014, PASJ, 66, R1, doi: 10.1093/pasj/pst019

[77] Taylor, A. N., & Watts, P. I. R. 2001, MNRAS, 328, 1027, doi: 10.1046/j.1365-8711.2001.04874.x

[78] Tegmark, M., Taylor, A. N., & Heavens, A. F. 1997, ApJ, 480, 22, doi: 10.1086/303939

[79] Uzgil, B. D., Aguirre, J. E., Bradford, C. M., & Lidz, A. 2014, ApJ, 793, 116, doi: 10.1088/0004-637X/793/2/116

[80] Viero, M. P., Moncelsi, L., Quadri, R. F., et al. 2013, ApJ, 779, 32, doi: 10.1088/0004-637X/779/1/32

[81] Visbal, E., & Loeb, A. 2010, JCAP, 11, 016, doi: 10.1088/1475-7516/2010/11/016

[82] Wang, W.-H., Lin, W.-C., Lim, C.-F., et al. 2017, ApJ, 850, 37, doi: 10.3847/1538-4357/aa911b

[83] Wyithe, J. S. B., & Loeb, A. 2008, MNRAS, 383, 606, doi: 10.1111/j.1365-2966.2007.12568.x

[84] York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, AJ, 120, 1579, doi: 10.1086/301513

[85] Yue, B., Ferrara, A., Pallottini, A., Gallerani, S., & Vallini, L. 2015, MNRAS, 450, 3829, doi: `10.1093/mnras/stv933`

[86] Zheng, Z., Cen, R., Weinberg, D., Trac, H., & Miralda-Escudé, J. 2011, ApJ, 739, 62, doi: `10.1088/0004-637X/739/2/62`

Chapter 3

SPECTRAL LINE DE-CONFUSION IN AN INTENSITY MAPPING SURVEY

Cheng, Y.-T., Chang, T.-C., Bock, J., Bradford, C. M., & Cooray, A. 2016, ApJ, 832, 165, doi: 10.3847/0004-637X/832/2/165

Spectral line intensity mapping (LIM) has been proposed as a promising tool to efficiently probe the cosmic reionization and the large-scale structure. Without detecting individual sources, LIM makes use of all available photons and measures the integrated light in the source confusion limit to efficiently map the three-dimensional matter distribution on large scales as traced by a given emission line. One particular challenge is the separation of desired signals from astrophysical continuum foregrounds and line interlopers. Here we present a technique to extract large-scale structure information traced by emission lines from different redshifts, embedded in a three-dimensional intensity mapping data cube. The line redshifts are distinguished by the anisotropic shape of the power spectra when projected onto a common coordinate frame. We consider the case where high-redshift [C II] lines are confused with multiple low-redshift CO rotational lines. We present a semi-analytic model for [C II] and CO line estimates based on the cosmic infrared background measurements, and show that with a modest instrumental noise level and survey geometry, the large-scale [C II] and CO power spectrum amplitudes can be successfully extracted from a confusion-limited data set, without external information. We discuss the implications and limits of this technique for possible LIM experiments.

3.1 Introduction

Line intensity mapping (LIM) has emerged as a promising tool to probe the three-dimensional structure of the Universe. Several emission lines have been proposed to uniquely trace the cosmic reionization process, revealing properties of the ionizing sources and the intergalactic medium at high redshifts, and to efficiently map the large-scale matter distribution in a large cosmic volume, suitable for cosmological studies at lower redshifts.

In contrast to traditional large-scale structure surveys, intensity mapping (IM) operates in the confusion-limited regime without thresholding to identify individual sources; rather, IM makes use of integrated light emission from all sources, including unresolved faint galaxies, to statistically measure properties of the light tracers and the underlying matter distribution. In addition, with LIM, where the tracer is a particular spectral line emission, the three-dimensional matter distribution can be faithfully represented on large cosmic scales. The 21-cm hyperfine emission from neutral hydrogen [9, 34, 50, 60], the CO rotational lines [4, 6, 20, 24, 27, 28, 35, 45, 47, 57], the [C ii] 157.7-μm fine structure line [19, 52, 56, 61], and the Lyman-α emission line [11, 21, 46, 53] are amongst the most studied such tracers in the LIM regime.

One of the main challenges in an LIM experiment is the separation of signals from the astrophysical foreground continuum emissions and line interlopers. The continuum foreground issue has been studied mostly in the context of 21cm IM, where the Galactic and extragalactic synchrotron and free-free radiations overwhelm the expected signals by several orders of magnitude [e.g., 3, 10, 18, 32, 37, 41, 54]; line interlopers, on the other hand, are a pressing issue for other line probes in the electromagnetic spectrum that is crowded with other line features.

Several studies have proposed strategies for deblending lines in an IM survey, by means of masking and cross-correlation. The masking technique makes use of an external galaxy catalog from galaxy surveys to identify and remove bright sources in order to reduce potential foreground contaminations [5, 52, 61]. On the other hand, cross-correlation of an LIM survey with an external data set tracing the same cosmic volume can help extract signals of interest; the method has been proposed, in particular, in the studies of reionization[7, 19, 21, 30, 52, 57], and has been successfully applied to extract LIM signals at lower redshifts against continuum foregrounds [8, 14, 36]. Aside from these two methods, Kogut et al. [26] makes use of the companion lines to directly identify [C ii] line intensity in each voxel. de Putter et al. [16] propose to use angular fluctuations of the light to reconstruct the 3D source luminosity density.

In an IM experiment, the intrinsic observing coordinates are angular and spectral coordinates defined by the instrument and survey geometry, which, given a known redshift, are mapped into comoving coordinates in the redshift space, before a power spectrum is computed. In the linear regime, any line tracers supposedly follow the matter distribution, which is isotropic in their respective real-space coordinates. If, however, without a priori knowledge of redshifts, lines at different redshifts embedded in an observing volume are blindly projected into the same comoving coordinates, they will exhibit different anisotropic 3D shapes in that frame, due to the incorrect redshift projection. This is key to the line separation technique we employ in this paper. The idea has been previously suggested by Visbal & Loeb [57] and Gong et al. [21], and recently investigated by Lidz & Taylor [29].

To demonstrate this technique, we consider a 3D LIM observing volume with high-redshift [C II] emissions blended with multiple lower-redshift CO rotational lines. We present a halo-model-based formalism to estimate [C II] and CO line strengths and power spectra across redshifts. After projecting the observed volume onto a common comoving frame, the resulting total power spectrum is a superposition of [C II] and CO power spectra at different redshifts, each with a different but predictable 3D shape due to the projection which we use as templates. We generate simulated data and use the Markov Chain Monte Carlo (MCMC) formalism to extract power spectrum parameters based on the templates.

The paper is organized as follows. In Sec. 3.2, we describe a model to estimate the [C II] and CO power spectra across redshifts, and the formalism for expressing the 3D power spectra of both lines in the comoving frame of [C II]. In Sec. 3.3, we discuss the details of the template fitting and MCMC implementation. The results are presented in Sec. 3.4. In Sec. 3.5, we discuss the implication and limitation of our method, and conclude in Sec. 3.6. Throughout this paper, we consider a flat ΛCDM cosmology with $n_s = 0.97$, $\sigma_8 = 0.82$, $\Omega_m = 0.26$, $\Omega_b = 0.049$, $\Omega_\Lambda = 0.69$, and $h = 0.68$, consistent with the latest measurement from *Planck* [44].

3.2 Power Spectrum Modeling

Here we provide an estimate of the [C II] and CO power spectra as a function of redshift. Our modeling is based on the halo-model formalism [12]. With the [C II] and CO luminosity ($L_{\rm CII}$, $L_{\rm CO}$) and halo mass (M) relations, the power spectrum of a spectral line intensity field can be calculated with the halo model.

We build our model (referred to as the CIB model hereafter) based on Planck Collaboration et al. [43]; the authors fitted the $L_{\rm IR}$ and and halo mass relation with cosmic infrared background (CIB) emission as measured by *Planck*. Details of the CIB model are provided in Appendix .1.

Then we connect the $L_{\rm IR}$ to $L_{\rm CII}$ and $L_{\rm CO}$. We adopt these relations based on observations and simulations in the literature. See Appendix .1 for more details.

Power Spectrum with Halo Model

With the luminosity-halo mass relation at hand, we derive the power spectrum in the halo-model framework. The comoving power spectrum consists of the one-halo and two-halo terms, which account for the correlation within the halos and between halos, respectively. The one-halo term of the [C II] or CO line is given by

$$P_{\mathrm{1h}}^{\mathrm{line}}(k,z) = \int_{M_{\mathrm{min}}}^{M_{\mathrm{max}}} dM \frac{dN}{dM} \, |u(k|M)|^2 \times \left(\frac{L_{\mathrm{line}}(M,z)}{4\pi D_L^2(z)} y_{\mathrm{line}}(z) D_A^2(z) \right)^2 , \tag{3.1}$$

where $L_{\mathrm{line}}(M,z)$ is the luminosity of [C II] or CO for a given halo mass M at redshift z, D_L is the proper luminosity distance, D_A is the comoving angular diameter distance, $u(k|M)$ is the Fourier transform of the normalized halo density profile, and we adopt the NFW profile in this work [38]. $y_{line}(z) \equiv d\chi/d\nu = \lambda_{line}(1+z)^2/H(z)$, where χ is the comoving distance, H is the Hubble parameter, ν is the observed frequency, and λ_{line} is the rest-frame wavelength of the line.

The two-halo term can be written as

$$P_{\mathrm{2h}}^{\mathrm{line}}(k,z) = [\int_{M_{\mathrm{min}}}^{M_{\mathrm{max}}} dM \frac{dN}{dM} \, |u(k|M)| \, b(M,z) \times \left(\frac{L_{\mathrm{line}}(M,z)}{4\pi D_L^2(z)} y_{\mathrm{line}}(z) D_A^2(z) \right)]^2 P_{\mathrm{lin}}(k), \tag{3.2}$$

where $b(M,z)$ is the halo bias [51], and $P_{\mathrm{lin}}(k)$ is the linear matter power spectrum.

We also consider the shot noise power spectrum due to the discretization of sources:

$$P_{\mathrm{sh}}^{line}(z) = \int_{M_{\mathrm{min}}}^{M_{\mathrm{max}}} dM \frac{dN}{dM} \left(\frac{L_{\mathrm{line}}(M,z)}{4\pi D_L^2(z)} y_{line}(z) D_A^2(z) \right)^2 . \tag{3.3}$$

The total comoving power spectrum is thus

$$P_{\mathrm{tot}}^{\mathrm{line}}(k,z) = P_{\mathrm{1h}}^{line}(k,z) + P_{2h}^{\mathrm{line}}(k,z) + P_{\mathrm{sh}}^{\mathrm{line}}(z). \tag{3.4}$$

Fig. 3.1 shows the comoving isotropic power spectrum of [C II] at $z = 6$ based on the CIB model, and the lower-redshift CO power spectra, which overlap in the same observing frequency range.

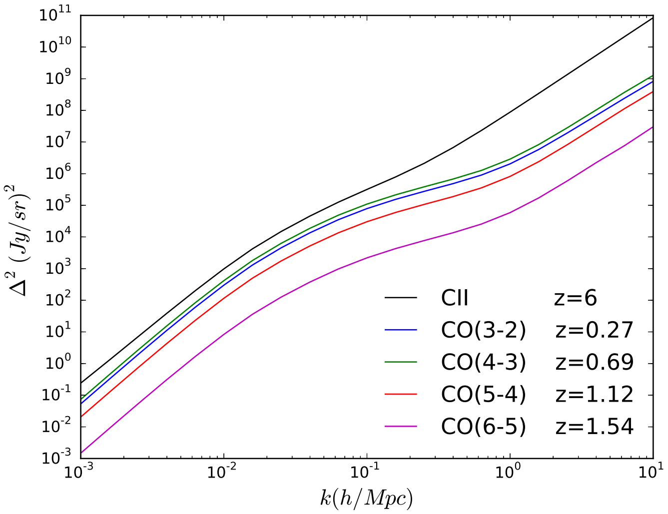

Figure 3.1: Comoving isotropic power spectra of [C II] at $z = 6$ and CO interlopers at lower redshfits in the same observing frequency range.

3D Power Spectrum

The [C II] and CO power spectra $P(\mathbf{k})$ is isotropic in real-space on large scales, so the power spectra only depend on $|\mathbf{k}|$ but not the direction of the $\mathbf{k}$ vector. However, an anisotropic feature comes in due to observational effects, which introduce the dependence on μ, the cosine of the angle between the $\mathbf{k}$ vector, and the line-of-sight direction. Below we discuss the projection effect and the redshift space distortions (RSDs) that give rise to the anisotropy in the 3D power spectrum.

Projection Effect

In a redshifted [C II] IM experiment, a CO line signal from redshift z_{CO} will blend with the [C II] signal from z_{CII} if they have the same observing frequency. The redshifts of these two lines follow

$$\nu_{\mathrm{obs}} = \frac{\nu_{\mathrm{CII}}}{(1+z_{\mathrm{CII}})} = \frac{\nu_{\mathrm{CO}}}{(1+z_{\mathrm{CO}})}, \tag{3.5}$$

where ν_{obs} is the observed frequency, and ν_{CII} and ν_{CO} are the rest-frame frequencies of the [C II] and CO lines, which are 1902 GHz and 115J GHz for the J to $J-1$ transition, respectively.

Both [C II] and CO power spectra are isotropic in their respective comoving frames. However, in the confusion limit, we may incorrectly project the observed [C II] and CO signals onto the comoving frame of [C II] for the power spectrum calculation.

In this case, the CO power spectrum is no longer isotropic. This is caused by the different redshift projection factors parallel and perpendicular to the line of sight, which makes the CO 3D data cube stretched more in one direction than the other. Below we provide the formalism for calculating the 3D power spectrum of a low-z CO signal projected onto the high-z [C II] comoving frame.

The projection in the direction parallel to the line of sight can be derived by considering an observed frequency range $d\nu_{\rm obs}$, which corresponds to either the [C II] from a comoving size $R^{\parallel}_{\rm CII}$ at z_{CII}, or the CO signal with comoving size $R^{\parallel}_{\rm CO}$ at $z_{\rm CO}$. The relation between $R^{\parallel}_{\rm CII}$ and $R^{\parallel}_{\rm CO}$ is

$$\begin{aligned}\frac{R^{\parallel}_{CO}}{R^{\parallel}_{CII}} &= \frac{d\chi(z_{CO})/d\nu_{\rm obs}}{d\chi(z_{CII})/d\nu_{\rm obs}} = \frac{y_{CO}(z_{CO})}{y_{CII}(z_{CII})}\\ &= \frac{\lambda_{CO}(1+z_{CO})^2/H(z_{CO})}{\lambda_{CII}(1+z_{CII})^2/H(z_{CII})}\\ &= \frac{(1+z_{CO})/H(z_{CO})}{(1+z_{CII})/H(z_{CII})}.\end{aligned} \tag{3.6}$$

Since Fourier wavenumber $k^{\parallel} \propto 1/R^{\parallel}$, we obtain

$$k^{\parallel}_{CO} = k^{\parallel}_{CII}\frac{H(z_{CO})(1+z_{CII})}{H(z_{CII})(1+z_{CO})} \equiv k^{\parallel}_{CII} r^{\parallel}(z_{CII}, J), \tag{3.7}$$

where J labels the CO transition from J to $J-1$, and $r^{\parallel}(z_{CII}, J)$ is the conversion factor for projecting the scale at $z_{\rm CO}$ to $z_{\rm CII}$ in the parallel direction.

The transverse scale of CO will be projected to the scale of [C II] corresponding to the same observed angle θ. Hence, the scale conversion relation in the perpendicular direction is

$$\theta = \frac{R^{\perp}_{CO}}{D_A(z_{CO})} = \frac{R^{\perp}_{CII}}{D_A(z_{CII})}, \tag{3.8}$$

where D_A is the comoving angular diameter distance. Thus

$$k^{\perp}_{CO} = k^{\perp}_{CII}\frac{D_A(z_{CII})}{D_A(z_{CO})} \equiv k^{\perp}_{CII} r^{\perp}(z_{CII}, J), \tag{3.9}$$

where $r^{\perp}(z_{CII}, J)$ is the conversion factor for projecting the scale at $z_{\rm CO}$ to $z_{\rm CII}$ in the perpendicular direction. Besides the shift in k value, the projection changes the comoving voxel volume $V_{\rm vox}$ and induces an amplitude change in the power spectrum. Since the power spectrum is proportional to $1/V_{\rm vox}$ at fixed intensity fluctuation, the projected CO power spectrum needs to be multiplied by the change in voxel volume $r^{\parallel}(r^{\perp})^2$.

The CO projected power spectrum $P_{\mathrm{CO}}^{\mathrm{prj}}$ can thus be written as

$$P_{CO}^{prj}(k_{CII}^{\perp}, k_{CII}^{\parallel}, z_{CII}, J) = r^{\parallel}(r^{\perp})^2 P_{CO}(k_{CO}, z_{CO}), \tag{3.10}$$

where $k_{\mathrm{CO}} = \sqrt{(k_{\mathrm{CO}}^{\perp})^2 + (k_{\mathrm{CO}}^{\parallel})^2}$, and P_{CO} is the comoving CO power spectrum.

For completeness, we also write down the [C II] power spectrum in the same coordinate,

$$P_{CII}^{prj}(k_{CII}^{\perp}, k_{CII}^{\parallel}, z_{CII}) = P_{CII}(k_{CII}, z_{CII}), \tag{3.11}$$

where $k_{CII} = \sqrt{(k_{CII}^{\perp})^2 + (k_{CII}^{\parallel})^2}$.

Redshift Space Distortions

Here we incorporate the RSD effects. We consider the linear Kaiser effect [23] describing the coherent motion of structure growth on large scales, which enhances the power spectrum, and the suppression on small scales due to the nonlinear virial motion, which we write as an exponential damping term [42]. The comoving one-halo and two-halo power spectrum can be written as [59]

$$\begin{aligned} P_{\mathrm{1h}}^{\mathrm{line}}(k, \mu, z) &= (1+\beta\mu^2)^2 \\ &\times \int_{M_{\mathrm{min}}}^{M_{\mathrm{max}}} dM \frac{dN}{dM} [\left(\frac{L_{\mathrm{line}}(M,z)}{4\pi D_L^2(z)} y_{\mathrm{line}}(z) D_A^2(z) \right)^2 \\ &\times |u(k|M)|^2 \, e^{-(k\sigma_v\mu)^2/2}], \end{aligned} \tag{3.12}$$

$$\begin{aligned} P_{\mathrm{2h}}^{\mathrm{line}}(k, \mu, z) &= P_{\mathrm{lin}}(k)(1+\beta\mu^2)^2 \\ &\times \{ \int_{M_{\mathrm{min}}}^{M_{\mathrm{max}}} dM \frac{dN}{dM} [\left(\frac{L_{\mathrm{line}}(M,z)}{4\pi D_L^2(z)} y_{\mathrm{line}}(z) D_A^2(z) \right) \\ &\times |u(k|M)| \, b(M,z) e^{-(k\sigma_v\mu)^2/2}]\}^2, \end{aligned} \tag{3.13}$$

where $(1+\beta\mu^2)^2$ is the Kaiser effect and $e^{-(k\sigma_v\mu)^2/2}$ is the exponential damping term. $\beta \equiv f/\bar{b}_{\mathrm{line}}$, where $f = dlnD/dlna$ is the logarithm derivative of the linear growth rate $D(z)$ with respect to the scale factor $a = 1/(1+z)$, and $\bar{b}_{\mathrm{line}}$ is the luminosity-weighted bias of the tracer, which we consider to be a constant on large scales. σ_v is the 1D velocity dispersion within halo mass M. Assuming the halos are isothermal, the velocity dispersion can be estimated as

$$\sigma_v^2 = \frac{GM}{2r_{\mathrm{vir}}}, \tag{3.14}$$

where $r_{\rm vir}$ is the virial radius of the halo.

Combining the projection and RSD effects, the projected CO power spectrum can be written as

$$\begin{aligned} P^{\rm prj}_{\rm CO(1h)}(k_{CII}, \mu_{CII}) &= r^{\parallel}(r^{\perp})^2(1 + \frac{f(z_{CO})}{\bar{b}_{CO}}\mu^2_{CO})^2 \\ &\times \int_{M_{\rm min}}^{M_{\rm max}} dM \frac{dN}{dM}[\left(\frac{L_{\rm line}(M, z_{CO})}{4\pi D^2_L(z_{CO})} y_{\rm CO}(z) D^2_A(z_{CO})\right)^2 \\ &\times |u(k_{CO}|M)|^2 \, e^{-(k_{CO}\sigma_v\mu_{CO})^2/2}], \end{aligned} \tag{3.15}$$

$$\begin{aligned} P^{\rm prj}_{\rm CO(2h)}(k_{CII}, \mu_{CII}) &= r^{\parallel}(r^{\perp})^2 P_{\rm lin}(k_{CO})(1 + \frac{f(z_{CO})}{\bar{b}_{CO}}\mu^2_{CO})^2 \\ &\times \{[\int_{M_{\rm min}}^{M_{\rm max}} dM \frac{dN}{dM}[\left(\frac{L_{\rm line}(M, z_{CO})}{4\pi D^2_L(z_{CO})} y_{CO}(z_{CO}) D^2_A(z_{CO})\right) \\ &\times |u(k_{CO}|M)| \, b(M, z_{CO}) e^{-(k_{CO}\sigma_v\mu_{CO})^2/2}]\}^2. \end{aligned} \tag{3.16}$$

We define $\mu_{\rm CII}$ and $\mu_{\rm CO}$ to be the cosine angle of the $\mathbf{k}_{\rm CII}$ and $\mathbf{k}_{\rm CO}$ vectors with respect to the line-of-sight direction, respectively. For the [C II] power spectrum in its own comoving frame, only the RSD effect needs to be considered, so the [C II] 2D power spectrum has the form given by Eq. (3.12) and Eq. (3.13).

To demonstrate the deblending technique, we consider a simple case where the $z_{\rm CII} = 6$ [C II] line is blended with the brightest CO line, CO(3-2) from $z = 0.27$, in an IM observing volume. This is our fiducial model. Fig. 3.2 shows the fiducial [C II] and projected CO power spectra in different $\mu_{\rm CII}$ values. We also show power spectra obtained by averaging over $\mu_{\rm CII}$, which we called the "ave-prj" power spectrum hereafter. For comparison, we plot the ave-prj power spectra using the SFR-M relation in Silva et al. [52, hereafter S15 model]. The ave-prj CO power spectra for other J transitions are shown in Fig. 3.3.

3.3 MCMC-Based Parameter Inference

We construct the [C II] and projected CO power spectra with the CIB model described above, and use them as templates to extract [C II] and CO components in a simulated intensity mapping data cube. Specifically, we use the MCMC formalism to extract power spectrum parameters.

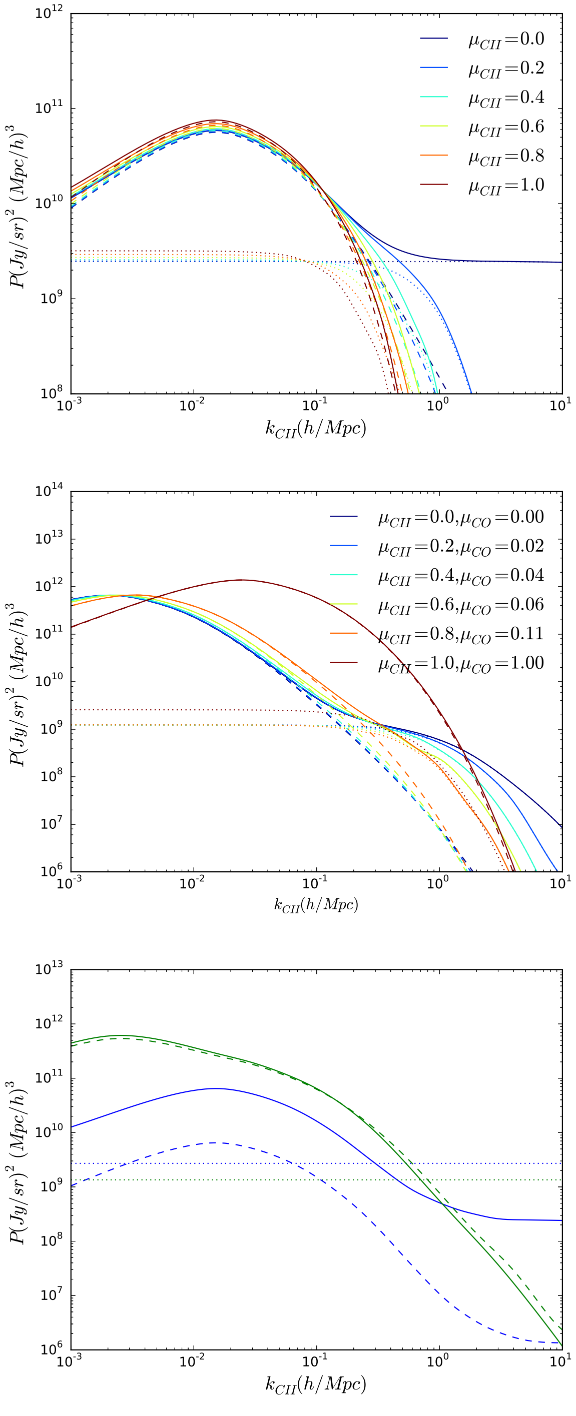

Figure 3.2: Top: projected [C II] power spectra with different μ_{CII} values in the fiducial model. The dotted lines are one-halo terms, the dash lines are two-halo terms, and the solid lines are the projected total power spectrum ($P^{\mathrm{prj}}_{\mathrm{1h}}+P^{\mathrm{prj}}_{\mathrm{2h}}$). Middle: projected CO power spectra for different μ_{CII} and the corresponding μ_{CO}. Bottom: the “ave-prj” power spectrum of [C II] (blue) and CO (green) in the fiducial model (solid lines). The dotted lines are the projected shot noise level of the CIB model for [C II] (blue) and CO (green), respectively. The ave-prj power spectra from S15 model are shown in dash lines.

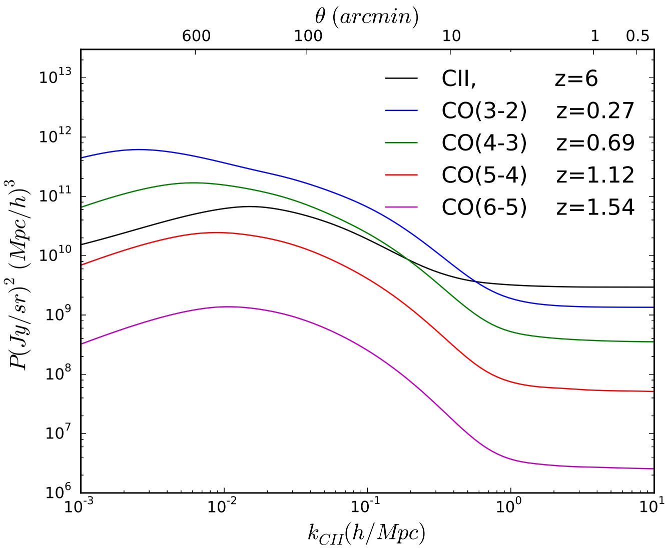

Figure 3.3: The "ave-prj" power spectra of [C II] and several CO contaminants in the fiducial model. The upper axis is the corresponding observed angular scale given by $\theta = 2\pi / k_{CII}\chi(z=6)$.

[C II] and CO Templates

In summary, the [C II] and CO power spectra are derived based on Eqns. (3.12), (3.13), (3.15), and (3.16) ablove. We write the 3D power spectrum in the following form.

$$P_{line}(k_{CII}, \mu_{CII}) = (1 + \frac{f(z_{\rm line})}{\bar{b}_{\rm line}}\mu_{line}^2)^2 \\ \times A_{\rm line} T_{\rm line}(k_{CII}, \mu_{CII}), \tag{3.17}$$

where $A_{\rm line}$ is a constant amplitude of the power spectrum, and $\bar{b}_{\rm line}$ the luminosity-weighted bias of the emission line. $T(k^{\perp}, k^{\parallel})$ is the power spectrum template, given by the sum of the one-halo and two-halo terms in Eq. (3.12), (3.13), (3.15), and (3.16) (without the Kaiser RSD term). Fig. 3.4 shows the [C II] and CO power spectrum templates T for the fiducial model.

We have assumed that the shape of the power spectrum templates T is independent of the luminosity–halo mass relation, so that only an amplitude parameter A is required to describe variations in the model. This assumption is valid on large scales, where the bias parameter is assumed to be scale-independent and only the linear Kaiser RSD effect is important. The large-scale power spectrum is proportional to $\bar{b}^2\bar{I}^2$, where $\bar{b}$ and $\bar{I}$ are the luminosity-weighted bias and mean intensity, respectively, and reduces to the linear form: $P(k,\mu) = (1 + \frac{f(z)}{\bar{b}}\mu^2)^2\bar{b}^2\bar{I}^2 P_{lin}(k)$. Thus the power spectrum can be easily described by $\bar{b}$ in the Kaiser term (where we assume f is

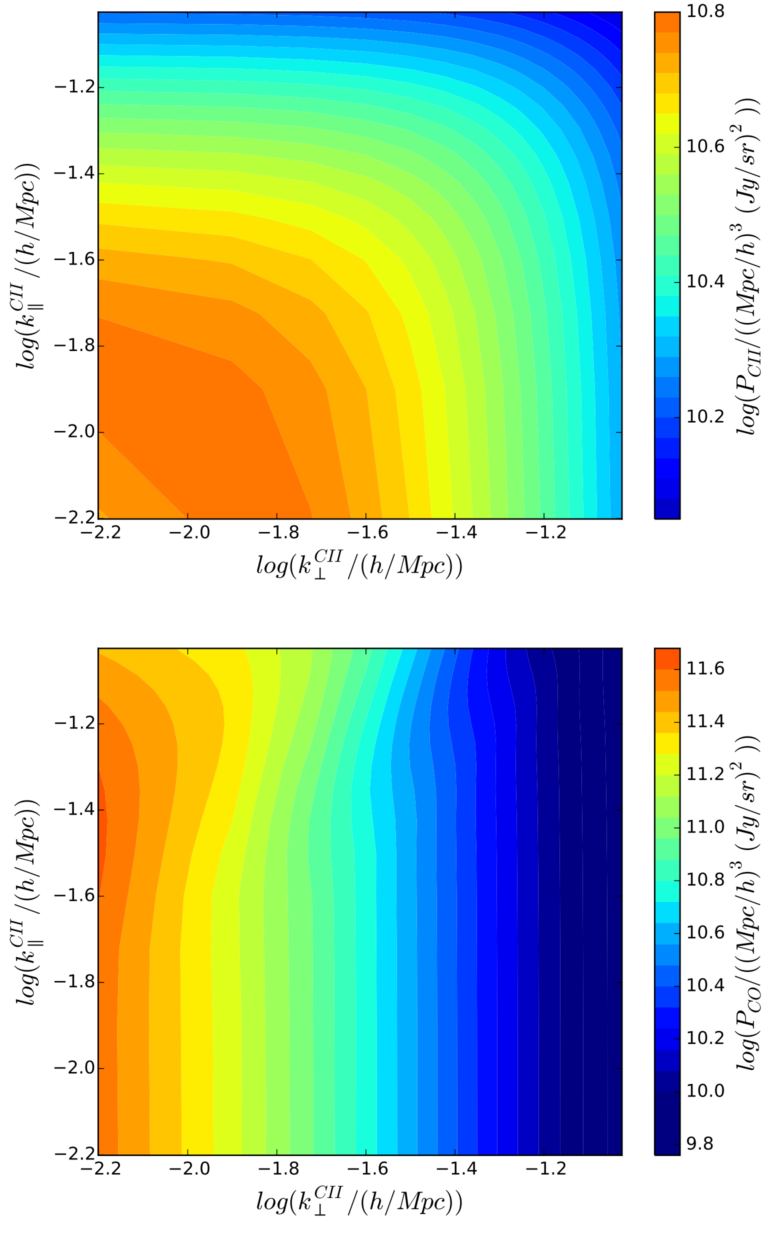

Figure 3.4: [C II] (top) and CO (bottom) power spectrum templates with fiducial model in the k range specified in Sec. 3.3.

fixed by our chosen cosmology) and an overall amplitude A that is proportional to $\bar{b}^2\bar{I}^2$. These are the two parameters we fit with MCMC. For this purpose, we restrict ourselves to large scales only. The exact k-space range we use is described in Sec. 3.3. Below we work interchangeably in the (k , μ) and ($k^{\perp}$, $k^{\parallel}$) space, where $k^{\parallel}$ is the Fourier wavenumber parallel to the light of sight and $k^{\perp}$ perpendicular to it, and $k = \sqrt{(k^{\perp})^2 + (k^{\parallel})^2}$ and $\mu = k^{\parallel}/k$.

k-space Range

We consider a cubic survey size with a spatial dimension of 10 × 10 deg^2. At $z = 6$, this corresponds to a linear comoving scale of 952 Mpc/h; assuming the same comoving scale in the light-of-sight direction, we have an intensity mapping survey volume of $(952)^3(Mpc/h)^3$. We restrict our analysis to large scales only, with $k_{CII} < 0.1$ h Mpc^{-1}. For [C II] at $z = 6$, the smallest accessible $k^{\perp}$ mode is $k_{CII}^{\text{min}} = 6.3 \times 10^{-3}\ h/Mpc$, and we choose the maximum to be $k_{CII}^{\text{max}} = 9.5 \times 10^{-2}\ h/Mpc$. We consider the same limit for $k^{\parallel}$. Thus we have 15×15 k space pixels in the mock data and the templates.

Mock Observed Power Spectrum Construction

We generate a mock data power spectrum P_{data}, which consists of the redshifted [C II] and CO power spectra, and a noise contribution δP:

$$\begin{aligned} P_{data}(k_{CII}, \mu_{CII}) &= P_{CII}(k_{CII}, \mu_{CII}) \\ &+ P_{CO}(k_{CII}, \mu_{CII}) + \delta P(k_{CII}, \mu_{CII}), \end{aligned} \tag{3.18}$$

where P_{CII} and P_{CO} are described in Eq. (3.17).

δP are random values that account for power spectrum noise. For each k-space pixel, δP is drawn from a Gaussian distribution with zero mean and variance $\sigma_P^2(k_{CII}, \mu_{CII})$:

$$\begin{aligned} \sigma_P(k_{CII}, \mu_{CII}) &= \frac{1}{\sqrt{N(k_{CII}, \mu_{CII})}} \\ &\times (P_{\text{data}}(k_{CII}, k_{\parallel}) + P_{\text{sh}}^{CII} + P_{\text{sh}}^{\text{CO}} + P_n), \end{aligned} \tag{3.19}$$

where the P_{data} term accounts for cosmic variance, P_{sh}^{CII} and P_{sh}^{CO} are the shot noise contributions (see Eq. 3.3), and P_n is the instrument noise power spectrum. In this work, we assume $P_n = 1.77 \times 10^9$ (Jy/sr)^2(Mpc/h)^3, which is consistent with the thermal noise level of current generation of planned [C II] IM experiments, such as the TIME-*pilot* [13]; though, TIME-*pilot* plans a smaller survey volume than considered here. $N(k_{CII}, \mu_{CII})$ is the number of pixels in each k bin.

We do not include the shot noise and instrument white noise contributions in the template and in the mock data, since we assume they can be measured and subtracted before the template fitting process. In a real experiment, these constant noises are the dominant "signals" at high-k, so we are able to infer the noise level from the high k modes and subtract them from the data.

MCMC Implementation

For a given set of parameters $\{A_{CII}, A_{CO}, b_{CII}, b_{CO}\}$, the model power spectrum P_{model} is given by

$$\begin{aligned} &P_{\text{model}}(k_{CII}, \mu_{CII}) \\ =&(1 + \frac{f(z_{CII})}{\bar{b}_{CII}}\mu^2_{CII})^2 A_{CII} T_{CII}(k_{CII}, \mu_{CII}) \\ +&(1 + \frac{f(z_{CO})}{\bar{b}_{CO}}\mu^2_{CO})^2 A_{CO} T_{CO}(k_{CII}, \mu_{CII}). \end{aligned} \tag{3.20}$$

The log-likelihood expression is

$$\begin{aligned} ln\ \mathcal{L} = -\frac{1}{2}\sum_k \{ &\frac{(P_{\text{data}}(k_{CII}, \mu_{CII}) - P_{\text{model}}(k_{CII}, \mu_{CII}))^2}{(\sigma_P(k_{CII}, \mu_{CII}))^2} \\ &+ ln\left[2\pi(\sigma_P(k_{CII}, \mu_{CII}))^2\right]\}. \end{aligned} \tag{3.21}$$

We set flat priors for A_{CII} and A_{CO} in the range of $\left[10^{-6}, 10^6\right]$ (fiducial value =1), and flat priors for $\bar{b}_{CII}$ and $\bar{b}_{CO}$ between $[0.1, 20]$, and zero otherwise.

We use Python package *emcee* v2.1.0 [17] to perform the MCMC analysis. We use an ensemble of 1000 walkers taking 1000 steps after 1000 burn-in steps.

3.4 Results

Fiducial Model

Fig. 3.5 shows the posterior distribution of the four parameters in the fiducial case, where $\left\{A_{CII},\ A_{CO},\ \bar{b}_{CII},\ \bar{b}_{CO}\right\} = \{1,\ 1,\ 7.20,\ 1.48\}$. For each point in the four-dimensional parameter space, we construct an ave-prj power spectrum by averaging over μ_{CII} of a 3D power spectrum specified by these parameters. Instead of examining the amplitude parameters A_{CII} and A_{CO}, we use the ave-prj power spectra to compare the input and output amplitudes.

Next, we consider more general cases by changing the input amplitudes A_{CII} and A_{CO} in the mock data. The [C II] and CO shot noise levels also vary accordingly with the clustering amplitudes. We first fix $A_{CII} = 1$ and run MCMC with $A_{CO} = [0.01, 0.1, 1, 10, 100]$. Fig. 3.6 (left) shows the 68% confidence interval of ave-prj [C II] power spectrum amplitude relative to the input, $P_{CII}^{\rm true}$. For comparison, we also calculate the amplitude of the noise power spectrum σ_P (see Eq. 3.19) relative to the best-fit value of the ave-prj [C II] power spectrum $P_{CII}^{\rm best}$. We define a quantity A_σ to be the median value of the ratio $\sigma_P/P_{CII}^{\rm best}$ over the 15×15 k-space pixels. A_σ serves as an indicator of the available information content level, which we discuss further in Sec. 3.5. The $S/N \equiv P_{CII}/\Delta P_{CII}$ of [C II] ave-prj power spectrum is shown in Fig. 3.6 (right), where ΔP_{CII} is the standard deviation of the ave-prj power spectrum amplitudes given by MCMC. We then fix $A_{CO} = 1$ and set $A_{CII} = [0.01, 0.1, 1, 10, 100]$ to repeat the exercise. The results are shown in Fig. 3.7.

In the three cases with high [C II] to CO ratios, where $A_\sigma < 1$, the mean of the ave-prj [C II] power spectrum can be reproduced within 10% deviation from the true value, with $S/N > 10$ in both tests. Not all of the true ave-prj [C II] power spectrum amplitudes, however, fall in the 68% interval of the MCMC distributions. To understand this inconsistency, we run 100 realizations of mock data with the fiducial case ($A_{CII} = A_{CO} = 1$). We find that in 63 of the 100 runs, the ave-prj $P_{CII}^{\rm true}$ does fall in the 68% interval of the MCMC distribution. This suggests that parameter degeneracy may exist in the fitting, so that some of the fitted amplitudes deviate slightly from the true values. This issue might be resolved by adopting tighter constraints on the input $\bar{b}_{CII}$ and $\bar{b}_{CO}$. We will investigate this degeneracy in future work.

For the rest, where $A_\sigma > 1$, the S/N given by MCMC degrade to less than 10 and one of the mean values is biased by a factor of four; we find A_σ to be a good indicator in determining the reliability of the fitting.

Fig. 3.8 shows the ave-prj power spectrum for all nine combinations of input A_{CII} and A_{CO} we discuss above. Reproducing the best-fit mock data with the four parameters $\{A_{CII},\ A_{CO},\ \bar{b}_{CII},\ \bar{b}_{CO}\}$ and comparing with the input mock data, we calculate the χ^2 in each cases, and the probability to exceed (PTE) the observed value if MCMC gives correct parameters. The PTE values are also shown in Fig. 3.8.

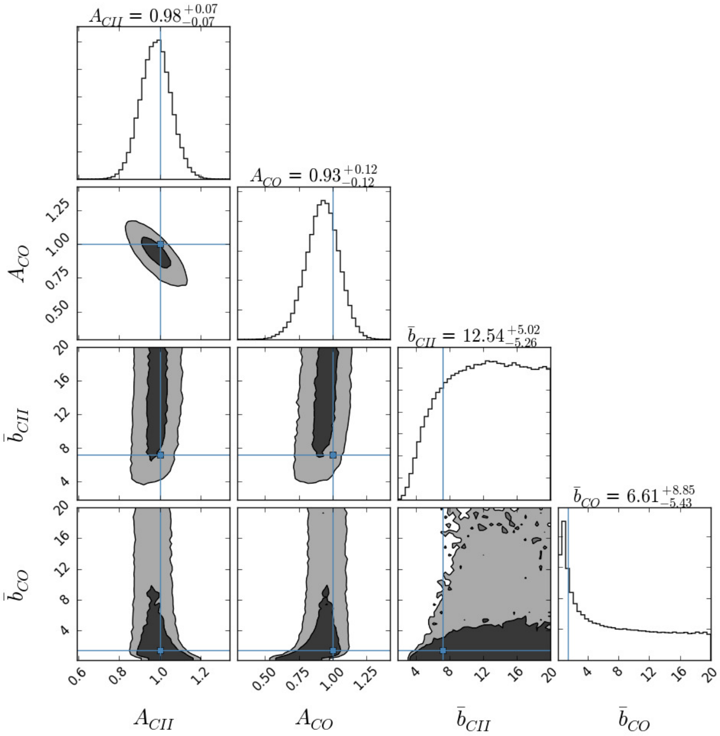

Figure 3.5: MCMC posterior distribution of the fiducial case. The contours marked the 68% and 95% confidence interval in the parameter space. Crosshairs indicate the input value of the data: $\{A_{CII} = 1,\ A_{CO} = 1,\ \bar{b}_{CII} = 7.20,\ \bar{b}_{CO} = 1.48\}$.

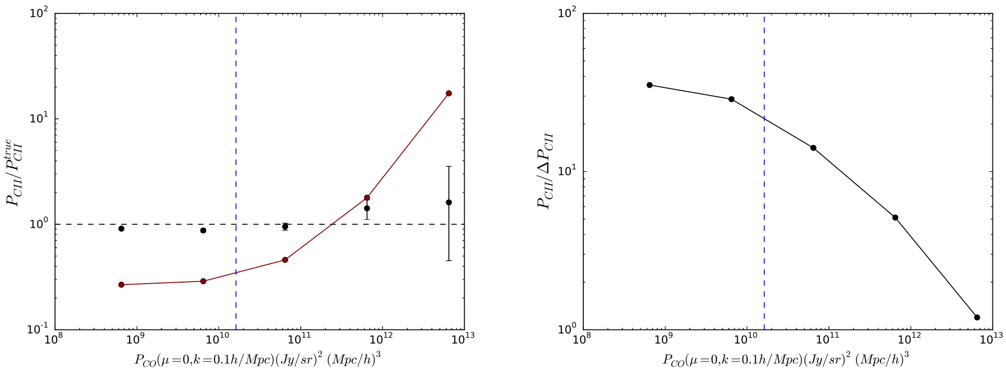

Figure 3.6: Left: the ave-prj power spectrum amplitude from MCMC relative to the input value in the mock data. The inputs are $A_{CII} = 1$, and $A_{CO} = [0.01, 0.1, 1, 10, 100]$ from left to right. The x-axis is expressed in the CO ave-prj power spectrum at $k_{CII} = 0.1\ h$ Mpc^{-1} for better comparison. The blue dashed line indicates the input [C II] prj-ave power spectrum at $k_{CII} = 0.1\ h$ Mpc^{-1}. The error bars are the 68% confident interval given by MCMC. The dark red line indicates the A_{σ} values (see the text). Right: S/N of the ave-prj [C II] power spectrum amplitude given by MCMC.

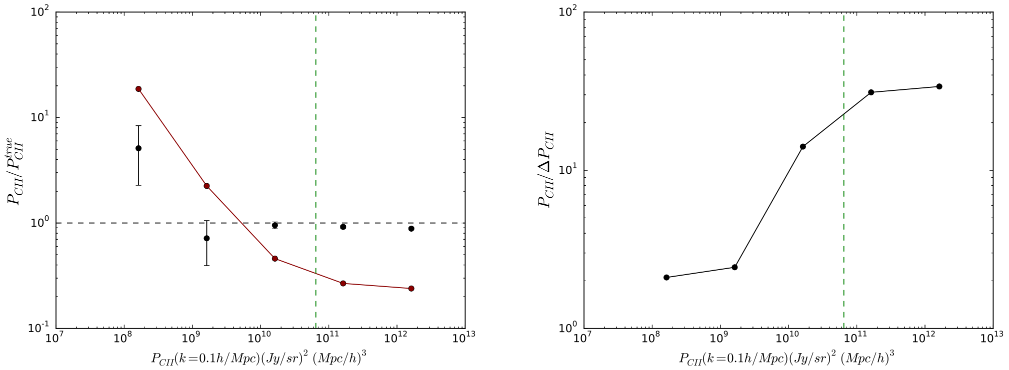

Figure 3.7: Left: the ave-prj power spectrum amplitude from MCMC relative to the input value in the mock data. The inputs are $A_{CO} = 1$, and $A_{CII} = [0.01, 0.1, 1, 10, 100]$ from left to right. The x-axis is expressed in the [C II] ave-prj power spectrum at $k_{CII} = 0.1$ h Mpc^{-1} for better comparison. The green dashed line indicates the input CO prj-ave power spectrum at $k_{CII} = 0.1$ h Mpc^{-1}. The error bars are the 68% confident interval given by MCMC. The dark red line indicates the A_σ values (see text). Right: S/N of the ave-prj [C II] power spectrum amplitude given by MCMC.

S/N Dependence on Noise Level

We assume the instrument noise can be subtracted from the data power spectrum before template fitting, but it still contributes a k-space noise in Eq. (3.19). Here we investigate the effect of instrument noise level P_n on the fitted results. We again fix input $A_{CII} = 1$ and $A_{CO} = [0.01, 0.1, 1, 10, 100]$, and run a series of cases by changing P_n to be $[0.1, 1, 10, 100, 1000]$ times the initial P_n value ($P_{n,i} = 1.77 \times 10^9\ (Jy/sr)^2(Mpc/h)^3$). The results are shown in Fig. 3.9. For comparison, We also calculate the theoretical S/N on the [C II] power spectrum, which can be expressed as

$$S/N = \sqrt{\sum_k \left(\frac{P_{CII}(k_{CII}, \mu_{CII})}{\sigma_P(k_{CII}, \mu_{CII})}\right)^2}, \tag{3.22}$$

where σ_P is given by Eq. (3.19), and summing over all k-space pixels.

Note that all of the highest P_n cases show biased results, and some of the large P_n cases also give amplitudes that deviate significantly from the input values. This may be indicative that given the survey size we consider in this work, $P_n \approx 10^{10}(Jy/sr)^2(Mpc/h)^3$ is the maximum allowed instrument noise for detecting the CIB-based [C II] signals. We will conduct a more detailed investigation in future work.

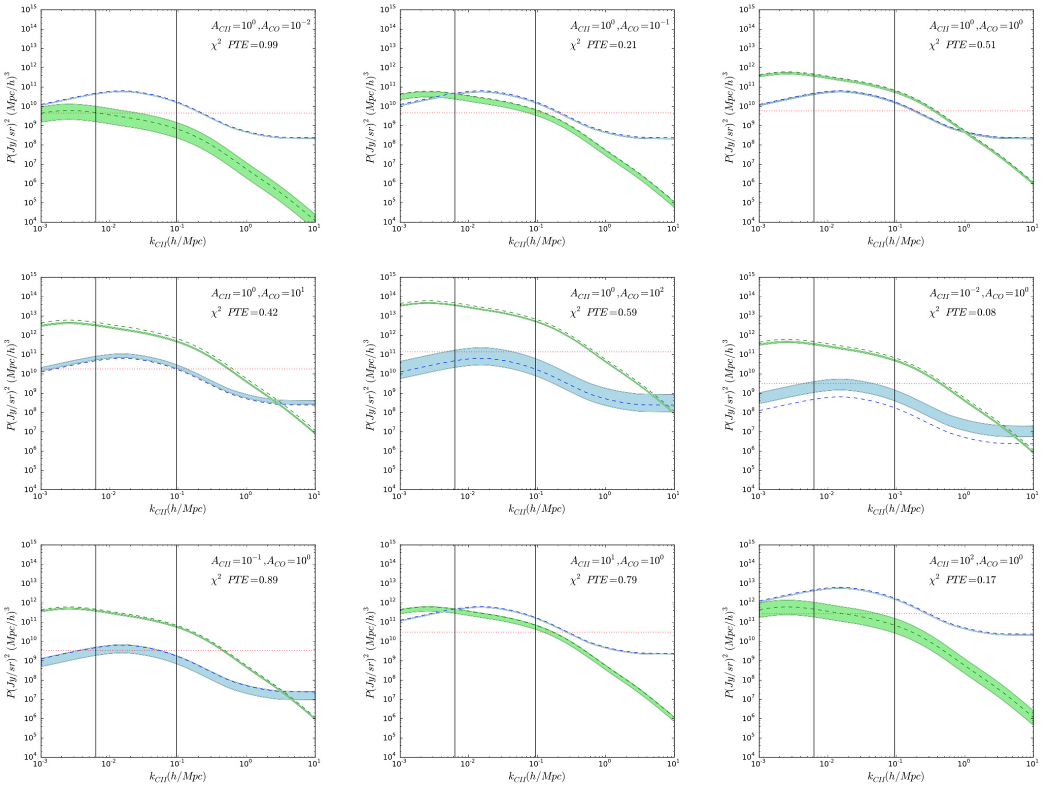

Figure 3.8: The ave-prj power spectrum with different combinations of input A_{CII} and A_{CO}. The blue and green regions are the 68% confidence interval of [C II] and CO ave-prj power spectrum respectively. The dashed lines are the input [C II](blue) and CO(green) ave-prj power spectrum of the mock data. The red dotted line indicates $P^{CII}_{sh} + P^{CO}_{sh} + P_n$, which is the constant power spectrum in the data that contribute to the noise in k space (see Eq. 3.19). The two vertical black lines are k^{min}_{CII} and k^{max}_{CII} marking the k space region we use in template fitting. The χ^2 PTE for each MCMC fit is also provided.

Multiple Foreground Lines

So far, we have only considered the brightest foreground CO (3-2) line at $z = 0.27$ to the $z = 6$ [C II] signal. Here we extend the technique to incorporate two more foreground lines: CO(4-3) from $z = 0.69$ and CO(5-4) from $z = 1.12$. The power spectra of $J > 5$ transition lines are more than two orders of magnitude lower than the expected [C II] signal, and we do not consider them in this paper.

The extra CO lines can be incorporated by extending Eq. (3.18) and Eq. (3.20) with two more CO terms, which introduce two more amplitudes and bias factors in the MCMC procedure. Thus we now fit the mock data with eight parameters in the same k-space defined before. We set all the input amplitude to be unity, so the mock

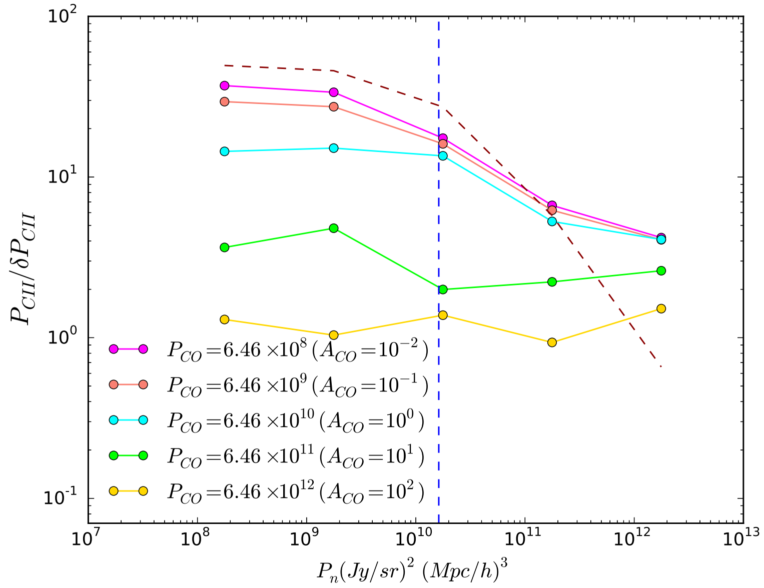

Figure 3.9: The S/N of the ave-prj [C II] power spectrum amplitude as a function of P_n and P_{CO}. The P_{CO} written on the top right corner is the ave-prj power spectrum at $k = 0.1\ h\,\mathrm{Mpc}^{-1}$. The blue dashed line marks the [C II] ave-prj power spectrum at $k = 0.1\ h\,\mathrm{Mpc}^{-1}$. The dark red dashed line is the S/N derived from mode counting (see the text).

data is consistent with the CIB model prediction.

The result is shown in Fig. 3.10. The [C II] ave-prj amplitude given by MCMC has an $S/N = 4.12$, and the input [C II] ave-prj amplitude slightly falls outside the 68% confidence interval. The technique appears to be valid even in the presence of multiple foreground lines that overwhelm the [C II] signals in the k range we consider, though at the cost of S/N for the extracted [C II] signal.

3.5 Discussion

Model Dependence

In reality, if the templates do not perfectly describe the true signal intensity field, there will be amplitude and shape discrepancies between the template and true signal power spectra. The overall amplitude discrepancy can be fully absorbed by the amplitude parameters A_{CII} and A_{CO} in our fitting process. Power spectrum shape difference may arise from different assumed L–M relations, but since we restrict our fitting to the large-scale clustering terms, the procedure is not susceptible to incorrect model assumptions. As a sanity check, we use templates generated from

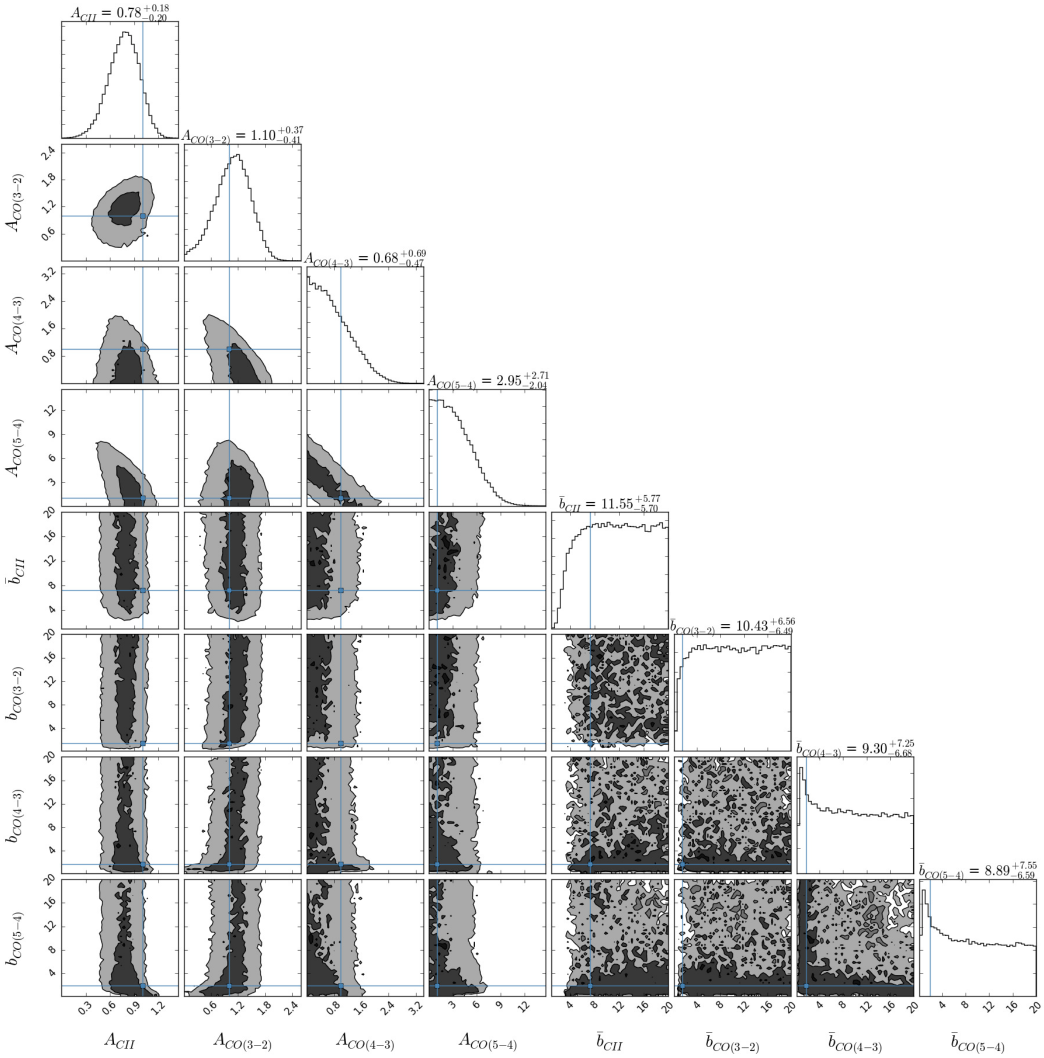

Figure 3.10: MCMC posterior distribution on the parameter space with three CO foreground lines. The contours marked the 68% and 95% confidence interval in the parameter space. Crosshairs indicate the input value of the data:
$\{\ A_{CII} = 1,\ A_{CO(3-2)} = 1,\ A_{CO(4-3)} = 1,\ A_{CO(5-4)} = 1,$
$\bar{b}_{CII} = 7.20,\ \bar{b}_{CO(3-2)} = 1.48,\ \bar{b}_{CO(4-3)} = 1.70,\ \bar{b}_{CO(5-4)} = 1.94\ \}$.

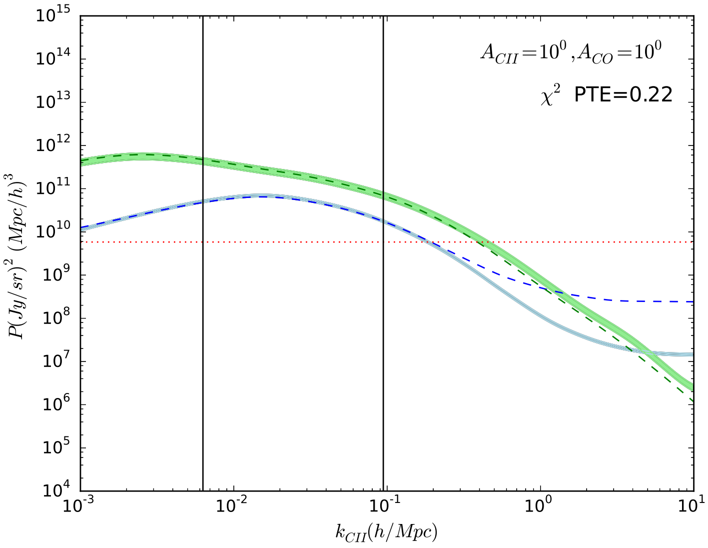

Figure 3.11: The ave-prj power spectrum of fiducial CIB model fitted with S15 template. The blue and green regions are the 68% confidence interval of [C II] and CO ave-prj power spectrum respectively. The dashed lines are the input [C II](blue) and CO(green) ave-prj power spectrum of the mock data. The red dotted line indicates $P_{\rm sh}^{CII} + P_{\rm sh}^{CO} + P_n$, which is the constant power spectrum in the data that contributes to the noise in k space (see Eq. 3.19). The two vertical black lines are $k_{CII}^{\rm min}$ and $k_{CII}^{\rm max}$ marking the k space region we use in template fitting. The χ^2 PTE for each MCMC fit is also provided.

the S15 model to fit the same set of mock data discussed in Sec. 3.4, which are constructed with the CIB model. Fig. 3.11 shows the ave-prj power spectrum from fitting the S15 template to the CIB model mock data. The results for all the nine scenarios considered before are also consistent with the results presented in Sec. 3.4. The template fitting technique is robust against model uncertainties as long as only large-scale information is considered.

Conversely, this model-independent property implies that the technique constrains only the overall amplitude of the power spectrum and is not sensitive to different L–M relation scenarios.

Model Uncertainties

While [C II] and CO modeling uncertainties do not bias the template fitting results, they affect the quality of the fit. In Sec. 3.2, we show that the CIB model considered in this work gives rise to a high [C II] power and a large bias factor. Our modeling of the different CO J rotational lines also determines the relative amplitudes of

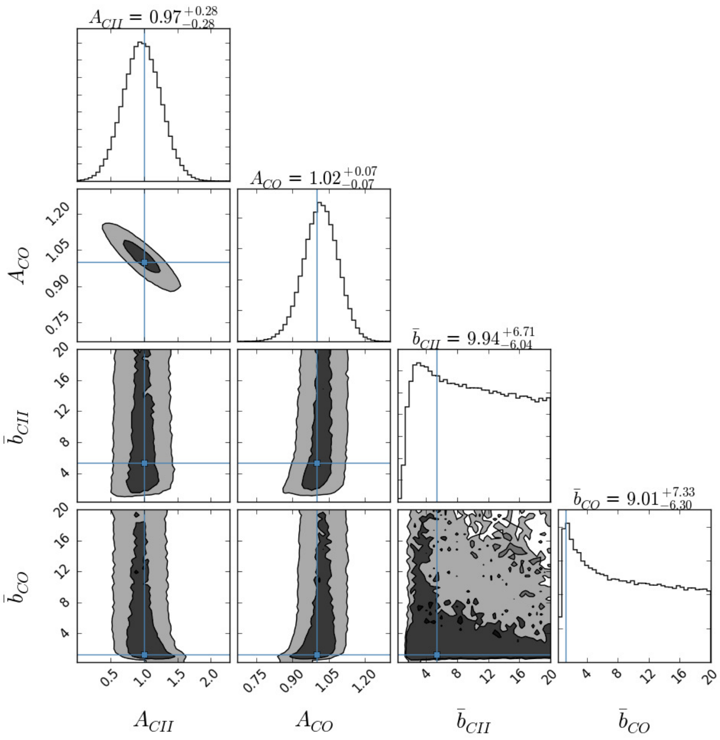

Figure 3.12: MCMC posterior distribution of the S15 case. The contours marked the 68% and 95% confidence interval in the parameter space. Crosshairs indicate the input value of the data: $\left\{A_{CII} = 1,\ A_{CO} = 1,\ \bar{b}_{CII} = 5.38,\ \bar{b}_{CO} = 1.29\right\}$.

interlopers and the severity of contamination.

We have therefore conducted a series of tests with different input [C II] and CO amplitudes and varying noise levels to account for the uncertainties. We find that if the true [C II] power is 10 times smaller than the fiducial CIB amplitude (second-left point in Fig 3.7), for instance, A_σ goes above unity, which implies a non-detection, contrary to the more optimistic fiducial case.

To get a sense of how much the results vary with the assumed models, we run simulations with mock data and templates generated from the S15 model. Compared to the CIB model, the S15 model has slightly lower bias values for both [C II] and CO lines and a similar CO power spectrum amplitude, while the [C II] power spectrum amplitdue is lower by about one order of magnitude (see the bottom panel of Fig. 3.2). Fig. 3.12 shows the MCMC result of the S15 model. The S/N on ave-prj [C II] power spectrum in this case is 4.5. For comparison, the fiducial CIB model has an S/N of ≈ 14. We find that the fitted [C II] power spectrum S/N depends sensitively on its amplitude, when overwhelmed by the CO foregrounds.

In a similar work by Lidz & Taylor (2016, in prep), the authors modeled the power spectrum with an SFR that follows the Schechter functional form [49], and the [C II] power spectrum they derived at $z = 7$ was about one order of magnitude smaller than our fiducial $z = 6$ prediction. Their assumed survey volume was about 5.6 times smaller than considered here. As a result, the [C II] amplitude constraints in their study using the Fisher matrix formalism appeared less optimistic than ours, but the two results are broadly in agreement. In reality, the feasibility of this deblending method will be highly dependent on the assumed survey geometry and S/N strengths. We will leave a more realistic [C II] and CO power spectrum modeling to future work, while cautioning that built-in detection margins are important for planned experiments.

Continuum Foreground

In this paper, we focus on the de-confusion technique that handles spectral line foregrounds. For completeness, we note that at the frequency range of interest, $\sim 200-300$ GHz, continuum foregrounds are non-negligible and generally stronger than line foregrounds. For our purpose, the main continuum foregrounds include the cosmic microwave background and CIB radiations; the two contribute comparably. Silva et al. [52] estimates that the dust continuum emission is of the order of 10^5 Jy sr^{-1}, which is two or three orders of magnitude (depending on model) higher than the [C II] intensity considered in this work. However, since the continuum signals are expected to be spectrally smooth, they dominate the low $k^{\parallel}$ modes in power spectrum space. We therefore expect to be able to mitigate the effect by removing or avoiding the one or two lowest $k^{\parallel}$ modes before template fitting. This is the same technique envoked in the well-studied field of 21-cm LIM (e.g., Liu & Tegmark [31], Parsons et al. [41], Switzer et al. [54]) and implemented on LIM data (e.g., Ali et al. [1], Keating et al. [24], Paciga et al. [40], Switzer et al. [55]). To test the impact of continuum foreground mode removal, we run a fiducial MCMC fit, same as the case in Sec. 3.4 but removing the lowest $k^{\parallel}$ modes. The results are shown in Fig. 3.13. The S/N on the [C II] power spectrum remains nearly the same as the fiducial one (Fig. 3.5). This simple test demonstrates that continuum foregrounds are unlikely to be a major concern, but we note that subtler issues, such as the exact number of $k^{\parallel}$ modes to be removed and the amount of residual continuum in the data, need to be quantified and further tested in the future.

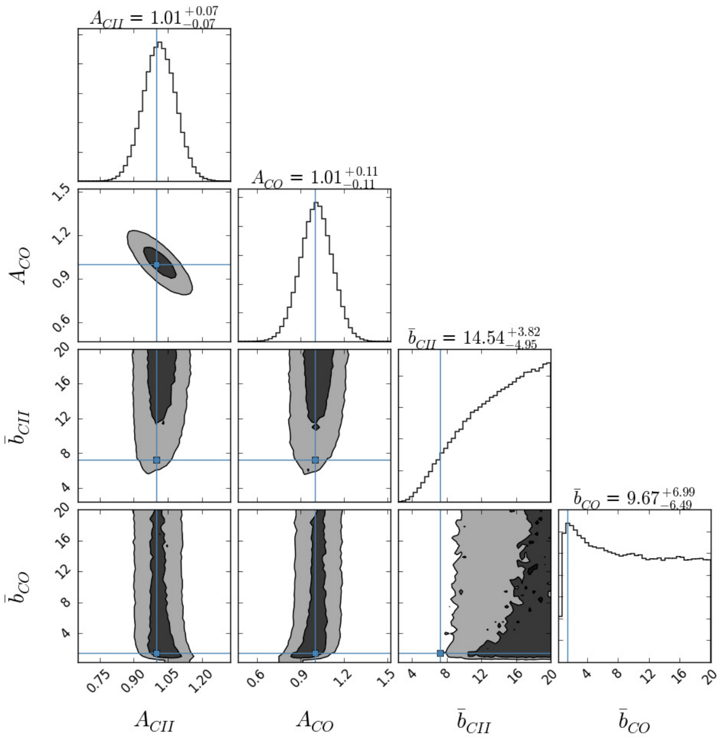

Figure 3.13: MCMC posterior distribution of the fiducial model with the lowest $k^{\parallel}$ mode removed.

Detection Limit

In Sec. 3.4, we define A_σ to be the median value of σ_P/P^{best}_{CII} and use it as an indicator of the available information content level; when $A_\sigma < 1$, or $\sigma_P < P^{\text{best}}_{CII}$, we recover an unbiased estimate of P^{true}_{CII}. Indeed, in Sec. 3.4, the scenarios with high CO to [C II] amplitudes result in biased extracted [C II] power spectrum amplitudes, however, the extracted amplitudes are all below the noise level (i.e. $A_\sigma < 1$). This suggests that the k-space noise level is a good indicator of the information content and sets the limit for extracting [C II] signals from the data, below which the MCMC likely returns biased results. In reality, σ_P and P^{best}_{CII} are quantities that can be directly inferred from real data, and serve to evaluate the reliability of the extracted results.

Constraint on Bias

In our template fitting procedure, the bias factors are loosely constrained by MCMC (see Fig. 3.5 and 3.10). This can be understood by looking at the Kaiser RSD term: $(\frac{f(z))}{\bar{b}_{line}}\mu^2)^2$, where $f \approx 1$, the bias is usually between 1 and 10, and μ^2 is a value between 0 and 1. Thus $\frac{f(z))}{\bar{b}_{line}}\mu^2$ is usually smaller than unity, sub-dominant to the overall amplitude change of the power spectrum, which is absorbed in the A

parameter. In addition, the MCMC constraint on $\bar{b}_{CO}$ is weaker than on $\bar{b}_{CO}$. This is due to the projection effect that makes the projected μ_{CO} very small for most of the k-space pixels (see Fig. 3.2), which are then not sensitive to the Kaiser effect.

The CIB model gives a high $\bar{b}_{CII}$ value that makes the Kaiser effect too small to be detected. This can be seen from the histogram of $\bar{b}_{CII}$ given by MCMC, which extends to large values that correspond to a non-detection of the Kaiser RSD effect. To test the ability to extract bias in our procedure, we run the fiducial case and set $\bar{b}_{CII}$ to be the same as the fiducial value of $\bar{b}_{CO} = 1.48$. In this case, the $\bar{b}_{CII}$ can be constrained by MCMC with an $S/N > 5$; while the $\bar{b}_{CO}$ is still unconstrained due to the projection effect described above. Therefore, if $\bar{b}_{CII}$ in reality is smaller than the value we considered in this work, we might be able to better constrain $\bar{b}_{CII}$. We will investigate strategies to better extract the bias information in future work.

3.6 Conclusion

We demonstrate the feasibility of deblending spectral line information in the IM regime. We consider a 3D LIM survey, where multiple spectral lines at different redshifts are embedded in the same observing volume, and make use of the anisotropic shape of their respective power spectra when projected into a common comoving coordinate to disentangle the information. We consider deblending high-redshift [C II] signals from the brighter, lower-redshift CO interlopers. We use the halo-model and CIB measurement to construct expected CO and [C II] templates across redshifts, and use the MCMC formalism to constrain power spectrum parameters. We show that this technique can reproduce the linear [C II] and CO power spectrum amplitudes, though with reduced signal-to-noise, given a range of CO signal strengths and noise levels. We establish an indicator to evaluate whether the fitted parameters are unbiased. Finally, we demonstrate the ability of extracting [C II] in the presence of multiple, stronger CO foreground lines. The technique can be extended to other line blending problems to extract information of both signal and interlopers in an intensity mapping experiment.

.1 Power Spectrum Modeling

$L_{IR} - M$ Relation: CIB Model

In Planck Collaboration et al. [43], the authors modeled the CIB emission as measured by *Planck*, and parametrized the CIB specific luminosity L_ν with the following three components.

1. A normalized spectral energy distribution, $\Theta(\nu, z)$, for all the galaxies:

$$\begin{aligned} \Theta(\nu, z) &\propto \nu^{\beta} B_{\nu}(T_d(z)) && ; \nu < \nu_0 \\ \Theta(\nu, z) &\propto \nu^{-\gamma} && ; \nu \geq \nu_0, \end{aligned} \tag{23}$$

   where B_ν is the Planck function, and T_d the redshift-dependent dust temperature:

$$T_d = T_0(1+z)^{\alpha}. \tag{24}$$

2. L−−M relation: they assumed the CIB luminosity is a log-normal function Σ of halo mass M with peak mass M_{eff} and variance $\sigma^2_{L/M}$,

$$\Sigma(M) = M \frac{1}{(2\pi\sigma^2_{L/M})^{1/2}} e^{-(log_{10}(M) - log_{10}(M_{eff}))^2/2\sigma^2_{L/M}}. \tag{25}$$

3. Redshift evolution of the L−−M relation: the global normalization is parametrized by

$$\Phi(z) = (1+z)^{\delta}. \tag{26}$$

   The L−−M ratio is assumed to increase with redshift (i.e. $\delta > 0$).

Combining these three components, the $L_\nu - M$ relation can be written as

$$L_\nu(M, z) = L_0 \Phi(z) \Sigma(M) \Theta(\nu, z), \tag{27}$$

with an overall normalization factor L_0.

With the *Planck* data, the authors constrained the model parameters as listed in Table 9 of Planck Collaboration et al. [43]. Here we adopt their best-fit values: $\{\alpha = 0.36,\ \beta = 1.75,\ \gamma = 1.7,\ \delta = 3.6,\ T_0 = 24.4\ K,\ M_{\rm eff} = 10^{12.6} M_\odot,\ \sigma^2_{L/M} = 0.5,\ L_0 = 0.02 L_\odot\}$. We integrate Eq. (27) over the wavelength range of $8 - 1000\ \mu m$ to obtain the total infrared luminosity $L_{\rm IR}$.

We can convert $L_{\rm IR}$ to SFR following Kennicutt [25]:

$$SFR/L_{IR} = 1.7 \times 10^{-10} M_\odot yr^{-1} L_\odot^{-1}. \tag{28}$$

As a sanity check, we calculate the resulting star-formation rate density (SFRD) by integrating the SFR over halo mass,

$$SFRD(z) = \int_{M_{min}}^{M_{max}} dM \frac{dN}{dM} SFR(M, z), \tag{29}$$

where dN/dM is the halo mass function [51], and we take $M_{\min} = 10^8\ M_\odot/h$, $M_{\max} = 10^{15}\ M_\odot/h$. We use this integration range for all halo mass integration throughout this work.

For comparison, we calculate the SFRD from the SFR–M relation given by Lidz et al. [28], Pullen et al. [45], and Silva et al. [52]. We also consider the SFRD given by Madau & Dickinson [33] and Robertson et al. [48], where the SFRD is modeled by the following four-parameter functional form

$$SFRD(z) = a\frac{(1+z)^b}{1+[(1+z)/c]^d}\left(M_\odot yr^{-1}Mpc^{-3}\right). \tag{30}$$

Madau & Dickinson [33] used the UV and IR galaxy counts and obtained the parameters $\{a, b, c, d\} = \{0.015, 2.7, 2.9, 5.6\}$; while Robertson et al. [48] used the joint constraint of galaxy counts and the CMB optical depth of $\tau = 0.066 \pm 0.12$ from Planck Collaboration et al. [44] and obtained $\{a, b, c, d\} = \{0.01376, 3.26, 2.59, 5.68\}$.

The SFRD(z) of the aforementioned models are plotted in Fig. .14. Lidz et al. [28] and Pullen et al. [45] use a simple scaling relation to model the SFR–M relation, and do not reproduce the SFRD peak at $z \sim 2-3$. In Silva et al. [52], the SFR–M relation is fitted in several redshift bins with mock galaxy catalogs from De Lucia & Blaizot [15] and Guo et al. [22]. The discontinuous features in the SFRD curve are the boundaries of the redshift bins. They predict an SFRD peak at a higher redshift and the model does not agree well with galaxy counts at $z \sim 0-2$. The SFRD from the CIB model is broadly consistent in shape with that of galaxy counts, but the CIB model predicts a systematically higher amplitude, especially at high redshifts. This results in a higher [C II] power spectrum amplitude. Furthermore, the steep SFR–M relation in the CIB model also results in a high bias factor for [C II]. We discuss the implications in Sec. 3.5.

$L_{CII} - SFR$ and $L_{CO} - SFR$ Relation

To connect L_{CII} to SFR, we adopt the following relation based on observations of nearby late-type galaxies [2]:

$$L_{CII}(M, z) = 1.59 \times 10^{-3} L_{IR}, \tag{31}$$

and use Eq. (28) to convert $L_{\rm IR}$ to SFR.

For CO luminosity, we use the empirical relation from Wang et al. [58]:

$$L_{CO(1-0)} = 3.2 \times 10^4 L_\odot\left(\frac{SFR}{M_\odot yr^{-1}}\right)^{3/5}. \tag{32}$$

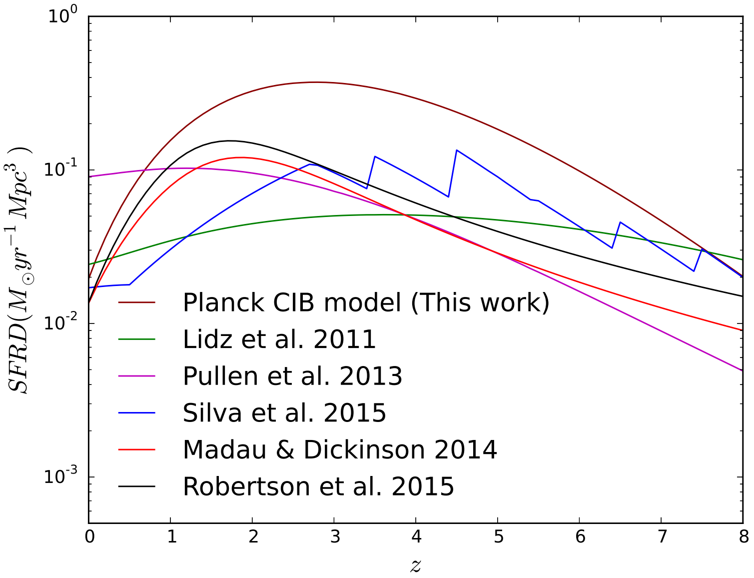

Figure .14: The SFRD of the *Planck* CIB model (this work) comparing with the literature [28, 33, 45, 48, 52].

For higher-J rotational transitions, we assume a constant ratio between $L_{CO(1-0)}$ and $L_{CO(J-(J-1))}$, using Eq. (16) of Obreschkow et al. [39] and assuming an excitation temperature of $T_e = 17$ K for all the galaxies.

References

[1] Ali, Z. S., Parsons, A. R., Zheng, H., et al. 2015, ApJ, 809, 61, doi: `10.1088/0004-637X/809/1/61`

[2] Boselli, A., Gavazzi, G., Lequeux, J., & Pierini, D. 2002, A&A, 385, 454, doi: `10.1051/0004-6361:20020156`

[3] Bowman, J. D., Morales, M. F., & Hewitt, J. N. 2009, ApJ, 695, 183, doi: `10.1088/0004-637X/695/1/183`

[4] Breysse, P. C., Kovetz, E. D., & Kamionkowski, M. 2014, MNRAS, 443, 3506, doi: `10.1093/mnras/stu1312`

[5] —. 2015, MNRAS, 452, 3408, doi: `10.1093/mnras/stv1476`

[6] Carilli, C. L. 2011, ApJL, 730, L30, doi: `10.1088/2041-8205/730/2/L30`

[7] Chang, T. C., Gong, Y., Santos, M., et al. 2015, Advancing Astrophysics with the Square Kilometre Array (AASKA14), 4. `https://arxiv.org/abs/1501.04654`

[8] Chang, T.-C., Pen, U.-L., Bandura, K., & Peterson, J. B. 2010, Nature, 466, 463, doi: 10.1038/nature09187

[9] Chang, T.-C., Pen, U.-L., Peterson, J. B., & McDonald, P. 2008, Physical Review Letters, 100, 091303, doi: 10.1103/PhysRevLett.100.091303

[10] Chapman, E., Abdalla, F. B., Harker, G., et al. 2012, MNRAS, 423, 2518, doi: 10.1111/j.1365-2966.2012.21065.x

[11] Comaschi, P., & Ferrara, A. 2016, MNRAS, 455, 725, doi: 10.1093/mnras/stv2339

[12] Cooray, A., & Sheth, R. 2002, PhysRep, 372, 1, doi: 10.1016/S0370-1573(02)00276-4

[13] Crites, A. T., Bock, J. J., Bradford, C. M., et al. 2014, in Proc. SPIE, Vol. 9153, Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy VII, 91531W, doi: 10.1117/12.2057207

[14] Croft, R. A. C., Miralda-Escudé, J., Zheng, Z., et al. 2016, MNRAS, 457, 3541, doi: 10.1093/mnras/stw204

[15] De Lucia, G., & Blaizot, J. 2007, MNRAS, 375, 2, doi: 10.1111/j.1365-2966.2006.11287.x

[16] de Putter, R., Holder, G. P., Chang, T.-C., & Dore, O. 2014, ArXiv e-prints. https://arxiv.org/abs/1403.3727

[17] Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306, doi: 10.1086/670067

[18] Furlanetto, S. R., Oh, S. P., & Briggs, F. H. 2006, PhysRep, 433, 181, doi: 10.1016/j.physrep.2006.08.002

[19] Gong, Y., Cooray, A., Silva, M., et al. 2012, ApJ, 745, 49, doi: 10.1088/0004-637X/745/1/49

[20] Gong, Y., Cooray, A., Silva, M. B., Santos, M. G., & Lubin, P. 2011, ApJL, 728, L46, doi: 10.1088/2041-8205/728/2/L46

[21] Gong, Y., Silva, M., Cooray, A., & Santos, M. G. 2014, ApJ, 785, 72, doi: 10.1088/0004-637X/785/1/72

[22] Guo, Q., White, S., Boylan-Kolchin, M., et al. 2011, MNRAS, 413, 101, doi: 10.1111/j.1365-2966.2010.18114.x

[23] Kaiser, N. 1987, MNRAS, 227, 1, doi: 10.1093/mnras/227.1.1

[24] Keating, G. K., Bower, G. C., Marrone, D. P., et al. 2015, ApJ, 814, 140, doi: 10.1088/0004-637X/814/2/140

[25] Kennicutt, Jr., R. C. 1998, ApJ, 498, 541, doi: 10.1086/305588

[26] Kogut, A., Dwek, E., & Moseley, S. H. 2015, ApJ, 806, 234, doi: 10.1088/0004-637X/806/2/234

[27] Li, T. Y., Wechsler, R. H., Devaraj, K., & Church, S. E. 2016, ApJ, 817, 169, doi: 10.3847/0004-637X/817/2/169

[28] Lidz, A., Furlanetto, S. R., Oh, S. P., et al. 2011, ApJ, 741, 70, doi: 10.1088/0004-637X/741/2/70

[29] Lidz, A., & Taylor, J. 2016, ArXiv e-prints. https://arxiv.org/abs/1604.05737

[30] Lidz, A., Zahn, O., Furlanetto, S. R., et al. 2009, ApJ, 690, 252, doi: 10.1088/0004-637X/690/1/252

[31] Liu, A., & Tegmark, M. 2011, PhRvD, 83, 103006, doi: 10.1103/PhysRevD.83.103006

[32] —. 2012, MNRAS, 419, 3491, doi: 10.1111/j.1365-2966.2011.19989.x

[33] Madau, P., & Dickinson, M. 2014, ARA&A, 52, 415, doi: 10.1146/annurev-astro-081811-125615

[34] Madau, P., Meiksin, A., & Rees, M. J. 1997, ApJ, 475, 429

[35] Mashian, N., Sternberg, A., & Loeb, A. 2015, JCAP, 11, 028, doi: 10.1088/1475-7516/2015/11/028

[36] Masui, K. W., Switzer, E. R., Banavar, N., et al. 2013, ApJL, 763, L20, doi: 10.1088/2041-8205/763/1/L20

[37] Morales, M. F., Bowman, J. D., & Hewitt, J. N. 2006, ApJ, 648, 767, doi: 10.1086/506135

[38] Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, ApJ, 462, 563, doi: 10.1086/177173

[39] Obreschkow, D., Heywood, I., Klöckner, H.-R., & Rawlings, S. 2009, ApJ, 702, 1321, doi: 10.1088/0004-637X/702/2/1321

[40] Paciga, G., Albert, J. G., Bandura, K., et al. 2013, MNRAS, 433, 639, doi: 10.1093/mnras/stt753

[41] Parsons, A. R., Pober, J. C., Aguirre, J. E., et al. 2012, ApJ, 756, 165, doi: 10.1088/0004-637X/756/2/165

[42] Peacock, J. A. 1992, in Lecture Notes in Physics, Berlin Springer Verlag, Vol. 408, New Insights into the Universe, ed. V. J. Martinez, M. Portilla, & D. Saez, 1, doi: 10.1007/3-540-55842-X_28

[43] Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2014, A&A, 571, A30, doi: 10.1051/0004-6361/201322093

[44] —. 2015, ArXiv e-prints. https://arxiv.org/abs/1502.01589

[45] Pullen, A. R., Chang, T.-C., Doré, O., & Lidz, A. 2013, ApJ, 768, 15, doi: 10.1088/0004-637X/768/1/15

[46] Pullen, A. R., Doré, O., & Bock, J. 2014, ApJ, 786, 111, doi: 10.1088/0004-637X/786/2/111

[47] Righi, M., Hernández-Monteagudo, C., & Sunyaev, R. A. 2008, A&A, 489, 489, doi: 10.1051/0004-6361:200810199

[48] Robertson, B. E., Ellis, R. S., Furlanetto, S. R., & Dunlop, J. S. 2015, ApJL, 802, L19, doi: 10.1088/2041-8205/802/2/L19

[49] Schechter, P. 1976, ApJ, 203, 297, doi: 10.1086/154079

[50] Scott, D., & Rees, M. J. 1990, MNRAS, 247, 510

[51] Sheth, R. K., & Tormen, G. 1999, MNRAS, 308, 119, doi: 10.1046/j.1365-8711.1999.02692.x

[52] Silva, M., Santos, M. G., Cooray, A., & Gong, Y. 2015, ApJ, 806, 209, doi: 10.1088/0004-637X/806/2/209

[53] Silva, M. B., Santos, M. G., Gong, Y., Cooray, A., & Bock, J. 2013, ApJ, 763, 132, doi: 10.1088/0004-637X/763/2/132

[54] Switzer, E. R., Chang, T.-C., Masui, K. W., Pen, U.-L., & Voytek, T. C. 2015, ApJ, 815, 51, doi: 10.1088/0004-637X/815/1/51

[55] Switzer, E. R., Masui, K. W., Bandura, K., et al. 2013, MNRAS, 434, L46, doi: 10.1093/mnrasl/slt074

[56] Uzgil, B. D., Aguirre, J. E., Bradford, C. M., & Lidz, A. 2014, ApJ, 793, 116, doi: 10.1088/0004-637X/793/2/116

[57] Visbal, E., & Loeb, A. 2010, JCAP, 11, 016, doi: 10.1088/1475-7516/2010/11/016

[58] Wang, R., Carilli, C. L., Neri, R., et al. 2010, ApJ, 714, 699, doi: 10.1088/0004-637X/714/1/699

[59] White, M. 2001, MNRAS, 321, 1, doi: 10.1046/j.1365-8711.2001.03956.x

[60] Wyithe, J. S. B., & Loeb, A. 2008, MNRAS, 383, 606, doi: 10.1111/j.1365-2966.2007.12568.x

[61] Yue, B., Ferrara, A., Pallottini, A., Gallerani, S., & Vallini, L. 2015, MNRAS, 450, 3829, doi: `10.1093/mnras/stv933`

Chapter 4

PHASE-SPACE SPECTRAL LINE DE-CONFUSION IN INTENSITY MAPPING

Cheng, Y.-T., Chang, T.-C., & Bock, J. J. 2020, ApJ, 901, 142, doi: 10.3847/1538-4357/abb023

Line intensity mapping (LIM) is a promising tool to efficiently probe the three-dimensional large-scale structure by mapping the aggregate emission of a spectral line from all sources that trace the matter density field. Spectral lines from different redshifts can fall in the same observed frequency and be confused, however, which is a major challenge in LIM. In this work, we develop a line de-confusion technique in map space capable of reconstructing the three-dimensional spatial distribution of line-emitting sources. If multiple spectral lines of a source population are observable in multiple frequencies, using the sparse approximation, our technique iteratively extracts sources along a given line of sight by fitting the LIM data to a set of spectral templates. We demonstrate that the technique successfully extracts sources with emission lines present at a few σ above the noise level, taking into account uncertainties in the source modeling and presence of continuum foreground contamination and noise fluctuations. As an example, we consider a TIME/CONCERTO-like survey targeting [C II] at the epoch of reionization, and reliably reconstruct the 3D spatial distribution of the CO interlopers and their luminosity functions at $0.5 \lesssim z \lesssim 1.5$. We also demonstrate a successful de-confusion for the SPHEREx mission in the near-infrared wavelengths.We discuss a formalism in which the reconstructed maps can be further cross-correlated with a (galaxy) tracer population to estimate the total interloper power. This technique is a general framework to extract the phase-space distribution of low-redshift interlopers, without the need of external information, for any line de-confusion problem.

4.1 Introduction

Line intensity mapping (LIM) has emerged as a promising tool to study the three-dimensional large-scale structures by mapping a particular spectral line emission, and infers the line-of-sight distance of the emission sources from the frequency-redshift relation. LIM measures the aggregate emission of all sources to constrain the bulk properties of the galaxies; whereas in traditional galaxy surveys, only the brighter sources can be individually detected. This relatively low spatial resolution and point-source sensitivity requirement in LIM enables the use of small apertures to efficiently scan a large survey volume out to high redshifts.

Several spectral lines have been proposed for LIM survey. The 21 cm hyperfine emission from neutral hydrogen [12, 45, 61, 71], the CO rotational lines [4, 6, 7, 9, 16, 26, 30, 34, 35, 39, 40, 47, 57, 60, 69], the [C II] 157.7 μm fine-structure line [26, 29, 63, 68, 72], and the Lyman-α emission line [17, 19, 20, 26, 31, 58, 64] are amongst the most studied lines in the LIM regime.

One of the main challenges in LIM is the astrophysical foreground contaminations, including the continuum emission and line interlopers. Although the continuum foregrounds are usually a few orders of magnitude brighter than the lines (a situation more severe for 21cm than for other lines), their smooth spectral feature can be used to distinguish from the line signals. This has been extensively studied in the context of 21 cm LIM [e.g., 2, 13, 27, 43, 49, 54, 67]. The line interlopers, which originate from sources residing in different redshifts emitting spectral lines in the same observed frequency channel, is another pressing issue for LIM experiments. The two most studied line de-confusion techniques, source masking and cross-correlation, typically rely on external data sets that trace the same cosmic volume: the masking technique makes use of a galaxy survey catalog to identify and remove bright interloper sources [5, 63, 66, 72], whereas cross-correlation of an LIM survey with an external (or internal) data set can help extract signals of interest [10, 11, 16, 20, 29, 31, 42, 48, 63, 69]. In addition to these two methods, Lidz & Taylor [41] and Cheng et al. [14] use the anisotropy of the interloper power spectrum arising from projection to the target line redshift to separate the lines. Gong et al. [28] distinguish the lines from the same projection effect but using the multipole power spectrum. de Putter et al. [22] propose to use angular fluctuations of the light to reconstruct the 3D source luminosity density.

Most of the existing line de-confusion methods (e.g. cross-correlation and power spectrum anisotropy) only extract the two-point statistics (power spectrum or correlation function) but lose the phase information of individual line maps, which are valuable for cosmological parameter constraints and systematics control in the data. With individual line maps, one can extract information beyond two-point statistics in the non-Gaussian intensity maps, especially ones from the epoch of reionization (EoR). For example, Breysse et al. [3] and Ihle et al. [33] show that the one-point statistics of the intensity field can help constrain the luminosity function model. In addition, individual line maps can be used directly as density tracers for various cross-correlation, multi-tracer analysis, de-lensing of the cosmic microwave background (CMB), and for performing consistency tests on different spatial regions with different foreground properties.

In this work, we develop a technique to extract individual line intensity maps from an LIM data set with blended interlopers. Using the fact that when multiple spectral lines emitted by a source are observable in an LIM survey, the redshift of the source can be pinned down by fitting to a set of spectral templates that are unique at each redshift. Without any external tracers or spectroscopic follow-up observations, individual line maps can be directly derived. For demonstration, we apply our technique to simulated data of an LIM survey targeting the EoR [C II] line with multiple low-redshift CO interlopers. In this case, the intensity field of the low-z CO lines ($0.5 \lesssim z \lesssim 1.5$) can be reconstructed since they can be detected in multiple spectral channels.

Kogut et al. [37] first explored the map-space line de-confusion using the multi-line wavelength information in the context of a pencil-beam spectroscopic survey. In this work, we explore the technique in the LIM regime that has a much lower sensitivity and the spectral resolution. In this regime, our template-fitting-based technique can obtain the signal-to-noise ratio (S/N) of the desired signals by using the data from multiple frequency channels.

The recent work by Moriwaki et al. [50] demonstrated the feasibility of LIM phase-space de-confusion with deep learning. They show that, in the absence of noise or foreground components, their algorithm can reconstruct the individual line maps that are mixed in the LIM data set. Their training data generation relies on the assumption of the signal clustering and the line luminosity model, whereas in this work, we develop a line de-confusion technique that only makes use of the spectral feature of the lines, which is more robust against the model uncertainty and the noise.

This chapter is organized as follows. First we introduce the model and the survey parameters we used to generate the mock LIM survey data in Sec. 4.2. Then, Sec. 4.3 describes our line de-blending technique. Sec. 4.4 presents the results on the fiducial setup. In Sec. 4.5, we present the performance of the technique with more practical considerations, and discuss its applications and extensions. The conclusions are given in Sec. 4.6. Throughout this paper, we consider a flat ΛCDM cosmology with

$n_s = 0.97$, $\sigma_8 = 0.82$, $\Omega_m = 0.26$, $\Omega_b = 0.049$, $\Omega_\Lambda = 0.69$, and $h = 0.68$, consistent with the measurement from Planck [55].

4.2 Mock Light Cone Construction

For each spatial pixel in an LIM survey, there are a set of spectral channel measurements. Hereafter, the term "light cone" refers to the collection of the spectral measurements in a single pixel. The line de-confusion method introduced in this work is performed on a pixel-by-pixel basis, which only utilizes the spectral information in an individual line of sight (light cone), without taking into account the spatial clustering information, which we leave for future work.

We test our de-confusion technique on simulated light cones. Since the clustering information is not relevant to our technique, we generate light cones that are based solely on the spectral line luminosity function models and not on the clustering properties. Thus, all light cones are independent from one another, and we also ignore the line-of-sight clustering in this work. This allows for both the light cone construction and de-confusion by parallelization to speed up without affecting the quantification of performance.

As a demonstration of the technique, we assume an LIM experiment targeting the redshifted [C II] fine-structure emission from the EoR: the LIM data set contains multiple low-z CO rotational transitions from low redshifts as interlopers. We note that the technique can be readily applied to any line-confusion problem at other wavelengths.

Line Signal Models

We model the line emissions of the redshifted [C II] emission and five low-redshift CO J-transitions: {CO(2–1), CO(3–2), CO(4–3), CO(5–4), CO(6-5)}. In reality, in the sub-mm spectral range of interest (generally in the ~ 200-300 GHz range), there are higher CO J lines (that are fainter), Galactic and extragalactic dust continuum emissions (the cosmic infrared background), CMB radiation, and atmospheric emissions that can all contribute to the measurements. Since none of them will produce strong spectral features that impact the performance of our technique, we will not include them in the light cones. Instead, in Sec. 4.5, we will demonstrate that the continuum foreground mitigation in the data analysis process has a negligible impact on our technique performance.

We use the [C II] and CO luminosity function models provided by Popping et al.

[56], in which a semi-analytic model including the effect of radiative transfer was used to estimate the CO and [C II] luminosity functions as constrained by current observations. Here we adopt their fitted Schechter luminosity functions to construct our light cones.

Survey Parameters

We consider a mock experiment that has similar survey parameters as the two ongoing EoR [C II] LIM experiments, TIME [18] and CONCERTO [38].

The mock survey covers 200–305 GHz with 70 evenly spaced spectral channels ($\delta\nu = 1.5$ GHz), and the $\Omega_{\rm pix} = 0.43^2$ arcmin^2 pixel size. We assume the instrument noise is white and has a Gaussian distribution, with four different per-pixel noise levels of standard deviation $\sigma_n = \{10^3, 5\times10^3, 10^4, 5\times10^4\}$ Jy sr^{-1}. These values are comparable to the range of expected instrument noise in TIME and CONCERTO.

Light Cone Generation

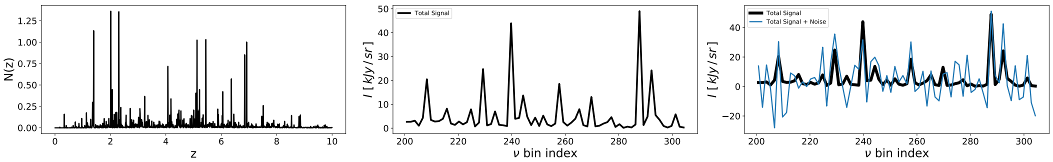

Figure 4.1: The steps of constructing a light cone. Left: $N(z)$, the effective number of ℓ_* sources in each redshift bin in this light cone. Middle: the signals from all six spectral lines and the sources in the 70 spectral channels. Right: the mock observed light cone (blue) consists of the signals (middle panel) and the $\sigma_n = 10$ kJy Gaussian noise. The black line is the signal component (same as the middle panel) for reference.

Based on the assumed luminosity function models and survey parameters, we populate the light cones with sources drawn from random realizations of the Schechter function model. We first define a fiducial Schechter function [1] for all of the lines as a function of redshift,

$$\Phi(\ell) = \phi_* \, (\mathcal{L}/\mathcal{L}_*)^{\alpha} \, e^{-(\mathcal{L}/\mathcal{L}_*)}. \tag{4.1}$$

We discretize the luminosity and redshift into bins of $\Delta(\mathcal{L}/\mathcal{L}_*)$ in 100 luminosity bins in log-space from $\mathcal{L}/\mathcal{L}_* = 10^{-3}$ to 10, and 2000 redshift bins of $\Delta z = 5\times10^{-4}$

---

[1] For the fiducial Schechter function, we choose the ϕ_* and α values of CO(1–0) for $z \leq 5$ and [C II] for $z > 5$ in [56]. We interpolate or extrapolate the Schechter function parameters to the desired redshift from the values of $z = [0, 1, 2, 3, 4, 6]$ given by [56].

in linear space from $z = 0$ to $z = 10$. The expectation value of the source counts N within each $\Delta\mathcal{L}$ and Δz bin is given by

$$\langle N \rangle = \Phi(\ell)[\Delta(\mathcal{L}/\mathcal{L}_*)][\Omega_{\text{pix}} D_A^2(z)\frac{d\chi}{dz}\Delta z], \tag{4.2}$$

where the last square bracket is the co-moving volume of the voxel defined by spatial pixel Ω_{pix} and redshift bin Δz, D_A is the co-moving angular diameter distance, and χ is the co-moving distance.

For each $(\mathcal{L}, z)$ bin, we assign its source counts to a Poisson random number with expectation value $\langle N \rangle$, and then integrate along $\mathcal{L}$ to get the total luminosity in each redshift bin $\mathcal{L}_{\text{tot}}(z)$. We define $N(z) \equiv \mathcal{L}_{\text{tot}}/\mathcal{L}_*$ as the "effective number counts" per redshift bin. The left panel of Fig. 4.1 shows the $N(z)$ of one example light cone.

Next, we assign the line luminosity signals in each redshift bin to $N(z)\,\ell_*^{\text{line}}$, where ℓ_*^{line} is the ℓ_* value of the line in the Popping et al. [56] model. We then project the line signals to their corresponding spectral channels to make the light cones. The middle panel of Fig. 4.1 shows an example light cone spectrum from the $N(z)$ shown in the left panel.

Finally, we add a Gaussian random fluctuation with a an rms σ_n value to each channel to account for the instrumental noise. The right panel of Fig. 4.1 shows the same light cone with a $\sigma_n = 10$ kJy Gaussian noise.

This light cone construction procedure assumes that all of the spectral lines have the same Schechter function parameters (ϕ_* and α) and thus the same luminosity function shape, and the CO spectral line energy distribution (SLED) is also fixed. That is, by construction, all of the sources have the same line luminosity ratio sets by the relative value of ℓ_*^{line}. In the main parts of this work, we use this fixed SLED model, and we test the impact of adding the SLED variation on the performance of our technique in Sec. 4.5.

We further point out that, even though the luminosity function has been sampled to the faint end ($10^{-3}\ell_*$), the light cone signals are still dominated by a few bright peaks. This indicates that the emission field can be well described by the few bright sources in the data. In other words, to extract the line emission field from a single line, one only needs to determine the redshift and luminosity of those bright sources. This is the main concept of our de-confusion technique, detailed in Sec. 4.3.

4.3 Methods

Formalism

The intensity of a light cone in frequency channel ν_i can be expressed as the linear combination of signals from all N_z redshift bins and the noise n_i,

$$I(\nu_i) = \sum_{j=1}^{N_z} \widetilde{A}_{ij}\, N(z_j) + n_i, \tag{4.3}$$

where $N(z_j)$ is the effective number of ℓ_* sources in redshift z_j (Sec. 4.2), and $\widetilde{A}_{ij}$ converts $N(z_j)$ to the observed intensity in channel ν_i.

$\widetilde{A}$ is an $N_{\rm ch} \times N_{\rm z}$ matrix, where we have $N_{\rm ch} = 70$ spectral channels, and we use $N_{\rm z} = 2000$ ($\Delta z = 5{\times}10^{-4}$) redshift bins. In principle, we can set Δz to infinitesimally small values; however, in practice the redshift resolution is limited by the instrument spectral resolution, as the SLEDs of nearby galaxies can be highly degenerate when their respective spectral lines fall in the same set of observed channels. Setting $N_{\rm z} = 2000$ in fact gives a much finer redshift resolution compared to the spectral channel width. We will use this fine redshift resolution as a starting point, and discuss the strategy to reduce redshift bins and therefore remove redundant information at the end of Sec. 4.3.

Most of the elements in $\widetilde{A}_{ij}$ are zeros except for which the sources at z_j emit a spectral line at the observed frequency ν_i. In this case,

$$\widetilde{A}_{ij} \equiv I_*^{\rm line}(z_j) = \ell_*^{\rm line}(z_j)\frac{1}{4\pi D_L^2(z_j)\delta\nu_i\Omega_{\rm pix}}, \tag{4.4}$$

where $\ell_*^{\rm line}$ is the line luminosity of an ℓ_* source in the model, D_L is the luminosity distance, and $I_*^{\rm line}$ is defined as the observed line intensity of an $\ell_*^{\rm line}(z_j)$ source at z_j. Fig. 4.2 shows the $I_*^{\rm line}$ in our assumed model[2]. Note that Eq. 4.3 assumes that all of the sources at the same redshift have the same SLED, and we will first build our technique based on this assumption. In reality, the SLED varies across galaxy type. In Sec. 4.5, we will show that our method also works in the realistic level of SLED variation.

Eq. 4.3 can be written in the matrix form,

$$\mathbf{I} = \widetilde{\mathbf{A}}\mathbf{N} + \mathbf{n}, \tag{4.5}$$

---

[2] The glitch in CO(5–4) at $z \sim 1$ is due to a feature in the Schechter luminosity function model in [56]. In their CO(5–4) luminosity function fit at $z = 1$, a slightly higher L_* and and a lower α value are derived that cause the slight apparent discontinuity in redshift, even though the luminosity function does not show an abrupt change at $z = 1$.

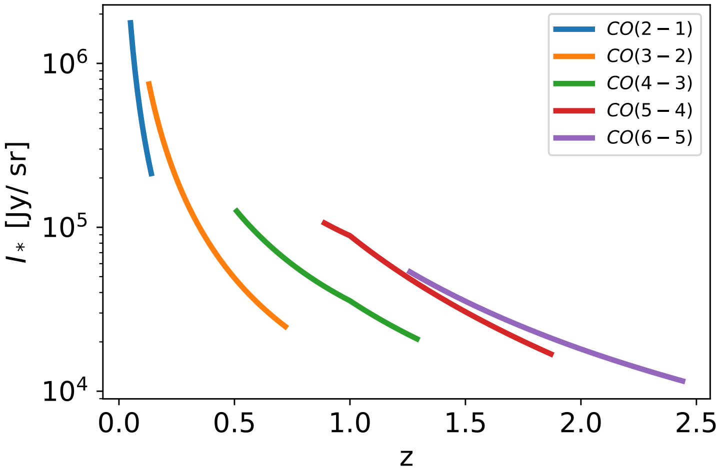

Figure 4.2: Intensity of the CO lines I_* from the sources of characteristic luminosity ℓ_*.

Table 4.1: Frequencies and Redshifts of the Six Defined Broad Bands

Name	Line	ν bin index	ν (GHz)	$\langle z\rangle(z_{\text{min}} - z_{\text{max}})$
J3 high	CO(3–2)	51–69 (19 bins)	200.75–227.75	0.61 (0.51–0.72)
J4 low	CO(4–3)	0–24 (25 bins)	268.25–304.25	0.61 (0.51–0.72)
J4 high	CO(4–3)	41–69 (29 bins)	200.75–242.75	1.09 (0.89–1.30)
J5 low	CO(5–4)	0–36 (37 bins)	250.25–304.25	1.09 (0.89–1.30)
J5 high	CO(5–4)	37–69 (33 bins)	200.75–248.75	1.59 (1.30–1.87)
J6 low	CO(6–5)	0–42 (43 bins)	241.25–304.25	1.56 (1.26–1.87)

NOTE: The CO(2–1) and CO(3–2) overlapping redshifts of $0.13 < z < 0.5$ can also be reconstructed, but they are only covered by four frequency channels, which makes it difficult to quantify the reconstruction performance with sufficient statistical power. Therefore, we ignore these redshifts in our analysis.

where $\mathbf{I}$ and $\mathbf{n}$ are N_{ch}-element column vectors, $\mathbf{N}$ is an N_z-element column vector, and $\widetilde{\mathbf{A}}$ is an $N_{\text{ch}} \times N_z$ matrix. The top panel of Fig. 4.3 shows the $\widetilde{\mathbf{A}}$ matrix in our model. The $\widetilde{\mathbf{A}}$ matrix is mostly zeros, and the six curves from left to right are the six spectral lines, CO(2–1), CO(3–2), CO(4–3), CO(5–4), CO(6–5), and [C II].

The goal of our line de-confusion technique is to solve for the source count vector $\mathbf{N}$ in a given observed light cone data $\mathbf{I}$ and a model $\widetilde{\mathbf{A}}$ matrix. With the $\mathbf{N}$ solution, the intensity map of individual spectral lines $\mathbf{I}_{\text{line}}$ can be reconstructed by $\mathbf{I}_{\text{line}} = \widetilde{\mathbf{A}}_{\text{line}}\mathbf{N}$, where $\widetilde{\mathbf{A}}_{\text{line}}$ is $\widetilde{\mathbf{A}}$ when only the target spectral line signals are turned on.

We further rewrite Eq. 4.3 by normalization and reducing the nuisance information. First, we normalize the columns in $\widetilde{\mathbf{A}}$ and move the normalization factor $\mathbf{I}_j^{\text{norm}}$ to

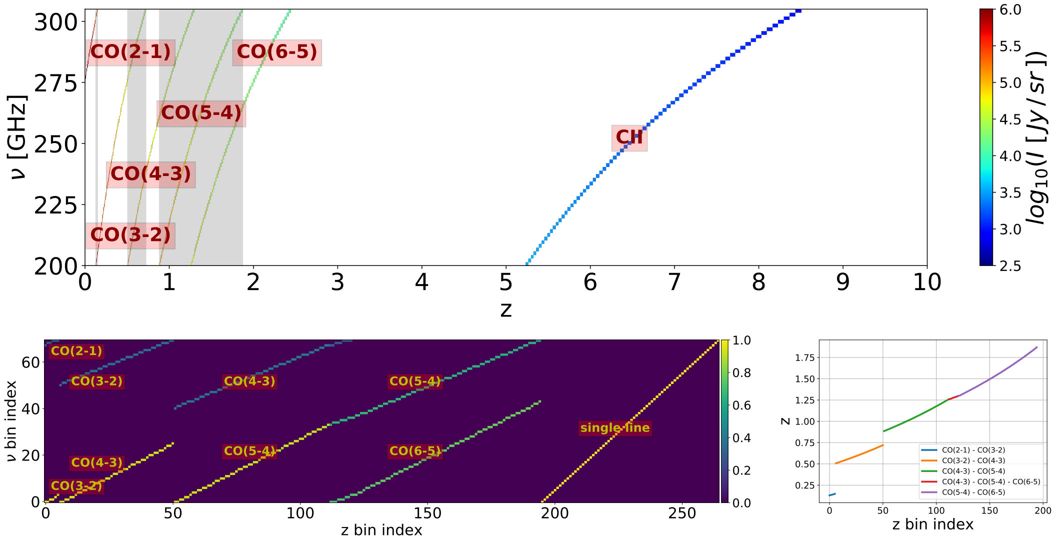

Figure 4.3: Top: $\widetilde{\mathbf{A}}$ matrix with our model on 70 frequency channels and 2000 redshift bins. $\widetilde{\mathbf{A}}$ are zeros (white) for the majority of the elements, and the six curves from left to right correspond to the six spectral lines, CO(2–1), CO(3–2), CO(4–3), CO(5–4), CO(6–5), and [C II]. The color scale indicates line intensities in the fiducial model. The gray shaded regions are the redshifts at which sources can be observed in multiple lines, and thus can be reconstructed in our technique. Bottom left: $\mathbf{A}$ matrix with $N_{\rm ch} \times N_{\rm z} = 70 \times 265$ size, which is the reduced and normalized $\widetilde{\mathbf{A}}$. The color scale represents the intensities with the fiducial model SLED normalized within each column, i.e. for all column j, $\sum_{i=1}^{N_{\rm ch}} \mathbf{A}_{ij}^2 = 1$. Bottom right: the redshift of the 195 multi-line redshift bins in $\mathbf{A}$, which are the redshift bins that have multiple CO lines observable in our mock survey. The colors label the pairs of detectable CO lines. The redshifts not covered by multiple CO lines cannot be reconstructed with our technique ($0.15 \leq z \leq 0.51$; $0.72 \leq z \leq 0.89$; $z \geq 1.87$).

their $\mathbf{N}$ element:

$$\mathbf{I} = \mathbf{A}\widetilde{\mathbf{N}} + \mathbf{n}, \tag{4.6}$$

where

$$\begin{aligned} \mathbf{A}_{ij} &= \widetilde{\mathbf{A}}_{ij} \,/\, \mathbf{I}_j^{\rm norm}, \\ \widetilde{\mathbf{N}}_j &= \mathbf{N}_j \, \mathbf{I}_j^{\rm norm}, \\ \sum_{i=1}^{N_{\rm ch}} \mathbf{A}_{ij}^2 &= 1. \end{aligned} \tag{4.7}$$

Next, we will reduce the nuisance or redundant elements in Eq. 4.6. The columns of $\mathbf{A}$ are the basis spanning the observed data space. In constructing $\widetilde{\mathbf{A}}$, we simply design the columns to be equally spaced redshift bins as shown in the top panel of

Fig. 4.3. However, this natural basis is highly degenerate. To remove the nuisance information, we first discard the redshift bins that are zero vectors in $\widetilde{\mathbf{A}}$. The sources in these redshifts do not emit lines in observable frequencies, and thus no information can be used to constrain their $N(z)$. Second, for the redshift bins containing only one CO line, their normalized columns in $\mathbf{A}$ are identical to other columns having a [C II] signal in the same frequency channel. In other words, given an observed data $\mathbf{I}$, we cannot distinguish the origin of the source with a single line emission. Therefore, we combine these identical columns into a single column. In conclusion, we keep the columns in $\mathbf{A}$ with redshift bins that can be observed with multiple spectral lines, plus an identity matrix for the redshift bins that only have a single detectable line.

For all of the redshifts that can be observed in multiple lines, we design the size of the redshift bins based on the following two competing considerations. On the one hand, the redshift bins have to be small enough to faithfully represent the emitting source distribution. On the other hand, finer redshift bins give larger $\widetilde{\mathbf{N}}$ size, and therefore more unknown parameters to be solved. The information can be compressed by combining some neighboring redshift bins, which are highly degenerate, since they have signals in the same channels with similar amplitudes. Therefore, we design the redshift bins using the following procedures: we (1) generate $\mathbf{A}$ with fine redshift bins ($\Delta z = 5 \times 10^{-4}$ from $z = 0$ to 10)[3], (2) keep the columns with multiple lines, (3) then identify the group of neighboring columns that have signals in the same sets of channels, and (4) keep the medium bin and discard the others. With this process, we get 195 non-degenerate columns. Hereafter, "multi-line redshift bins" refers to these 195 redshift bins that can be detected in multiple channels. Finally, we append the N_{ch}-sized identity matrix (that accounts for the single-line redshift bins) to these 195 columns to generate $\mathbf{A}$. The bottom left panel in Fig. 4.3 shows matrix $\mathbf{A}$ which has size $N_{\mathrm{ch}} \times N_z = 70 \times 265$ (195 multi-line redshifts plus 70 columns of identity matrix). The bottom right panel shows the redshift of the 195 multi-line bins, and their color labels are the pairs of detectable CO lines. We also define six broad bands from these pairs of lines by binning groups of channels. Table 4.1 lists the definition of the broad bands. Note that the CO(2–1) and CO(3–2) overlapping redshifts of $0.13 < z < 0.5$ can also be reconstructed, but they are only covered by four frequency channels, which makes it difficult to quantify the reconstruction performance with sufficient statistical power. Therefore, we ignore these redshifts

---

[3] Note that with these fine bins, the same frequency bin can map to multiple redshift bins instead of an one-to-one mapping.

in our analysis.

Our line de-confusion technique can only solve $\widetilde{\mathbf{N}}$ in the 195 multi-line redshift bins. The last 70 elements in $\widetilde{\mathbf{N}}$ that correspond to the identity matrix in $\mathbf{A}$ are nuisance parameters, since they represent degenerate “single-line” signals from different redshifts — that is to say, we cannot reconstruct the signals in the single-line redshifts, which are the regions not covered by the lines in the bottom left panel of Fig. 4.3 ($0.15 \leq z \leq 0.51$; $0.72 \leq z \leq 0.89$; $z \geq 1.87$). In the following analysis, we will only focus on the reconstruction of $\widetilde{\mathbf{N}}$ in the 195 multi-line redshift bins.

Sparse Approximation

The key step in our de-confusion technique is to solve for $\widetilde{\mathbf{N}}$ in Eq. 4.6, given the observed spectrum $\mathbf{I}$ and model $\mathbf{A}$. This type of linear system has been extensively studied in the context of CMB map making, in which $\mathbf{I}$, $\mathbf{n}$, and $\widetilde{\mathbf{N}}$ in Eq. 4.6 can be analogized to the time-ordered data, time-stream noise, and the pointing matrix, respectively. However, contrary to the map-making problem, our system is an ill-posed problem as there are more unknown variables ($N_z = 265$) than the input data points ($N_{\rm ch} = 70$). Thus the standard map-making algorithm [e.g., 65] cannot be applied.

In Eq. 4.6, the columns of $\mathbf{A}$ form a basis for $\mathbf{I}$, and the solution $\widetilde{\mathbf{N}}$ is the linear combination coefficient. The columns of $\mathbf{A}$ form an over-complete basis, since the $\mathbf{A}$ matrix is only of rank $N_{\rm ch}$, and thus the solution $\widetilde{\mathbf{N}}$ is not unique. Indeed, for any given observed data $\mathbf{I}$, there are infinite $\widetilde{\mathbf{N}}$ that can perfectly fit the input.

Nevertheless, Eq. 4.6 can be solved with the “sparse” condition, which means the preferred solution of $\widetilde{\mathbf{N}}$ is the one with a small number of nonzero elements. With this constraint, we can solve Eq. 4.6 with the following well-defined optimization problem:

$$\underset{\left\|\widetilde{\mathbf{N}}\right\|_0}{\operatorname{argmin}} \frac{1}{N_{\rm ch}} \left\|\mathbf{I} - \mathbf{A}\widetilde{\mathbf{N}}\right\|_2^2 < \epsilon^2, \tag{4.8}$$

where the ℓ_0-norm $\|\cdot\|_0$ is the number of nonzero elements, and ϵ sets the threshold of error tolerance of the fit.

This type of problem is known as “sparse approximation,” which has been extensively studied in the context of signal processing and compressive sensing [8, 24]. The sparse approximation algorithms solve the sparse representation of the signal in a “dictionary” that is composed of a set of “atoms,” and represent the signal in the data in terms of the linear combination of a few atoms in the dictionary. In Eq. 4.8, the dictionary is the matrix **A**, and the atoms are the column vectors of **A**.

The sparse approximation can only be applied if $k \equiv \left\|\widetilde{\mathbf{N}}\right\|_0 \ll N_{\mathrm{ch}}$. Note that the sparsity of the problem is quantify by k/N_{ch} but not k/N_z, since k/N_z can always be designed to be arbitrarily small by choosing a large basis (fine redshift bins). However, the degree of freedom in the solution is restricted by the input size N_{ch}, and thus the $k \ll N_{\mathrm{ch}}$ condition prohibits the algorithm from using more parameters than the input degree of freedom to over-fit the data.

In general, LIM light cones are not sparse, since there is always a large number of faint sources in the typical luminosity function (e.g. Schechter function), so all of the elements in $\widetilde{\mathbf{N}}$ are nonzero. However, as mentioned in Sec. 4.2, the light cone signals are dominated by only a few bright sources, and the intensity field can be well described by them. Consequently, the parameter k in our problem can be quantified by the "effective" number of these bright sources per voxel that contribute most of the emission. Following [15], we define the effective number N_{eff} as:

$$N_{\mathrm{eff}}(z) \equiv \frac{\left(V_{\mathrm{vox}} \int d\ell \, \Phi(\ell, z) \, \ell\right)^2}{V_{\mathrm{vox}} \int d\ell \, \Phi(\ell, z) \, \ell^2}, \tag{4.9}$$

where V_{vox} is the voxel size of the redshift bin. Note that Φ is the number of sources per luminosity per voxel, and therefore N_{eff} is dimensionless and is proportional to the voxel size. N_{eff} can be interpreted as the reciprocal of the effective shot noise in LIM, which is analogous to the $1/N$ shot noise in a galaxy power spectrum. If the luminosity function Φ follows the Schechter function form, then N_{eff} is (approximately) the number of sources brighter than ℓ_*, which contributes the majority of the emission.

We can estimate k by the cumulative $N_{\mathrm{eff}}(z)$ along the line of sight per light cone. Fig. 4.4 shows the cumulative N_{eff} in our model. While $N_{\mathrm{eff}} \sim 100$ from $z = 0$ to 10, the only relevant range is $z \lesssim 2.5$, where CO lines fall in the observed frequency range (see Fig. 4.3, top panel). Above $z \sim 2.5$, only the high-z [C II] lines can be observed, but they are much fainter then the CO signals and the assumed noise level, so they can be treated as background fluctuations. Therefore, for $z \lesssim 2.5$, we find $N_{\mathrm{eff}} \sim 10 \ll N_{\mathrm{ch}} = 70$, so the sparse condition is qualified in our problem.

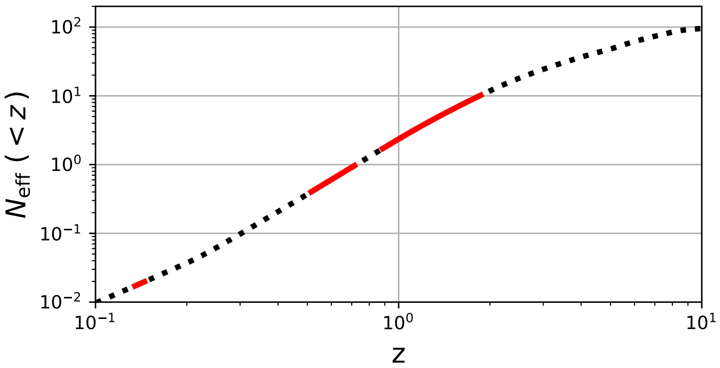

Figure 4.4: The redshift cumulated $N_{\rm eff}$ in our model (black dashed line). The red segments mark the multi-line redshift ranges that have multiple observable CO lines. Thus the de-confusion technique can be applied.

The Matching Pursuit Algorithm

We use the matching pursuit (MP) algorithm first introduced by Mallat & Zhang [46] to solve for Eq. 4.8. The MP algorithm iteratively selects an "atom" in the "dictionary" to project out part of the signals in the data, and keep track of the current solution of the signal and residual for the next step until the stopping criteria is met. In our case, the column vectors in **A** are the atoms that form the dictionary space. A detailed description of the MP algorithm is in Appendix .1.

In each step of the MP algorithm, the selected atom is the one that has the maximum inner product with the residual. The S/N of the signals in each step is the ratio of that maximum inner product to the instrument noise level σ_n (see Appendix .2 for the proof). Therefore, if we set the stopping criteria to be the maximum inner product smaller than m times of σ_n, then this is an m-σ detection threshold on the signals (e.g., $m = 5$ for a 5σ detection). Note that the detection threshold here is based on the combined information in multiple spectral channels projected onto the dictionary space.

The choice of detection threshold "m" is a trade-off between the purity and completeness of the source extraction. Higher "m" values give a higher purity map, whereas lower "m" values pick out fainter sources at the cost of increased false detections from noise. The optimal value of "m" depends on the instrument sensitivity and the purpose of the reconstruction map. For example, to reconstruct the line luminosity function, one might use a higher threshold to reduce the false detections at the faint end; whereas to constrain the large-scale structure, a lower threshold is preferred to reduce the shot noise in the power spectrum. An analytical formalism to determine the optimal threshold is to make use of the Fisher information framework [15], where one calculates the expectation value of the desired observable (e.g. power spectrum) as a function of threshold for a given signal model and noise level, and estimates the threshold that optimizes the Fisher information. Alternatively, one can simply perform test simulations with different thresholds to determine the optimal value. Both approaches can provide guidance on choosing the optimal threshold for the problem at hand, and a detailed investigation is beyond the scope of this paper.

4.4 Results

We present line de-confusion results in the simple case where mock light cones and the template (**A**) are both generated from the same signal model (Sec. 4.2). We demonstrate that in this scenario, our technique is capable of extracting low-z CO signals in the presence of realistic instrumental noise. We discuss the robustness of the performance against uncertainties in the signal model and contamination from astrophysical foregrounds, and extend the application to spectral lines in different wavelengths in Sec. 4.5.

We quantify the reconstruction performance by computing two statistics on the true and reconstructed data: (1) the Pearson correlation coefficient (Sec. 4.4) and (2) the voxel intensity distribution (VID; Sec. 4.4). The former quantifies the phase-space information, whereas the latter captures the one-point statistics that describes the distribution of voxel intensities.

Finally, we present results with a variety of instrument noise levels σ_n and the reconstruction threshold m. For each test, we use 2500 mock light cones to calculate the correlation coefficient and VID, and estimate errors with 100 noise realizations.

Visualization of Example Results

Fig. 4.5 visualizes the reconstruction results on one of the spectrum bins (274 GHz) that contains five lines (CO(3–2), CO(4–3), CO(5–4), CO(6–5), [C II]) in the observable band. We place the 400 light cones into a 20×20 pixel map, and show the true and the reconstructed intensity fields. Note that the input signals do not exhibit the spatial clustering because each light cone is generated independently. The top left panel shows the total signal intensity from all five lines, and the top right panel shows the observed intensity including the total line signals and a Gaussian instrument noise with $\sigma_n = 10^4$ Jy sr^{-1}. In this channel, three of the lines (CO(4–3), CO(5–4), CO(6–5)) are in the multi-line regime, so they can be reconstructed with

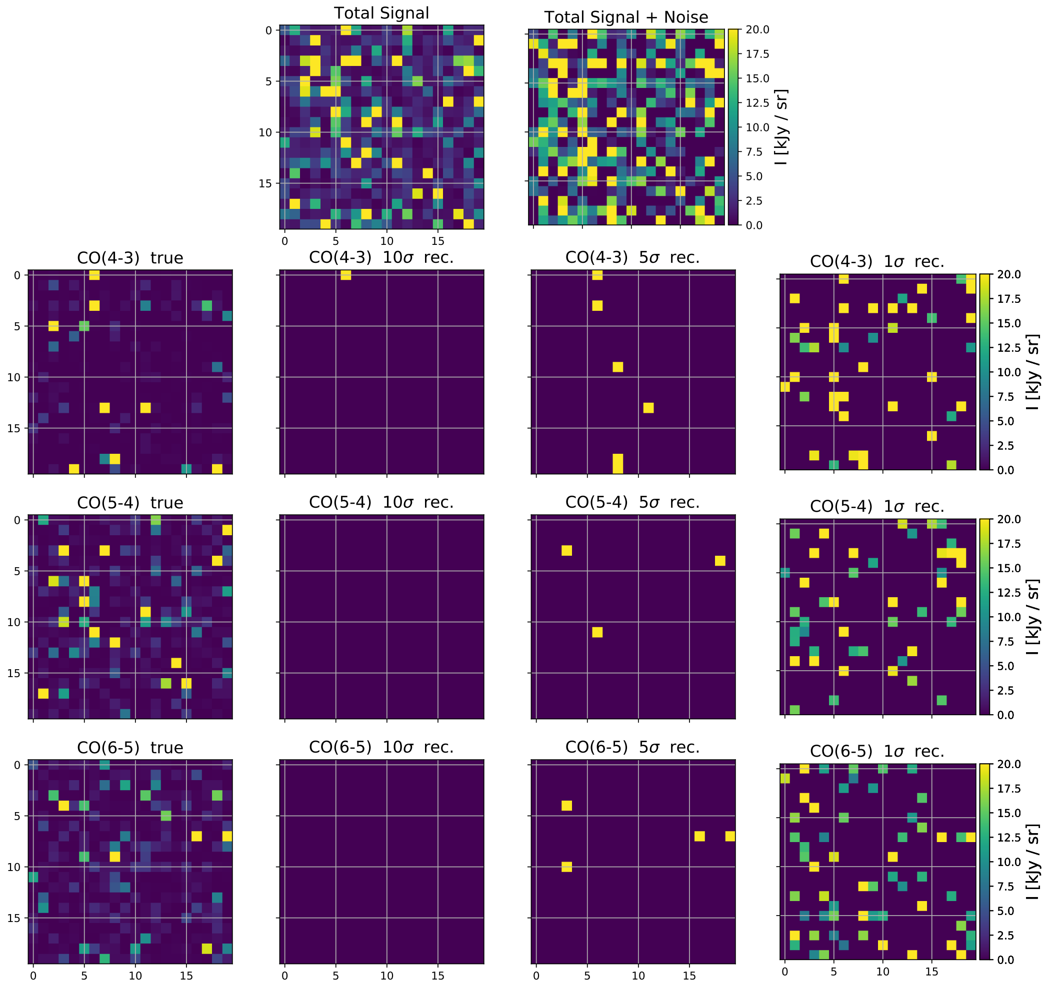

Figure 4.5: Visualization of the phase-space reconstruction with $\sigma_n = 10^4$ Jy sr^{-1}. We place the 400 toy model light cones into a 20×20 pixel map, and show the reconstruction results on one of the spectrum bins (274 GHz). The top left panel is the true signal intensity map from all of the spectral lines (CO(3–2), CO(4–3), CO(5–4), CO(6–5), [C II]). The top right panel is the observed intensity map including line signals (top left panel) and noise. The bottom three panels show the true input and the 10σ, 5σ, and 1σ reconstructed (right) emission field of the three spectral lines in the multi-line regime where the signal can be reconstructed.

our algorithm. The three bottom panels compare the true input to the reconstructed intensity maps for these lines with 10σ, 5σ, and 1σ reconstruction threshold as the MP algorithm stopping criteria.

The choice of threshold is a trade-off between the completeness and the purity in the reconstructed map. As shown in Fig. 4.5, in the high threshold case (10σ), the MP algorithm only extracts a few bright sources that are above the reconstruction threshold. As the threshold decreases, more sources are being reconstructed, at the cost of increased false positive detection from the noise fluctuations or the interlopers. This example provides a visual depiction of the reconstruction algorithm. To further quantify the reconstruction performance, we consider two summary statistics in the following sections.

Pearson Correlation Coefficient

We quantify the reconstruction performance by the Pearson correlation coefficient between the true and the reconstructed maps in each channel. The Pearson correlation coefficient is defined by

$$r = \frac{\sum_{i=1}^{N_{\rm lc}} \left(I^i_{\rm true} - \langle I_{\rm true}\rangle\right)\left(I^i_{\rm rec} - \langle I_{\rm rec}\rangle\right)}{\sqrt{\sum_{i=1}^{N_{\rm lc}} \left(I^i_{\rm true} - \langle I_{\rm true}\rangle\right)^2}\sqrt{\sum_{i=1}^{N_{\rm lc}} \left(I^i_{\rm rec} - \langle I_{\rm rec}\rangle\right)^2}}, \tag{4.10}$$

where $N_{\rm lc} = 2500$ is the number of light cones, $I^i_{\rm true}$ and $I^i_{\rm rec}$ are the true and the reconstructed line intensity maps, respectively, at the ith light cone. Fig. 4.6 shows the results of the correlation coefficient with a $\sigma_n = 10^4$ Jy sr^{-1} noise level and a 5σ reconstruction threshold. Our reconstructed map achieves $\sim 80\%$ correlation with the true input map at $z \lesssim 1.5$.

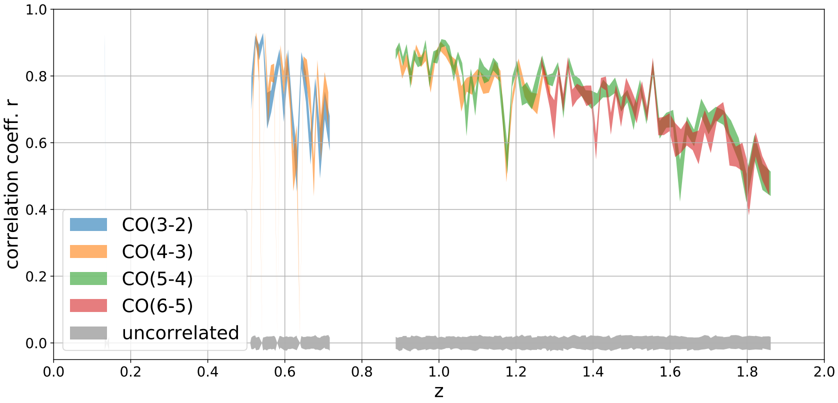

Figure 4.6: Pearson correlation coefficient r between the true and the reconstructed maps on 2500 light cones with $\sigma_n = 10^4$ Jy sr^{-1} 5σ reconstruction. The bands are the 1σ scatter of 100 noise realizations with the sample line signal. The gray bands are the r value with the white noise map for reference.

Fig. 4.7 shows the correlation coefficient r within a range of noise level σ_n and the reconstruction threshold m. For simplicity, we show the average r value of each broad band defined in Table 4.1. The key findings are summarized below.

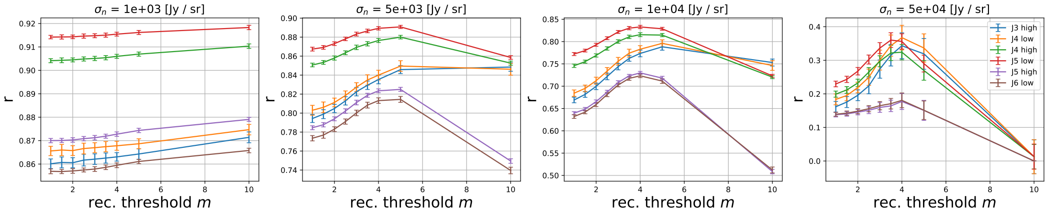

Figure 4.7: Average correlation coefficient r in each broad band (Table 4.1) within a range of noise level σ_n and the reconstruction threshold m. The error bars are the rms of 100 noise realizations.

- At a fixed m value, r decreases as σ_n increases. This is because under the same purity (same m), the detection threshold $m\sigma_n$ is higher for higher σ_n, and thus fewer sources have been reconstructed.
- The six defined bands correspond to three pairs of lines from different redshift bands (Table 4.1). The lines within each pair are strongly correlated because they are the signals from the same sources, and thus they are reconstructed in the same MP iteration. The pairs of lines in the same redshift are reconstructed in the same MP iteration, and as a result, they are highly correlated.
- Because of the purity and completeness trade-off, the maximum correlation r happens at the intermediate threshold m (except for the lowest-noise $\sigma_n = 10^3$ Jy sr^{-1} case, discussed in the next enumerated point).
- Correlation coefficient r has very low dependency on $\sigma_{\rm th}$ in the $\sigma_n = 10^3$ Jy sr^{-1} case. This can be understood by comparing the noise level σ_n to the quantity I_*, the intensity of the ℓ_* source in the Schechter function (Fig. 4.2). At the redshift range in which we perform the reconstruction ($0.5 \lesssim z \lesssim 1.9$), $10^4 \lesssim I_* \lesssim 10^5$ Jy sr^{-1}, which indicates that $m\sigma_n < I_*$ for all of the m values considered in Fig. 4.7 ($m = 1 \sim 10$). In the Schechter luminosity function, the sources $\gtrsim \ell_*$ contribute the majority of the information in the intensity field [15], and thus the correlation coefficient r is not sensitive to the change in reconstruction threshold if $m\sigma_n \ll \ell_*$.

VID

The Pearson correlation coefficient traces the phase-space variations between the true and reconstructed maps, but it cannot distinguish a systematic constant offset, i.e., if the reconstructed line signals are systematically lower or higher than the true input. Therefore, we check the consistency of the reconstructed and true input maps using the one-point statistics, VID. Note that we do not directly compare the mean intensity of each map since we only reconstruct the bright sources that are above the threshold and neglect all faint sources in the reconstructed map, so the mean intensity is not expected be faithfully recovered.

The VID of an LIM map contains information beyond the power spectrum, and is valuable for LIM targeting a late-time universe where the large-scale structure is highly non-Gaussian and cannot be fully described by two-point statistics. For example, Breysse et al. [3] showed that the VID can constrain the luminosity function model parameters, and Ihle et al. [33] demonstrated that a joint analysis of the power spectrum and VID improved the constraining power on the source luminosity function.

Fig. 4.8 compares the VID of the true and reconstructed maps in the six broad bands with noise level $\sigma_n = 10^4$ Jy sr^{-1} and reconstruction threshold $m = 4$. The gray lines compose the VID of the total observed map, which includes signals from all of the lines and noise. The black lines are the VID of the target line maps, and the red data points are the VID of the reconstructed target line map.While the VID of the total observed map is one to two orders of magnitude above that of the target line signal, our reconstruction technique can faithfully recover the VID of the signal to slightly below the I_*-scale, the characteristic source luminosity in the Schechter function.

In this realistic survey setup with the assumed signal model, we show that our method can successfully reconstruct the VID of the CO signal down to $\sim \ell_*$ scales. This provides a strong constraint on the CO luminosity function, as both are 1D statistics of the intensity field and are closely related. The CO luminosity function at various redshifts provides valuable insight on the formation and evolution of galaxies across cosmic time. Specifically, CO is a tracer of H_2 gas in the interstellar medium, and can therefore be used to study the evolution of the molecular gas content and its distribution as a function of time [23, 59, 70]. We note though that the expected S/N on the luminosity function depends on the assumed model.

4.5 Discussion

Model Uncertainty

For the results presented in Sec. 4.4, the light cone signals and the dictionary template **A** are both generated from the same assumed signal model (Sec.4.2). However, in reality, the variation in SLED across galaxies will affect the reconstruction per-

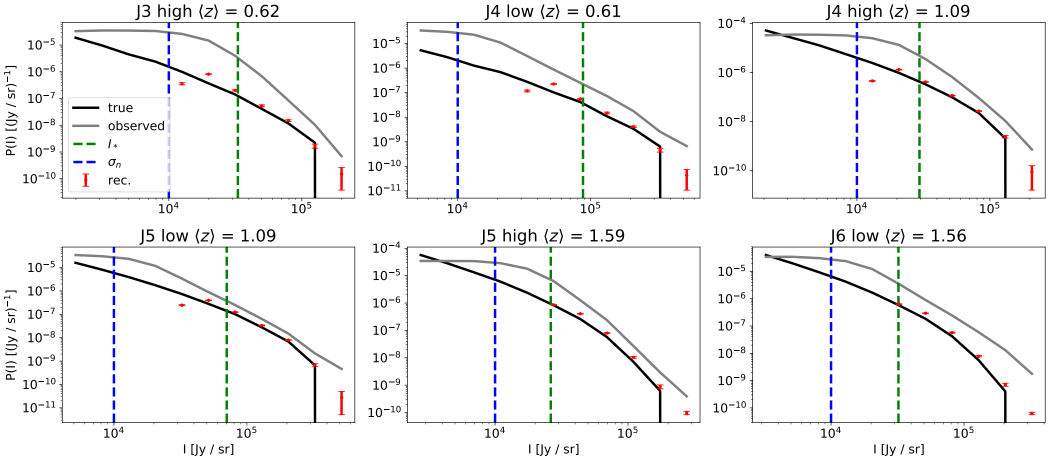

Figure 4.8: VID of the true input (black) and the reconstructed (red) maps in the six broad bands with noise level $\sigma_n = 10^4$ Jy sr^{-1} and reconstruction threshold $m = 4$. The error bars are the rms of 100 noise realizations. The gray curves are the VID of the total observed map, which includes signals from all of the lines and noise. For reference, the blue and green dashed lines mark the noise level σ_n and I_*, respectively.

formance. To test how the SLED uncertainties affect the reconstruction, we apply three different SLED model variations and bias levels (at 20%, 50%, and 100%) to the mock data, and run the reconstruction with the same dictionary template **A**. We detail the definition of variation and bias below. Daddi et al. [21] measured multiple CO lines of ULIRGs at $z \sim 1.5$, and estimated a $\sim 20\%$ variation on the CO SLED ratio for their sample. Therefore, an assumed 50% or 100% variation can be more extreme than realistic variations.

SLED model variation

First, we test the case with SLED model variations. For each line of each source in the mock light cone, we assign a line luminosity $L^{\rm line}$:

$$L^{\rm line} = L^{\rm line}_{\rm fid}(1 + \delta_L), \tag{4.11}$$

where $L^{\rm line}_{fid}$ is the fiducial luminosity from our model, and δ_L is a zero-mean Gaussian random variable with the standard deviations of 0.2, 0.5, and 1.0 for the 20%, 50%, and 100% variation cases, respectively.

Fig. 4.9 shows the reconstructed correlation coefficients r with a 4σ reconstruction threshold ($m = 4$) with different SLED variations. We see that introducing 20%

SLED variations in the model has a very mild impact on the reconstructed correlation; whereas with 50% (100%) variation, the correlation coefficient drops by about 10% (~ 50%). Fig. 4.10 compares the VID. We see that the SLED fluctuation has a negligible impact on the VID reconstruction in the 20%, 50%, and 100% variation cases. Hence, we conclude that our technique is robust against realistic level of CO SLED fluctuations.

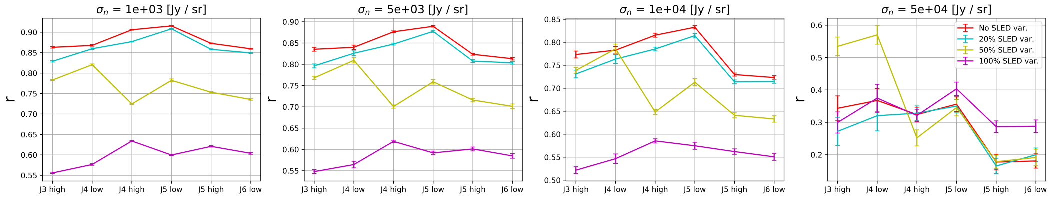

Figure 4.9: Comparing r of the no SLED variation (red), 20% variation (cyan), 50% variation (yellow), and 100% variation (purple) with the 4σ reconstruction threshold. The values are the average of r within the channels of the band, and the error bars are the rms of 100 noise realizations of all of the spectral bins in each band.

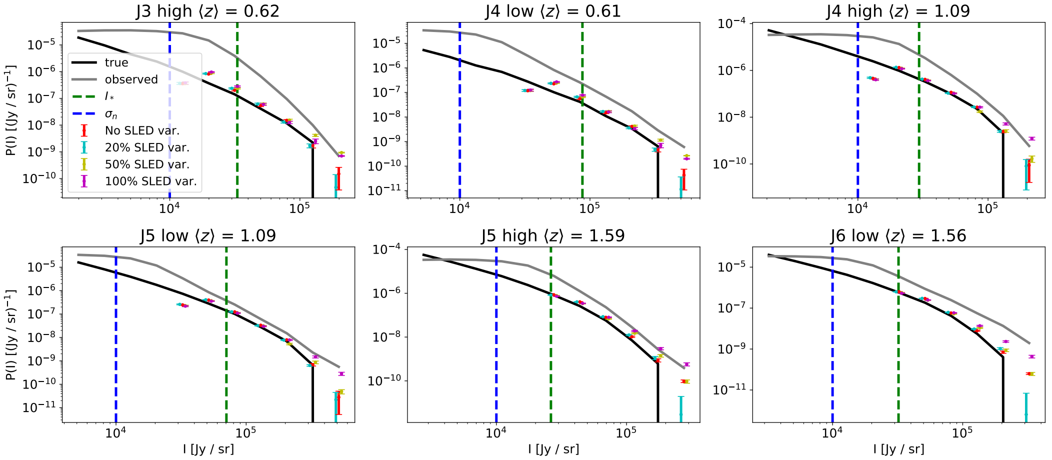

Figure 4.10: VID of the no SLED variation (red), 20% variation (cyan), 50% variation (yellow), and 100% variation (purple), and the true (black) maps in the six broad bands with noise level $\sigma_n = 10^4$ Jy sr^{-1} and reconstruction threshold $m = 4$. The error bars are the rms of 100 noise realizations. The gray curves are the VID of the total observed map, which includes signals from all of the lines and noise. For reference, the blue and green dashed lines mark the noise level σ_n and I_*.

Model offset

In addition to SLED variation, we also test whether our model gives biased estimates of the average SLED by assigning line luminosity $L^{\rm line}$ as

$$L^{\rm line} = L^{\rm line}_{\rm fid}(1 + b_L), \tag{4.12}$$

where b_L is a constant offset that we assign to the model lines, and it is applied to all of the sources in the mock light cones. For the 20% bias level, we apply $b_L = (+0.1, -0.1, +0.1, -0.1)$ for the four CO lines CO(3–2), CO(4–3), CO(5–4), and CO(6–5) to ensure the SLED ratio between neighboring lines is 20%, since our algorithm is only sensitive to the line ratio between two neighboring CO lines.

Fig. 4.11 and Fig. 4.12 show the reconstructed correlation coefficients r and VID with a 4σ with three SLED bias levels. Similar to Fig. 4.9, the correlation only drops significantly when the bias is tuned to 100%. Therefore, our technique is also robust against a realistic level of potential CO SLED bias.

Application to LIM in Other Wavelengths

The technique developed in this work is not restricted to the [C II] and CO lines' blending problem. It can in principle be applied to a range of LIM experimental setups. As a demonstration, we apply our method to reconstruct near-infrared lines in an SPHEREx-like survey.

SPHEREx is an ongoing NASA MIDEX mission to conduct an all-sky near-infrared spectro-imaging survey [25].[4] SPHEREx will carry out the first all-sky spectral survey at wavelengths between 0.75 and 5 μm with 96 spectral channels and a $6''.2$ pixel size. Lyα (121.6 nm), Hα (656.3 nm), Hβ (486.1 nm), [O II] (372.7 nm), and [O III] (500.7 nm) are the five prominent lines detectable by SPHEREx across a range of redshifts especially in the LIM regime.

The line signal model is described in Appendix .3. We generate a near-infrared LIM mock data with a $6''.2 \times 6''.2$ pixel size and a 5σ point-source sensitivity of $m_{\rm AB} = 22$ (similar depth as the SPHEREx deep fields), and run our de-confusion algorithm on the mock light cones. Fig. 4.13 shows the results of the correlation coefficients between the true and a 3σ-threshold reconstructed intensity maps. The reconstructed map achieves $\sim$ 80% correlation with the true input map at $z \lesssim 3$, and decreases toward higher redshifts, as the I_* of the lines approaches the noise

---

[4] http://spherex.caltech.edu.

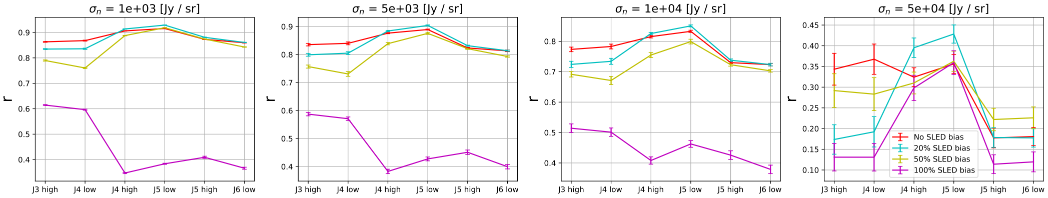

Figure 4.11: Comparison r of the no SLED bias (red), 20% bias (cyan), 50% bias (yellow), and 100% bias (purple) with a 4σ reconstruction threshold. The values are the average of r within the channels of the band, and the error bars are the rms of 100 noise realizations of all of the spectral bins in each band.

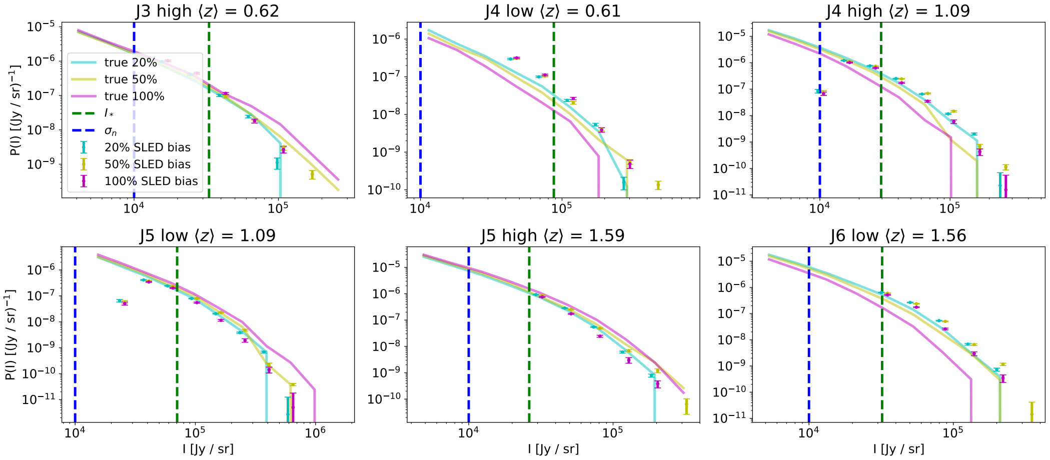

Figure 4.12: VID of no SLED bias (red), 20% bias (cyan), 50% bias (yellow), 100% bias (purple), and the true maps (solid lines) in the six broad band with noise level $\sigma_n = 10^4$ Jy /sr and reconstruction threshold $m = 4$. The error bars are the rms. of 100 noise realizations. The grey curves are the VID of the total observed map that includes signals from all the lines and noise. For reference, the blue and green dashed lines mark the noise level σ_n and I_*.

level σ_n. At $z \sim 5$, the brightest line, Hα, is redshifted into SPHEREx bands with a high spectral resolution of $R \sim 130$ and suffers less signal dilation, resulting in an increase of S/N on a single source detection, and thus r slightly rebounds at this redshift.

To account for uncertainties in modeling the SLED, we apply a realistic level of line luminosity variations from Moustakas et al. [51]. We apply 10%, 50%, and 100% SLED variations (Eq. 4.11) to the line luminosity ratio of H$_\alpha$/Hβ, [O II]/Hα, and [O III]/[O II], respectively. Fig. 4.14 shows the correlation coefficients between the true and a 3σ-threshold reconstructed intensity maps. Comparing with the fixed SLED case (Fig. 4.13), only the $H\beta$ line shows a significant decrease in the

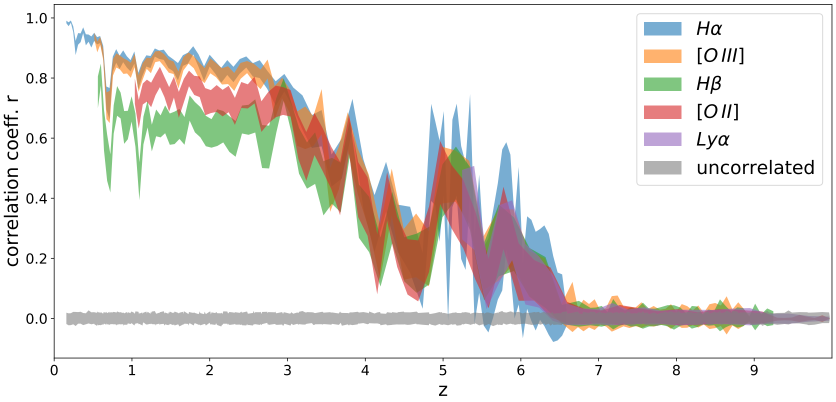

Figure 4.13: Pearson correlation coefficient r between the true and the reconstructed maps on 2500 light cones for SPHEREx-like mock data. The bands are the rms. of the value in 100 noise realizations with the sample line signal. The gray bands are the correlation coefficient with the uncorrelated white noise map for reference.

performance. The intensity map of the three brighter lines (Hα, [O II], and [O III]) can still be extracted with $\gtrsim 70\%$ correlation compared to the true input.

We conclude that our algorithm can reasonably well reconstruct the phase-space LIM signal in a SPHEREx-like experiment, given the expected variation of (redshifted) optical line ratios. The technique can be generalized to different LIM experimental applications, and the reconstructions are fairly robust against uncertainties in the SLED modeling.

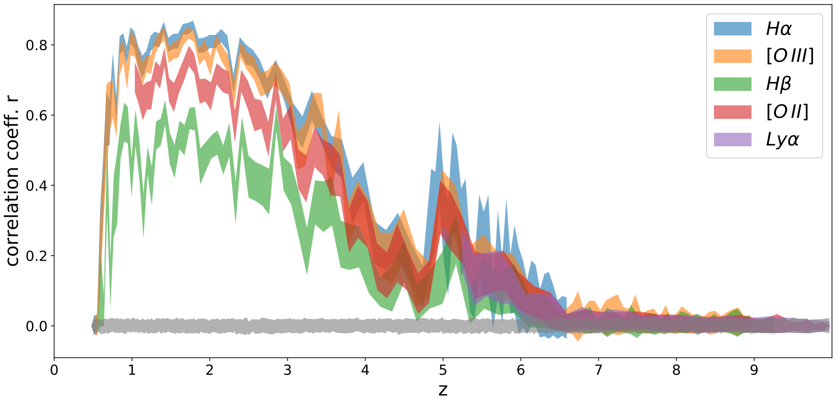

Figure 4.14: Pearson correlation coefficient r between the true and the reconstructed maps on 2500 light cones for SPHEREx-like mock data with a realistic level of SLED variation. The bands are the rms of the value in 100 noise realizations with the sample line signal. The gray bands are the correlation coefficient with the uncorrelated white noise map for reference.

Foreground Subtraction

In addition to line interlopers, LIM data are subjected to strong continuum foreground from various sources. For the frequency range considered in this work ($\sim$ 200-300 GHz), the dominant foregrounds are the atmospheric emission, dust continuum and the CMB; whereas for LIM in the near infrared, e.g. SPHEREx [25], the zodiacal light and the galaxy stellar continuum are the dominant continuum foregrounds. Even though these foregrounds are brighter than the sought-after line signals, their spectral responses are expected to be smooth and are distinct from the spectral line features, so that the continuum foregrounds can be separated and mitigated, for example by a smooth function fit such as a low-order polynomial[5]. Here we test how the foreground mitigation process affects our line reconstruction results.

We consider two cases of foreground mitigation. First we emulate the foreground removal process in the presence of an approximately constant foreground in both the spatial and spectral dimensions, for example the zodiacal light. In this case, before running the reconstruction, we subtract the mean value of the whole data cube, i.e. the mean intensity in $N_{\rm lc} \times N_\nu$ voxels. The second case is to emulate the continuum subtraction process of the galaxy stellar or dust continuum, which are expected to have smooth spectra but different in each light cone, since each light cone contains different galaxies with different continuum spectrum. We fit and subtract a first-order polynomial function to the spectrum of each light cone before running the reconstruction.

The results of a 4-σ-threshold reconstruction with different noise σ_n level are shown in Fig. 4.15 and Fig. 4.16. We can see compared to the no background subtraction case, the r value is even higher in these two tests. This is because our reconstruction only extracts the bright lines, and the fainter lines act as a background for the MP algorithm. The signals from the fainter lines introduce not only fluctuations but also a bias in the data, since the line signals are always positive unlike the zero-mean noise. The reconstruction performance is improved after background subtraction because this bias level is also removed during this process. For the VID results, we see that there is no significant difference compared to the no background subtraction case. This is again due to the fact that the background level is much fainter than the brightness of the sources being extracted with our algorithm, so the background

---

[5] Some foreground components are also spatially smooth (e.g. zodiacal light) that can be filtered in the spatial domain as well.

subtraction has no impact on the reconstruction. In conclusion, the background subtraction in the LIM data reduction pipeline will not affect our line reconstruction technique.

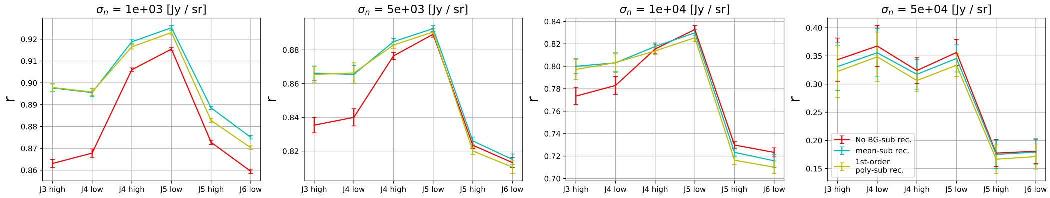

Figure 4.15: Comparing r of the reconstructed map with the no background subtraction (red), subtracting the mean of the whole data cube (cyan), and subtracting a first-order polynomial in the spectral direction for each light cone (yellow), using a 4σ reconstruction threshold. The values are the average of r within the channels of the band, and the error bars are the rms of 100 noise realizations of all of the spectral bins in each band.

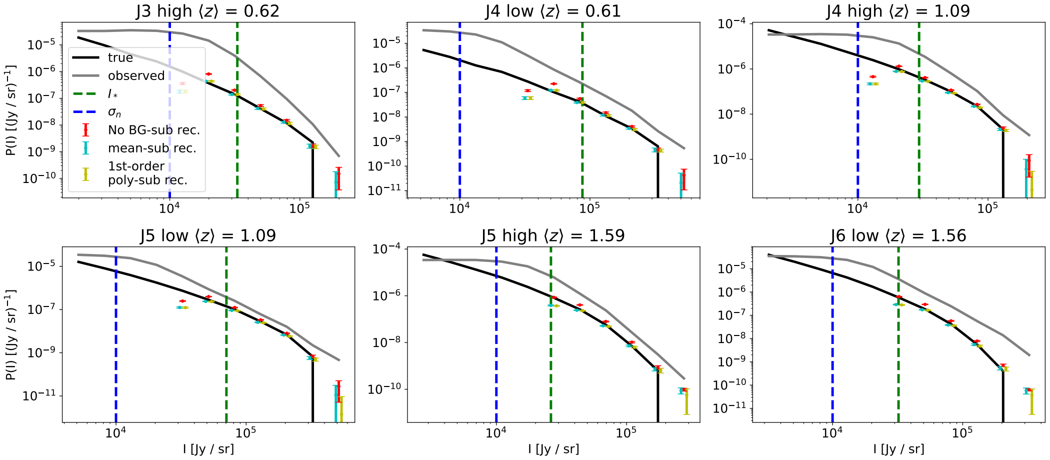

Figure 4.16: VID of no background subtraction (red), subtracting the mean of the whole data cube (cyan), subtracting a first-order polynomial in spectral direction for each light cone (yellow), and the true (black) maps in the six broad band with noise level $\sigma_n = 10^4$ Jy /sr and reconstruction threshold $m = 4$. The error bars are the rms of 100 noise realizations. The grey curves are the VID of the total observed map that includes signals from all the lines and noise. For reference, the blue and green dashed lines mark the noise level σ_n and I_*.

Prior with External Catalogs

Our analysis uses the LIM data itself without invoking any external information. In practice, initial LIM survey fields are designed to in part overlap with existing photometric or spectroscopic galaxy surveys, and thus there will be information provided by external galaxy catalogs to aid the line de-confusion problem. Ignoring redshift uncertainties of the external catalogs, one simple approach to incorporate the external information is to force the MP algorithm to first select the redshift bins that contain galaxies from the catalog. After iterating through the catalog sources, we then continue the normal MP procedure until hitting the stopping criteria.

To test the effect of including prior knowledge from external catalogs, we generate a mock catalog by selecting sources with CO(5–4) flux greater than 150 $L_\odot$ Mpc^{-2} (6×10^{-17} W m^{-2}) in each light cone. The source density in the catalog is ~ 1.2 per light cone (integrated along line of sight) for a 0.43^2 arcmin^2 pixel solid angle. The flux cut corresponds to an L_* galaxy at $z \sim 2$. According to Helgason et al. [32], such an L_* galaxy has an absolute magnitude $M_{\rm AB} \sim -23$ in the optical, which gives an apparent magnitude of $m_{\rm AB} \sim 21.8$, approximately the depth of the assumed optical catalog.

With this mock external catalog, we identify the redshift bins containing the catalog sources, regardless of the noise level and threshold value. After projecting out these components, we run the MP algorithm on the residual data as per usual until we hit the stopping criteria. The results of a 4-σ-threshold reconstruction with different noise levels are shown in Fig. 4.17. The reconstruction shows improved results for all noise levels. However, the huge improvement in the highest-noise case cannot be interpreted as a successful reconstruction of source intensities. At this high noise level, the signals are well buried under the noise, and when we fit the data with the catalog source redshift templates, the extracted components are dominated by noise rather than signal amplitude. Thus the improved correlation is merely due to the position information imposed by the external catalogs.

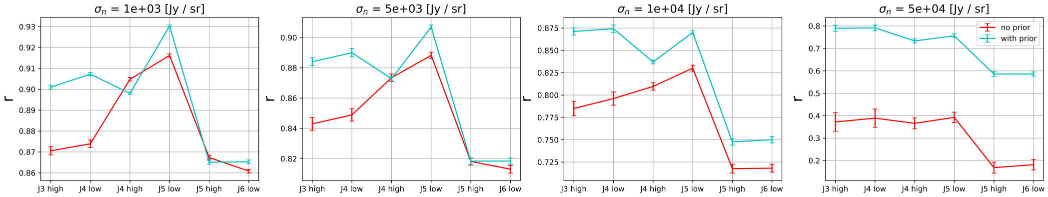

Figure 4.17: Comparing r of the reconstructed map with no external catalog prior (red), and utilizing the external catalog prior to fit the catalog sources before running the MP reconstruction (cyan), using 4-σ reconstruction threshold. The values are the average of r within the channels of the band, and the error bars are the rms of 100 noise realizations of all the spectral bins in each band.

Comparing with the Limit of No Interlopers

In the case of no interloper lines (i.e. only single-line emission in the data), the best estimator of the single-line intensity map is the observed map (regardless of foregrounds). To compare our reconstruction performance with this limiting case, we calculate the correlation coefficient between the input single-line map with the same single-line map plus the instrument noise.

Fig. 4.18 shows the results using $\sigma_n = 10^4$ Jy sr^{-1}. Comparing to Fig. 4.6, which has the same noise level, the single-line-plus-noise case has a lower correlation than the reconstruction in Fig. 4.6. Especially for the fainter lines (e.g., CO(3–2) at $z \sim 0.6$ and CO(4–3) at $z \sim 1.1$), our reconstruction map has a much better correlation. This can be explained by the fact that in our algorithm, the sources are detected in the template space rather than in a single voxel. That is, if a source can be observed in two frequency channels, we extract the source by projecting the signals in these two channels to the template space, which is effectively combining the information from both channels. Consequently, we are able to achieve a better S/N on the fainter lines because of the greater sensitivity of their brighter counterpart.

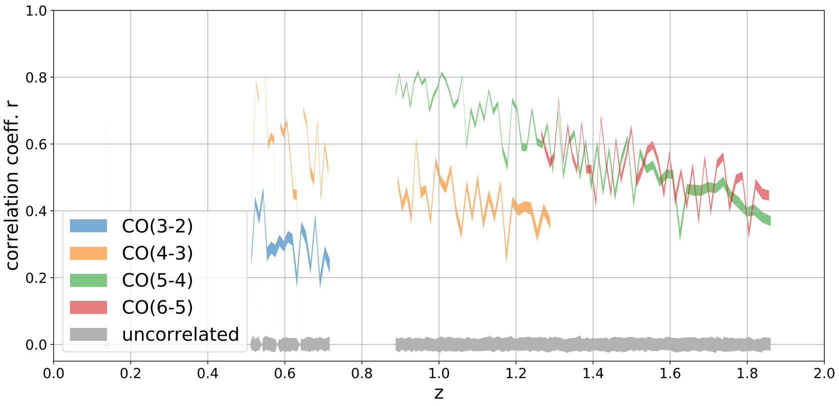

Figure 4.18: Pearson correlation coefficient r between the true line maps and the same maps adding the $\sigma_n = 10^{-4}$ Jy sr^{-1} noise on 2500 light cones. The bands are the rms of the value in 100 noise realizations with the sample line signal. The gray bands are the r value shown with the uncorrelated white noise map for reference.

Improving Cross-correlation Uncertainty

Another useful application of the reconstruction technique is for a more precise measurement of cross-correlation between LIM and other tracers. Cross-correlation analysis not only serves as a validation of a cosmological signal in LIM, since the cross-correlation is less susceptible to foreground contamination and other systematic effects, but also provides valuable astrophysical and cosmological information.

While cross-correlating the total observed LIM map with an external tracer gives an unbiased estimator of cross spectrum between the target line and the external tracer, the presence of continuum foregrounds and interlopers increases the error of this measurement. Here we present a simple argument that for cross-correlation analysis, using a reconstructed map instead of the total observed map can effectively reduce the error bars on the cross-power spectrum.

The cross-power spectrum errors between two fields δP_{12} in a single mode are given by

$$\delta P_{1,2}^2 = \frac{1}{2}\left(P_{1,2}^2 + \delta P_1 \delta P_2\right), \tag{4.13}$$

where $P_{1,2}$ is the cross-power spectrum, and δP_1 and δP_2 are the errors on the auto power spectrum in the two fields. For a single k mode, $\delta P_1 = P_1$, $\delta P_2 = P_2$, where P_1 and P_2 are the total power spectra (including signals and noise) measured in two fields. Eq. 4.13 can be expressed in terms of the cross-correlation coefficient: $r_{1,2} = P_{1,2}/\sqrt{P_1 P_2}$,

$$\delta P_{1,2}^2 = \frac{1}{2} P_{1,2}^2 \left(1 + \frac{1}{r_{1,2}^2}\right). \tag{4.14}$$

Say we have an external galaxy sample that traces one of the target CO lines, then we can write the total observed LIM data as the combination of target CO line ($I_{\rm CO}$), other interloper lines ($I_{\rm interlopers}$), and the noise (I_n),

$$I_{\rm tot} = I_{\rm CO} + I_{\rm interlopers} + I_n. \tag{4.15}$$

The expectation value of the cross spectrum between the observed total map and galaxy is the same as the cross spectrum with only the target CO, since the other components are not correlated with the large-scale structure at the same redshift, so $\langle P_{\rm g,tot} \rangle = \langle P_{\rm g,co} \rangle$.

For simplicity, we assume the galaxies are perfectly correlated with the target CO line field on the scale of interest, so $r_{\rm g,co} = 1$. This implies the galaxy field and CO field always have the same correlation r with any given field x, $r_{\rm g,x} = r_{\rm co,x}$.

From Eq. 4.14, the error on the galaxy-CO cross spectrum measured by cross-correlating the galaxy field with the total observed LIM data is

$$\begin{aligned}\delta P_{\mathrm{g,co}}^{\mathrm{tot}} &= P_{\mathrm{g,tot}}\sqrt{\frac{1}{2}\left(1+\frac{1}{r_{\mathrm{g,tot}}^2}\right)} \\ &= P_{\mathrm{g,co}}\sqrt{\frac{1}{2}\left(1+\frac{1}{r_{\mathrm{co,tot}}^2}\right)}.\end{aligned} \tag{4.16}$$

On the other hand, if we cross-correlate the galaxies with the reconstructed CO map ($\mathrm{CO^r}$), the error is

$$\delta P_{\mathrm{g,co^r}} = P_{\mathrm{g,co^r}}\sqrt{\frac{1}{2}\left(1+\frac{1}{r_{\mathrm{g,co^r}}^2}\right)}. \tag{4.17}$$

However, $P_{\mathrm{g,co^r}}$ is a biased estimator of $P_{\mathrm{g,co}}$ because of the error in the reconstruction. If we assume that the reconstructed map roughly preserves the same power as the true map, $P_{\mathrm{co^r}} \approx P_{\mathrm{co}}$, then we can write[6] $P_{\mathrm{g,co^r}} = r_{\mathrm{g,co^r}}P_{\mathrm{g,co}} = r_{\mathrm{co,co^r}}P_{\mathrm{g,co}}$. Therefore, we have to de-bias cross spectrum $P_{\mathrm{g,co^r}}$ by factor $1/r_{\mathrm{co,co^r}}$, $P_{\mathrm{g,co}} = P_{\mathrm{g,co^r}}/r_{\mathrm{co,co^r}}$. The value of $r_{\mathrm{co,co^r}}$ cannot be directly inferred from the data, so we have to estimate it by simulating the possible range of signals; this introduces an extra error term to the $\delta r_{\mathrm{co,co^r}}$ due to the uncertainty in $r_{\mathrm{co,co^r}}$,

$$\begin{aligned}\delta P_{\mathrm{g,co}}^{\mathrm{r}} &= P_{\mathrm{g,co}}\sqrt{\left(\frac{\delta P_{\mathrm{g,co^r}}}{P_{\mathrm{g,co^r}}}\right)^2+\left(\frac{\delta r_{\mathrm{co,co^r}}}{r_{\mathrm{co,co^r}}}\right)^2} \\ &= P_{\mathrm{g,co}}\sqrt{\frac{1}{2}\left(1+\frac{1}{r_{\mathrm{co,co^r}}^2}\right)+\left(\frac{\delta r_{\mathrm{co,co^r}}}{r_{\mathrm{co,co^r}}}\right)^2}.\end{aligned} \tag{4.18}$$

In summary, the S/N on the galaxy-CO cross spectrum $P_{\mathrm{g,co}}$ using the total observed map and the reconstructed map is

$$\begin{aligned}SNR^{\mathrm{tot}} &= \frac{P_{\mathrm{g,co}}}{\delta P_{\mathrm{g,co}}^{\mathrm{tot}}} = \frac{1}{\sqrt{\frac{1}{2}\left(1+\frac{1}{r_{\mathrm{co,tot}}^2}\right)}}, \\ SNR^{\mathrm{r}} &= \frac{P_{\mathrm{g,co}}}{\delta P_{\mathrm{g,co}}^{\mathrm{r}}} = \frac{1}{\sqrt{\frac{1}{2}\left(1+\frac{1}{r_{\mathrm{co,co^r}}^2}\right)+\left(\frac{\delta r_{\mathrm{co,co^r}}}{r_{\mathrm{co,co^r}}}\right)^2}}.\end{aligned} \tag{4.19}$$

---

[6] $P_{\mathrm{g,co^r}} = r_{\mathrm{g,co^r}}\sqrt{P_g P_{\mathrm{co^r}}} \approx r_{\mathrm{g,co^r}}\sqrt{P_g P_{\mathrm{co}}} = r_{\mathrm{g,co^r}}P_{\mathrm{g-co}}$, where the last equality uses the assumption $r_{\mathrm{g,co}} = 1$.

Note that the S/N in both cases converges to unity when r is unity and the de-bias error is zero, which is the limit of sample variance[7].

Fig. 4.19 shows the S/N^{tot} and S/N^{r} with two different values of $\delta r_{\text{co,co}^{\text{r}}}$ and $r_{\text{co,tot}}$ as a function of $r_{\text{co,co}^{\text{r}}}$. According to the calculations in Sec. 4.5, the value of $r_{\text{co,co}^{\text{r}}}$ ranges from ~ 0.7 to ~ 0.9 (except for the most noisy case, $\sigma_n = 5 \times 10^4$ Jy sr^{-1}). To estimate the realistic $r_{\text{co,tot}}$ value, we calculate the correlation of the input CO line maps and the observed map with $\sigma_n = 10^4$ Jy sr^{-1}, the same noise level as in Fig. 4.6. The results are shown in Fig. 4.20. We find that $r_{\text{co,tot}}$ are around 0.2-0.5 in this case. With this range of parameters, Fig. 4.19 indicates that S/N^{r} is better than S/N^{tot}, which means that in our model, using the reconstructed map instead of the observed map in cross-correlation can reduce the uncertainty.

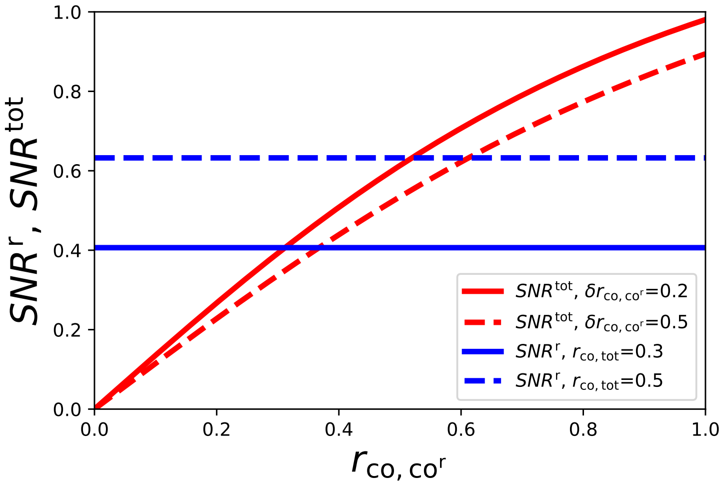

Figure 4.19: Comparing the S/N on the cross-power spectrum of one CO line and the external tracer, using the total observed map (S/N^{tot}) and the reconstructed CO map (S/N^{r}). We consider two different values of $\delta r_{\text{co,co}^{\text{r}}}$ and $r_{\text{co,tot}}$ that cover the range of realistic parameter values in our model. The value of $r_{\text{co,co}^{\text{r}}} \sim 0.7 - 0.9$ according to the results in Sec 4.4, and therefore using the reconstructed map instead of the total observed map in cross-correlation can reduce the uncertainty (i.e. $SNR^{\text{r}} > SNR^{\text{tot}}$).

---

[7] The power spectrum cross-correlation coefficient can be derived from the Pearson correlation (Eq. 4.10) with a weighting on pixels. Since our pixels are generated and reconstructed independently of each other, the Pearson correlation coefficient here is an unbiased estimator of the power spectrum correlation coefficient. Thus here we will use the value of Pearson correlation we derived for the power spectrum correlation coefficient.

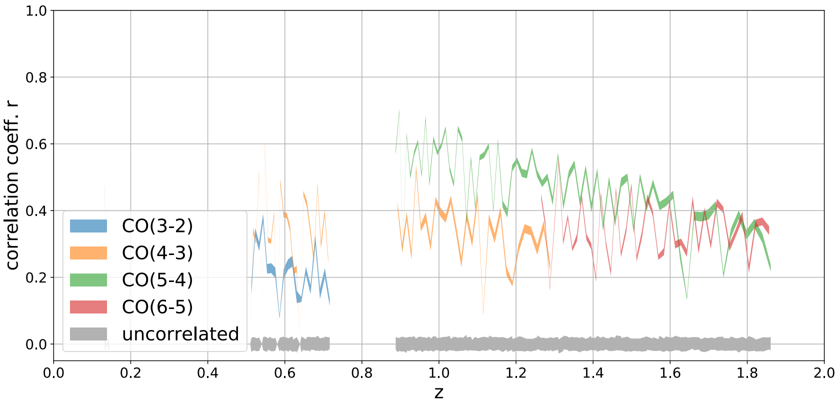

Figure 4.20: Pearson correlation coefficient r between the true input and the total observed maps on 2500 light cones with $\sigma_n = 10^4$ Jy sr^{-1}. The bands are the rms of the value in 100 noise realizations with the sample line signal. The gray bands are the r value shown with the uncorrelated white noise map for reference.

Estimating the complete interloper population

Here we discuss a potential extension of the technique, which is capable of estimating the *total* power of an interloper population. The method uses an incompletely reconstructed interloper sample and an incomplete, external tracer of the interloper density field. Once the interloper contribution is fully quantified, the high-redshift signal of interest in an LIM dataset can be estimated without bias. As an example, we write the observed intensity of a [C II] LIM dataset as

$$I^{\text{obs}} = I_{\text{CO}} + I_{\text{C\,II}} + \delta_n^{\text{obs}}, \tag{4.20}$$

where I_{CO} is the total CO interloper intensity, $I_{\text{C\,II}}$ is the C II signal, and δ_n^{obs} is the instrumental noise. For simplicity, here we only consider the contribution of one CO rotational line as the foreground.

Given a reconstruction threshold, we reconstruct the bright CO emissions using our technique as:

$$I_{\text{CO}}^{\text{rec}} = \alpha I_{\text{CO}_\text{b}} + \delta_n^{\text{rec}}, \tag{4.21}$$

where $I_{\text{CO}}^{\text{rec}}$ is the reconstructed CO intensity and I_{CO_b} is the intensity of bright CO sources in the reconstruction, which is a subset of the total CO population. There are two sources of error in the reconstruction: a multiplicative term α proportional to the bright CO intensity, where α can be greater or smaller than unity, and an additive term δ_n^{rec} describing random (or misidentified) fluctuations about the true CO intensity, which is uncorrelated with the CO field.

The auto power spectrum of the reconstructed CO map is

$$\begin{aligned}\left\langle I_{\mathrm{CO}}^{\mathrm{rec}} I_{\mathrm{CO}}^{\mathrm{rec}}\right\rangle &= \bar{\alpha}^2 \left\langle I_{\mathrm{CO_b}} I_{\mathrm{CO_b}}\right\rangle + \left\langle \delta_n^{\mathrm{rec}} \delta_n^{\mathrm{rec}}\right\rangle \\ &= \bar{\alpha}^2 \left\langle I_{\mathrm{CO_b}}\right\rangle^2 b_{\mathrm{CO_b}}^2 P(k) + N^{\mathrm{rec}},\end{aligned} \tag{4.22}$$

where, in the linear regime that we consider here, $b_{\mathrm{CO_b}}$ is the cosmological clustering bias of the bright CO sources, $P(k)$ is the matter density field, and N^{rec} is the auto power spectrum of the δ_n^{rec} term as a noise bias.

Using an external galaxy sample g in the same redshift range as the CO interlopers, we cross-correlate g with the observed and reconstructed maps, respectively, and consider only linear clustering scales:

$$\langle g I_{\mathrm{obs}}\rangle = r \langle I_{\mathrm{CO}}\rangle b_{\mathrm{CO}} b_{\mathrm{g}} P(k), \tag{4.23}$$

$$\left\langle g I_{\mathrm{CO}}^{\mathrm{rec}}\right\rangle = \left\langle g \alpha I_{\mathrm{CO_b}}\right\rangle = \bar{\alpha} r_b \left\langle I_{\mathrm{CO_b}}\right\rangle b_{\mathrm{CO_b}} b_{\mathrm{g}} P(k), \tag{4.24}$$

where $\langle\alpha\rangle = \bar{\alpha}$, b_g is the bias of the galaxy tracer, $b_{\mathrm{CO_b}}$ is the bias of the bright CO population, and r (r_b) is the astrophysical stochastic cross-correlation parameter between the galaxy and CO (CO$_b$) populations.

Finally, we construct an estimator of the full CO power spectrum as

$$\begin{aligned}\left\langle \widehat{I_{\mathrm{CO}} I_{\mathrm{CO}}}\right\rangle &= \left\langle I_{\mathrm{CO}}^{\mathrm{rec}} I_{\mathrm{CO}}^{\mathrm{rec}}\right\rangle \left(\frac{\left\langle g I_{\mathrm{CO}}^{\mathrm{obs}}\right\rangle}{\left\langle g I_{\mathrm{CO}}^{\mathrm{rec}}\right\rangle}\right)^2 \\ &= \left(\bar{\alpha}^2 \left\langle I_{\mathrm{CO_b}}\right\rangle^2 b_{\mathrm{CO_b}}^2 P(k) + N^{\mathrm{rec}}\right) \left(\frac{r \langle I_{\mathrm{CO}}\rangle b_{CO}}{\bar{\alpha} r_b \left\langle I_{\mathrm{CO_b}}\right\rangle b_{CO_b}}\right)^2 \\ &= \left(\left\langle I_{\mathrm{CO_b}}\right\rangle^2 b_{\mathrm{CO_b}}^2 P(k) + \frac{N^{\mathrm{rec}}}{\bar{\alpha}^2}\right) \left(\frac{r \langle I_{\mathrm{CO}}\rangle b_{CO}}{r_b \left\langle I_{\mathrm{CO_b}}\right\rangle b_{CO_b}}\right)^2 .\end{aligned} \tag{4.25}$$

If we can pick a reconstruction threshold such that the reconstruction of a subset of the CO population (bright CO sources) includes the majority of CO emitters, then $r \approx r_b$. Furthermore, if the reconstruction noise is negligible, $N^{\mathrm{rec}}/\bar{\alpha}^2 \to 0$, then this estimator becomes an unbiased estimator of the full CO power spectrum on large scales:

$$\left\langle \widehat{I_{\mathrm{CO}} I_{\mathrm{CO}}}\right\rangle = \left\langle I_{\mathrm{CO}}^{\mathrm{rec}} I_{\mathrm{CO}}^{\mathrm{rec}}\right\rangle \left(\frac{\left\langle g I_{\mathrm{CO}}^{\mathrm{obs}}\right\rangle}{\left\langle g I_{\mathrm{CO}}^{\mathrm{rec}}\right\rangle}\right)^2 = \langle I_{\mathrm{CO}}\rangle^2 b_{\mathrm{CO}}^2 P(k). \tag{4.26}$$

In principle, this argument holds regardless of the luminosity limit of the CO population used for the reconstruction *and* regardless of the magnitude limit of the galaxy sample used for the cross-correlation estimate, as long as the galaxy and CO samples overlap spatially. This is potentially a powerful approach to access the entire interloper CO population without the need to identify the faint, undetected source contributions. Subtracting this CO estimate then provides an unbiased estimate of the high-redshift C II power spectrum in the same LIM dataset, which is highly desirable.

With an external galaxy catalog along, one can also estimate the total CO power on large scales where $r \approx 1$, with the estimator: $\left\langle gI_{\rm CO}^{\rm obs}\right\rangle^2/\langle gg\rangle$. However, the advantage of the Eq. 4.25 estimator is that it only requires $r \approx r_b$, which can be valid on smaller scales where $r < 1$, as long as the reconstruction has sufficient quality in terms of the noise level purity and completeness. Therefore, with the reconstructed map, one can potentially extract the total CO power to smaller scales.

We note that a high reconstruction threshold m is required in order to achieve a (nearly) noiseless reconstruction ($N^{\rm rec}/\bar{\alpha}^2 \to 0$). On the other hand, $r \approx r_b$ becomes invalid if the threshold is too high such that majority of the fainter sources is missed in the reconstruction. In addition, a high threshold tends to boost the shot noise in the reconstructed power spectrum, as well as uncertainties in the CO power spectrum estimator (Eq. 4.25). There is clearly a trade-off between the fidelity of the reconstructed signals and the uncertainty to estimate the desired signals. A detailed simulation is necessary to determine the optimal threshold to minimize the effect of bias from $N_{\rm rec}$ and the variance from the reconstruction shot noise, and evaluate the performance of the CO power spectrum estimator. We leave this investigation to future work.

Comparing Foreground Cleaning Capability with Masking

Our algorithm can serve as a foreground mitigation method for C II LIM measurement by identifying the bright CO foreground signals. In this section, we quantify this "foreground cleaning" performance using our algorithm and compare it with the masking method, where a “cleaned map” is obtained by masking out voxels that contain bright CO sources identified with an external source catalog.

In the following, we compare two cases: (1) in the masking case, the “cleaned map” is the observed map (including all of the lines and instrument noise) masked using external CO catalogs; and (2) the “cleaned map” derived from our algorithm is the observed map subtracting the CO reconstructed map.

The external catalog considered here is the same as the one described in Sec. 4.5: a flux cut on CO(5–4) at the level of 150 $L_\odot$ Mpc^{-2} (6×10^{-17} W m^{-2}), which gives ~ 1.2 galaxies per light cone (0.43^2 arcmin^2 pixel solid angle) and corresponds to a $m_{\rm AB} \sim 21.8$ threshold in the optical band. Note that this masking threshold is comparable to the "case A" masking in Sun et al. [66], although here we consider a simpler model that ignores the scatter in the line luminosity model. For this masking scenario, the "cleaned map" is the observed map minus any voxels that contain the sources in the external catalog.

In our algorithm, the CO sources and their spectra are reconstructed and removed from the data iteratively, and the residual can be regarded as a "cleaned map" that is free of bright CO sources. However, we are only capable of cleaning the multi-line redshift bins in our algorithm. The signal identified in the single-line redshift bins in our algorithm is the combination of the remaining CO and C II signals. If we remove all of the reconstructed single-line signals, we will over-subtract C II in the cleaned map. Therefore, for the single-line redshift bins ($0.15 \leq z \leq 0.51$; $0.72 \leq z \leq 0.89$; $z \geq 1.87$), we clean the data by masking voxels that contain external catalog sources.

As a figure of merit, we calculate the S/N of the respective C II shot-noise power spectra (since there is no clustering signal in our mock light cones) while including the residual CO as part of the noise contribution.

The error on the shot-noise power spectrum $\delta P_{\rm sh}$ in a map is

$$\delta P_{\rm sh} = \frac{P_{\rm tot}}{\sqrt{N_{\rm mode}}} = \frac{V_{\rm vox}\sigma_{\rm tot}^2}{\sqrt{N_{\rm mode}}}, \tag{4.27}$$

where $P_{\rm tot}$ is the total power spectrum of the cleaned map on the shot-noise scales, which is proportional to the total voxel variance in the map, $\sigma_{\rm tot}^2$. $N_{\rm mode}$ is the number of k-space modes used to measure the shot noise. $N_{\rm mode}$ is usually of the order of the total number of voxels, so we choose $N_{\rm mode} = 6000$, similar to the number of voxels in TIME. The shot-noise power of the C II signal is given by

$$P_{\rm CII} = V_{\rm vox}\sigma_{\rm CII}^2, \tag{4.28}$$

where $\sigma_{\rm CII}^2$ is the voxel variance of the C II signal map. The S/N of the C II shot-noise power spectrum is then:

$$\frac{P_{\rm CII}}{\delta P_{\rm sh}} = \frac{\sigma_{\rm CII}^2}{\sigma_{\rm tot}^2}\sqrt{N_{\rm mode}}. \tag{4.29}$$

We test the foreground cleaning performance with the same set of 2500 mock light cones described in Sec. 4.4, and calculate the C II shot-noise S/N using Eq. 4.29. The

$\sigma^2_{\rm CII}$ is the variance of the input C II map, and $\sigma^2_{\rm tot}$ is the variance of the cleaned map. For simplicity, $\sigma^2_{\rm CII}$ and $\sigma^2_{\rm tot}$ are the variance from all of the frequency channels.

In Fig. 4.21, the black dotted lines and the red lines show the C II shot-noise S/N with the cleaned map obtained from masking and from our algorithm, respectively. For reference, the black dashed lines are the C II shot-noise S/N of the observed map before cleaning, and thus the map includes the contribution from all of the lines and the instrument noise.

According to Fig. 4.7, the optimal threshold that gives the maximum r value is $m \sim 4$. In Fig. 4.21, $m = 4$ is marked with an orange dashed line, and we see that in the highest-noise case for this threshold, masking (black dotted line) using an external (deep) catalog performs slightly better than reconstruction (red line) because it is difficult for the MP algorithm to extract the signals from noisy data directly. For realistic noise levels (between $\sigma_n = 5 \times 10^3$ and 10^4 Jy sr^{-1}), the reconstruction outperforms masking.

We also compare this result with the limiting case where there are only C II and instrumental noise in the data (green line), i.e. $\sigma^2_{\rm tot} = \sigma^2_{\rm CII} + \sigma^2_n$. For a small reconstruction threshold m, the reconstructed S/N is better than this limit, which indicates that the reconstruction over-fits and misidentifies noise fluctuation as a signal and removes them from the cleaned data. We see that for the two realistic noise levels at $m \sim 4$, the reconstructed S/N is lower than this limit, and thus indicates that overfitting is not an issue at this threshold. Also note that in the highest-noise case, the masking SNR is close to the noise-plus-C II limit (green line), for the following reason: since the external catalog goes much deeper than the noise level, noise fluctuation dominates over the line signals after masking, and thus variance in the masked map is close to the noise variance.

Finally, we point out that the reconstructed C II shot-noise S/N (red line) converges to a constant instead of increasing with smaller threshold m values. This is because of the fact that in the low threshold limit, the reconstruction residual in the cleaned map is subdominant compared to the (masked) single-line redshift bin signals being added back to the residual, and thus the cleaned map S/N does not depend on the threshold value.

To sum up, for a realistic noise level, our reconstruction performs better than masking in terms of foreground cleaning capability, given our signal model and the external catalog considered in this work. We note that this conclusion depends on the line luminosity function model and the depth of the external catalog for masking, and we leave a more detailed analysis to future work.

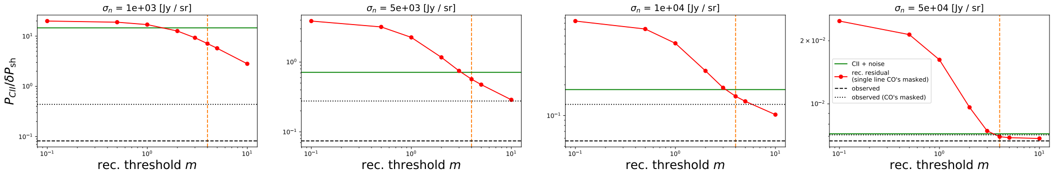

Figure 4.21: S/N on [C II] shot-noise power spectrum after CO line foreground removal with masking (black dotted) and sparse reconstruction (red) with different instrument noise level σ_n and reconstruction threshold m. The orange dashed line marks the $m = 4$ threshold for reference. We also show the S/N before cleaning (black dashed), and the limiting case of the no CO signals (green) for comparison.

Technique Extensions

In this section, we outline some directions for extending the current framework to further improve upon the line reconstruction in future work.

Template Generalization

Currently we use a single SLED template for all redshift bins in the reconstruction; this can be easily generalized to incorporate multiple spectral templates to account for the redshift evolution and SLED variation of the signals. In addition, the extension can help differentiate the emission from different types of galaxies that have different SLEDs. For instance, if we have two different SLED models for early- and late-type galaxies, respectively, we can incorporate them by having two columns in $\mathbf{A}$ for every redshift bin such that the reconstruction can infer not only the redshift and luminosity of the sources, but also their galaxy type from the SLED templates.

Alternative Sparse Approximation Algorithm

The MP algorithm adopted in this work optimizes the ℓ_0 norm in Eq. 4.8, which is the direct sum of the nonzero elements in $\widetilde{\mathbf{N}}$. We can improve the algorithm by including prior information on the expected value of each element in $\widetilde{\mathbf{N}}$, which depends on the voxel size and source luminosity function. For instance, instead of using the MP algorithm, one can obtain the sparse solution by solving the following

ℓ_1-norm regularization equation:

$$\underset{\widetilde{\mathbf{N}}}{\operatorname{argmin}} \left\| \mathbf{I} - \mathbf{A}\widetilde{\mathbf{N}} \right\|_2^2 + \lambda \left\| \mathbf{w} \cdot \widetilde{\mathbf{N}} \right\|_1 , \tag{4.30}$$

where the parameter λ determines the regularization strength for preventing over-fitting, which has a similar effect as the stopping criteria in the MP algorithm; the prior information on the number density of the sources in each redshift bin can be encoded in the weight vector $\mathbf{w}$ in this expression.

Clustering Information

In this work, we only perform the pixel-by-pixel line de-confusion using the information in the spectral correlation due to the multiple lines emitted from the same source to reconstruct the signals. The clustering information of the galaxies, which is neglected in this work, could provide additional information on the emission field. We can generalize this framework by incorporating the clustering information from the known galaxy two-point correlation, and perform the reconstruction on an ensemble of pixels to simultaneously fit for the spectral correlation and clustering. For example, if we have a theoretical model for the line-of-sight two-point correlation function of $\mathbf{N}(z)$, $\xi_{\rm th}(\mathbf{N}(z))$, we can add another ℓ_2-norm regularization term to Eq. 4.30,

$$\lambda_{\rm clus} \left\| \xi(\widetilde{\mathbf{N}}(z)) - \xi_{\rm th}(\mathbf{N}(z)) \right\|_2^2 . \tag{4.31}$$

This will enforce the algorithm to give higher priority to solutions close to the theoretical correlation function. Similarly, with an external catalog that traces the same large-scale structure, one can also constrain the algorithm with cross-correlation:

$$\lambda^x_{\rm clus} \left\| \xi^x(\widetilde{\mathbf{N}}(z), \delta_{\rm ext}) - \xi^x_{\rm th}(\mathbf{N}(z), \delta_{\rm ext}) \right\|_2^2 , \tag{4.32}$$

where $\delta_{\rm ext}$ is the density field of the external tracer, e.g. galaxy samples, and ξ^x is the cross-correlation between $\mathbf{N}(z)$ and $\delta_{\rm ext}$. We leave further investigation that makes use of clustering information to future work.

4.6 Conclusion

We develop a spectral line de-confusion technique for LIM experiments, where different spectral lines emitted by sources at different redshifts can be observed in the same frequency channel and then confused. Unlike most of the previously proposed methods that decompose the line signals in the power spectrum space, we perform a phase-space de-confusion that reconstructs the individual line intensity maps, if multiple spectral lines of a redshifted source population are observable. The reconstructed line intensity maps are direct data products of an LIM experiment, and can be used to trace the underlying density field for various science applications.

Our method is based on the information that multiple spectral lines emitted by redshifted sources are mapped onto distinct observed frequencies, which give deterministic features in the observed spectrum that can be fitted by a template. With a set of spectral template models and assuming the sparse approximation, we fit the LIM data iteratively with the MP algorithm.

As an example, we consider an LIM survey with similar survey parameters as the ongoing EoR [C II] experiments, TIME, and CONCERTO. The intervening CO line intensity maps at $0.5 \lesssim z \lesssim 1.5$ can be extracted with our technique, since multiple CO rotational transitions are observable. We demonstrate that with the assumed signal model and realistic noise level, our reconstructed CO maps reach $\sim 80\%$ spatial correlation with the true maps. In addition, in our assumed signal model and realistic survey setup, the VID of individual lines can be correctly extracted with a high S/N ratio down to the ℓ_*-scale. The CO luminosity function derived from the VID measurement can provide information on galaxy formation and evolution as traced by the CO distribution across cosmic time. The reconstruction performance is robust against a realistic level of line ratio uncertainties and continuum foreground mitigation process.

In addition to probing the large-scale luminosity and density fields, the reconstructed line intensity maps can also be used for a variety of applications. As a demonstration, we show that using the reconstructed map instead of the original LIM dataset can effectively reduce uncertainties in cross-correlation measurements, and improve the performance of interloper masking to reveal the high-redshift line emissions. Furthermore, given that the reconstructed intensity map, even if incomplete, traces the matter density on large scales, we construct an estimator capable of estimating the total interloper power. The estimator invokes the cross-correlation of the reconstructed map with an external density tracer such as galaxies in the linear clustering regime. This approach has the potential to fully specify the interloper and high-redshift source populations and warrants future investigation.

While we mainly discuss the application for an EoR [C II] LIM experiment in this paper, this technique is not restricted to this setup. We demonstrate that our technique can successfully extract redshifted optical line signals from a SPHEREx-like experiment in the near-infrared. The technique is a general framework that can be readily applied to mitigate line de-confusion problems in LIM experiments and enhance the science returns.

In this section, we describe the detailed steps in the matching pursuit (MP) algorithm. The MP algorithm iteratively selects an atom in the dictionary to project out part of the signals in the data, and keep track of the current solution of the signal **f** and residual **R** for the next step until the solution meets the stopping criteria. In Eq. 4.8, for a given signal **I** and matrix **A**, we define a set of vectors $\{\psi_i\}$ to be the column vectors in **A** (i.e. the atoms in the dictionary). The MP algorithm works as follows:

1. Initialize at step $t = 0$: $\mathbf{f}_0 = \mathbf{0}$, $\mathbf{R}_0 = \mathbf{I}$, $\widetilde{\mathbf{N}}_0 = \mathbf{0}$.
2. Compute the inner product of $\mathbf{R}_0$ and ψ_i's:
$$\{u_{0i}\} = \{\langle \mathbf{R}_0, \psi_i \rangle\}\,. \tag{33}$$
3. Select the element γ to be updated by
$$\gamma = \underset{i}{\operatorname{argmax}}\,\{u_{0i}\}\,. \tag{34}$$
4. If $u_{0\gamma}$ meets the stopping criteria, end the process and return $\mathbf{N} = \mathbf{0}$. Otherwise, proceed to step 5.
5. Update the current $\mathbf{f}_t$, $\mathbf{R}_t$, and record the amplitude of the new solution in γ-th element of vector $\widetilde{\mathbf{N}}_t$:
$$\mathbf{f}_{t+1} = \mathbf{f}_t + u_{t\gamma}\,\psi_\gamma, \tag{35}$$
$$\mathbf{R}_{t+1} = \mathbf{R}_t - u_{t\gamma}\,\psi_\gamma, \tag{36}$$
$$\widetilde{\mathbf{N}}_{t+1}(\gamma) = \widetilde{\mathbf{N}}_t(\gamma) + u_{t\gamma}. \tag{37}$$
6. Compute the inner product of $\mathbf{R}_{t+1}$ and ψ_i's:
$$\left\{u_{(t+1)i}\right\} = \{\langle \mathbf{R}_{t+1}, \psi_i \rangle\}\,. \tag{38}$$
7. Select the element γ to be updated by
$$\gamma = \underset{i}{\operatorname{argmax}}\left\{u_{(t+1)i}\right\}. \tag{39}$$
8. If $u_{(t+1)\gamma}$ meets the stopping criteria, go back to step 5 for the next iteration. Else, proceed to step 9.

9. Return the final solution $\mathbf{N} = \widetilde{\mathbf{N}}_{t+1} \,/\, \mathbf{I}^{\rm norm}$.

As described in Sec. 4.3, the stopping criteria is set by comparing $u_{t\gamma}$ with the noise σ_n. This follows the fact that $\mathrm{var}(u_{t\gamma}) = \sigma_n^2$ (see Appendix .2 for the proof), so if we set $u_{(t+1)\gamma} < m\,\sigma_n$ in step 8, this is effectively setting an "m-σ" detection threshold (e.g., $m = 5$ for a 5σ detection).

Fig. .22 illustrates the steps of the MP algorithm solving **N** of an example light cone. In this example, we set $\sigma_n = 10$ kJy, and the detection threshold $m = 5$. In this light cone, there are six ℓ_* sources in the multi-line redshift bins at z = [0.54, 1.06, 1.20, 1.24, 1.79, 1.82], shown in the top left panel. Since **N** is the effective number of ℓ_* sources per redshift bin, the amplitudes in the six corresponding redshift bins are equal to unity. The top middle panel is the total line signal in this light cone $\mathbf{I}_{\rm true}$, including the emission from those six sources as well as that from other sources in the single-line redshifts. The top right panel shows the observed data after adding noise to the signal $\mathbf{I}_{\rm true}$, which is also the $\mathbf{R}_0$ vector in the first step of the MP algorithm. The blue dashed lines mark the noise level $\pm\sigma_n = 10$ kJy for reference.

.1 The Matching Pursuit (MP) Algorithm

In the first iteration, the MP algorithm selects the 14th z-bin index ($z = 0.54$) with amplitude $u_{1\gamma} \sim 1$, so the $z = 0.54$ source is successfully extracted in this iteration. The gray dashed spectrum in the left panel of the second row is the template signal extracted in this step ($u_{1\gamma}\psi_\gamma$). The updated values of $\mathbf{N}_1$, $\mathbf{f}_1$, and $\mathbf{R}_1$ from step 5 are shown in the second row of Fig. .22.

Then we proceed to the second iteration. The MP algorithm selects the 37th column ($z = 0.65$). The gray dashed spectrum in the left panel of the third row is the template signal fitted in this step ($u_{2\gamma}\psi_\gamma$). However, there is no $z = 0.65$ source in the input, which means the MP algorithm misidentifies the emission from noise or other lines as the signal. The third row of Fig. .22 shows the updated values of $\mathbf{N}_2$, $\mathbf{f}_2$, and $\mathbf{R}_2$ from the second iteration.

In the third iteration, the 109th ($z = 1.24$) column in the dictionary is selected. The fourth row of Fig. .22 shows the updated values of $\mathbf{N}_3$, $\mathbf{f}_3$, and $\mathbf{R}_3$ from this iteration. Note that this time the MP algorithm picks up a correct redshift, while it overestimates the amplitude by $\sim 40\%$.

In the fourth iteration, the algorithm selects the 224th column from the dictionary. The 224th column is not in the multi-line redshift bins (first 195 columns), and thus

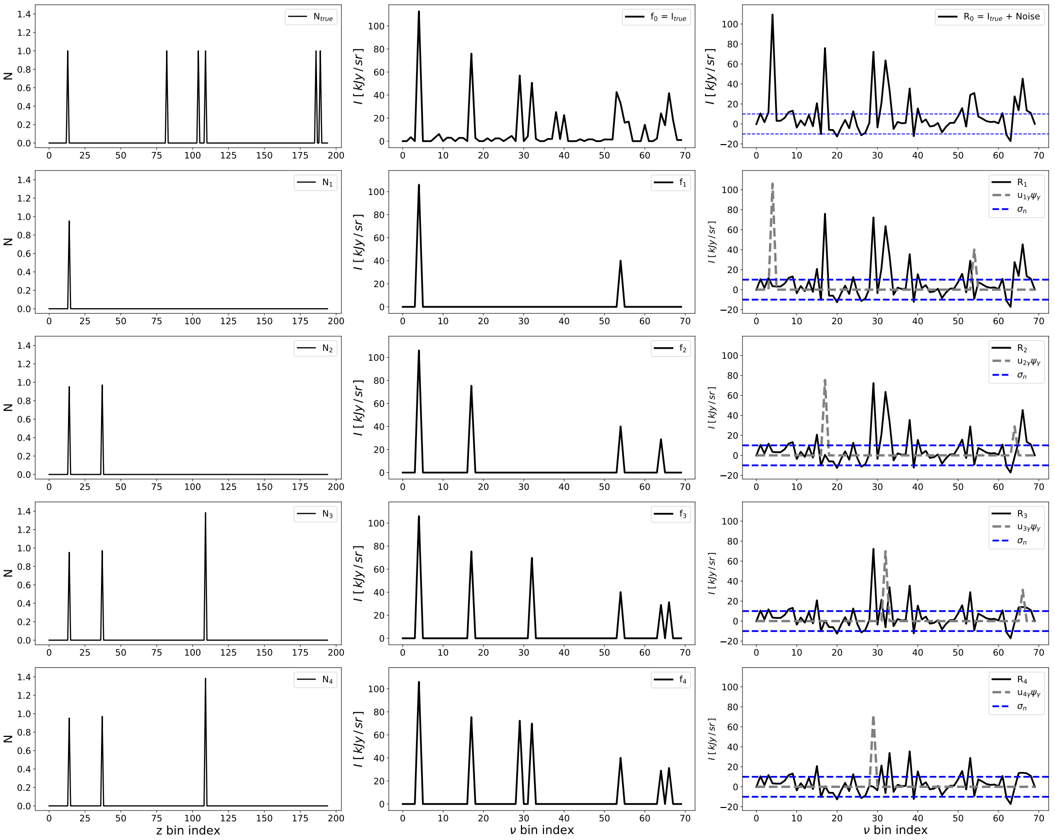

Figure .22: Illustration of how the MP algorithm solves for the source vector **N**. See the text for a detailed description.

the $\mathbf{N}_4$ in the bottom row of Fig. .22 remains unchanged, whereas $\mathbf{f}_4$ and $\mathbf{R}_4$ are updated with a single-peak signal.

In the fifth iteration, the stopping criteria in step 8 is met ($u_{5\gamma} < 5\sigma_n$), so the reconstruction terminates and returns the last row of Fig. .22 as the reconstruction results for this light cone.

In summary, in this example, two of the six multi-line redshift sources have been reconstructed, in addition to one misidentified source. Comparing the final reconstructed light cone signal (bottom middle panel of Fig. .22) to the true input light cone (top middle panel of Fig. .22), we can see that the MP reconstruction captures the strong peaks in the data, and the remaining signals are close to the noise level.

.2 Proving $\mathrm{var}(u_{t\gamma}) = \sigma_n^2$

In Sec. .1, the residual of step t $\mathbf{R_t}$ can be expressed in the linear combination of the dictionary atoms and noise:

$$\mathbf{R}_t = \sum_i c_i \psi_i + \mathbf{n}, \tag{40}$$

where c_i's represents the constant coefficient. Then we derive

$$\begin{aligned} u_{t\gamma} &\equiv \langle\, \mathbf{R}_t, \psi_\gamma \rangle \\ &= \sum_i c_i \langle\, \psi_i, \psi_\gamma \rangle + \langle\, \mathbf{n}, \psi_\gamma \rangle \\ &= \mathrm{const} + \sum_j n_j \, \psi_{\gamma j}. \end{aligned} \tag{41}$$

The first term does not depend on the noise, so it is a constant term that is not contributing to the variance. Also note that $\langle\, \psi_i, \psi_\gamma \rangle \neq \delta_{i\gamma}$ since the dictionary $\{\psi_i\}$ is normalized but not orthogonal. With this expression, we can calculate the variance:

$$\begin{aligned} \left\langle u_{t\gamma} \right\rangle &= \mathrm{const} + \sum_j \left\langle n_j \right\rangle \, \psi_{\gamma j} = \mathrm{const} \\ \left\langle u_{t\gamma}^2 \right\rangle &= \mathrm{const}^2 + \sum_j \left\langle n_j^2 \right\rangle \, \psi_{\gamma j}^2 \\ &= \mathrm{const}^2 + \sigma_n^2 \sum_j \psi_{\gamma j}^2 \\ &= \mathrm{const}^2 + \sigma_n^2. \end{aligned} \tag{42}$$

Therefore, we get

$$\mathrm{var}(u_{t\gamma}) = \left\langle u_{t\gamma}^2 \right\rangle - \left\langle u_{t\gamma} \right\rangle^2 = \sigma_n^2. \tag{43}$$

.3 SPHEREx Line Signal Model

In this section, we describe the line signal model in the SPHEREx wavelengths. We model five lines from $z = 0$ to 10 in SPHEREx band: Lyα (121.6 nm), Hα (656.3 nm), Hβ (486.1 nm), [O II] (372.7 nm), and [O III] (500.7 nm).

Since the optical lines are associated with the star formation activities, we model the signal with the following steps: we start with the halo mass function, and use the star formation rate (SFR)—halo mass (M) relation, and the SFR—line luminosity relation to paint the spectral line signals to each halo.

We use the publicly available halo mass function calculator *HMFcalc* [52][8] to obtain the halo mass function based on the Sheth et al. [62] model. For the SFR–M relation,

---

[8] `http://hmf.icrar.org/`.

we use the model from Behroozi et al. [1], in which the SFR–M relation is derived based on several observational constraints. [9]

For the SFR–line luminosity relation, we assume a linear relation for all of the lines. For Lyα, we use the prescription provided by Fonseca et al. [26] with their fiducial values: $\gamma_{\mathrm{Ly}\alpha} = 1$, $f_{\mathrm{esc}}^{UV} = 0.2$, $f_{\mathrm{esc}}^{\mathrm{Ly}\alpha} = 0.2$, $E_{\mathrm{UV}} = 1.0$ in their equation 8 and 15, and derive the conversion factor:

$$\frac{SFR}{M_{\odot}/\mathrm{yr}} = 2.29 \times 10^{-41} \frac{L_{\mathrm{Ly}\alpha}}{\mathrm{erg/s}}. \tag{44}$$

For other spectral lines, we adopt the relation from Kennicutt [36] and Ly et al. [44]:

$$\frac{SFR}{M_{\odot}/\mathrm{yr}} = (7.9 \pm 2.4) \times 10^{-42} \frac{L_{H\alpha}}{\mathrm{erg/s}}, \tag{45}$$

$$\frac{SFR}{M_{\odot}/\mathrm{yr}} = (1.4 \pm 0.4) \times 10^{-41} \frac{L_{[OII]}}{\mathrm{erg/s}}, \tag{46}$$

$$\frac{SFR}{M_{\odot}/\mathrm{yr}} = (7.6 \pm 3.7) \times 10^{-42} \frac{L_{[OIII]}}{\mathrm{erg/s}}, \tag{47}$$

and for the Hβ line, we use the fixed line ratio $H\beta/H\alpha = 0.35$ [53].

References

[1] Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013, ApJ, 770, 57, doi: 10.1088/0004-637X/770/1/57

[2] Bowman, J. D., Morales, M. F., & Hewitt, J. N. 2009, ApJ, 695, 183, doi: 10.1088/0004-637X/695/1/183

[3] Breysse, P. C., Kovetz, E. D., Behroozi, P. S., Dai, L., & Kamionkowski, M. 2017, MNRAS, 467, 2996, doi: 10.1093/mnras/stx203

[4] Breysse, P. C., Kovetz, E. D., & Kamionkowski, M. 2014, MNRAS, 443, 3506, doi: 10.1093/mnras/stu1312

[5] —. 2015, MNRAS, 452, 3408, doi: 10.1093/mnras/stv1476

[6] —. 2016, MNRAS, 457, L127, doi: 10.1093/mnrasl/slw005

[7] Breysse, P. C., & Rahman, M. 2017, MNRAS, 468, 741, doi: 10.1093/mnras/stx451

[8] Candes, E. J., Romberg, J., & Tao, T. 2006, IEEE Trans. Inf. Theor., 52, 489, doi: 10.1109/TIT.2005.862083

---

[9] The $SFR(M, z)$ is downloaded from the author's webpage (`https://www.peterbehroozi.com/data.html`). The model is only available at $0 < z < 8$, so we use $z = 8$ model for $z > 8$.

[9] Carilli, C. L. 2011, ApJL, 730, L30, doi: 10.1088/2041-8205/730/2/L30

[10] Chang, T. C., Gong, Y., Santos, M., et al. 2015, Advancing Astrophysics with the Square Kilometre Array (AASKA14), 4. https://arxiv.org/abs/1501.04654

[11] Chang, T.-C., Pen, U.-L., Bandura, K., & Peterson, J. B. 2010, Nature, 466, 463, doi: 10.1038/nature09187

[12] Chang, T.-C., Pen, U.-L., Peterson, J. B., & McDonald, P. 2008, Physical Review Letters, 100, 091303, doi: 10.1103/PhysRevLett.100.091303

[13] Chapman, E., Abdalla, F. B., Harker, G., et al. 2012, MNRAS, 423, 2518, doi: 10.1111/j.1365-2966.2012.21065.x

[14] Cheng, Y.-T., Chang, T.-C., Bock, J., Bradford, C. M., & Cooray, A. 2016, ApJ, 832, 165, doi: 10.3847/0004-637X/832/2/165

[15] Cheng, Y.-T., de Putter, R., Chang, T.-C., & Doré, O. 2019, ApJ, 877, 86, doi: 10.3847/1538-4357/ab1b2b

[16] Chung, D. T., Viero, M. P., Church, S. E., et al. 2019, ApJ, 872, 186, doi: 10.3847/1538-4357/ab0027

[17] Comaschi, P., & Ferrara, A. 2016, MNRAS, 455, 725, doi: 10.1093/mnras/stv2339

[18] Crites, A. T., Bock, J. J., Bradford, C. M., et al. 2014, in Proc. SPIE, Vol. 9153, Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy VII, 91531W, doi: 10.1117/12.2057207

[19] Croft, R. A. C., Miralda-Escudé, J., Zheng, Z., Blomqvist, M., & Pieri, M. 2018, MNRAS, 481, 1320, doi: 10.1093/mnras/sty2302

[20] Croft, R. A. C., Miralda-Escudé, J., Zheng, Z., et al. 2016, MNRAS, 457, 3541, doi: 10.1093/mnras/stw204

[21] Daddi, E., Dannerbauer, H., Liu, D., et al. 2015, A&A, 577, A46, doi: 10.1051/0004-6361/201425043

[22] de Putter, R., Holder, G. P., Chang, T.-C., & Dore, O. 2014, ArXiv e-prints. https://arxiv.org/abs/1403.3727

[23] Decarli, R., Walter, F., Gónzalez-López, J., et al. 2019, ApJ, 882, 138, doi: 10.3847/1538-4357/ab30fe

[24] Donoho, D. L. 2006, IEEE Trans. Inf. Theor., 52, 1289, doi: 10.1109/TIT.2006.871582

[25] Doré, O., Bock, J., Ashby, M., et al. 2014, ArXiv e-prints. https://arxiv.org/abs/1412.4872

[26] Fonseca, J., Silva, M. B., Santos, M. G., & Cooray, A. 2017, MNRAS, 464, 1948, doi: 10.1093/mnras/stw2470

[27] Furlanetto, S. R., Oh, S. P., & Briggs, F. H. 2006, PhysRep, 433, 181, doi: 10.1016/j.physrep.2006.08.002

[28] Gong, Y., Chen, X., & Cooray, A. 2020, ApJ, 894, 152, doi: 10.3847/1538-4357/ab87a0

[29] Gong, Y., Cooray, A., Silva, M., et al. 2012, ApJ, 745, 49, doi: 10.1088/0004-637X/745/1/49

[30] Gong, Y., Cooray, A., Silva, M. B., Santos, M. G., & Lubin, P. 2011, ApJL, 728, L46, doi: 10.1088/2041-8205/728/2/L46

[31] Gong, Y., Silva, M., Cooray, A., & Santos, M. G. 2014, ApJ, 785, 72, doi: 10.1088/0004-637X/785/1/72

[32] Helgason, K., Ricotti, M., & Kashlinsky, A. 2012, ApJ, 752, 113, doi: 10.1088/0004-637X/752/2/113

[33] Ihle, H. T., Chung, D., Stein, G., et al. 2019, ApJ, 871, 75, doi: 10.3847/1538-4357/aaf4bc

[34] Keating, G. K., Marrone, D. P., Bower, G. C., et al. 2016, ApJ, 830, 34, doi: 10.3847/0004-637X/830/1/34

[35] Keating, G. K., Bower, G. C., Marrone, D. P., et al. 2015, ApJ, 814, 140, doi: 10.1088/0004-637X/814/2/140

[36] Kennicutt, Robert C., J. 1998, Annual Review of Astronomy and Astrophysics, 36, 189, doi: 10.1146/annurev.astro.36.1.189

[37] Kogut, A., Dwek, E., & Moseley, S. H. 2015, ApJ, 806, 234, doi: 10.1088/0004-637X/806/2/234

[38] Lagache, G., Cousin, M., & Chatzikos, M. 2018, A&A, 609, A130, doi: 10.1051/0004-6361/201732019

[39] Li, T. Y., Wechsler, R. H., Devaraj, K., & Church, S. E. 2016, ApJ, 817, 169, doi: 10.3847/0004-637X/817/2/169

[40] Lidz, A., Furlanetto, S. R., Oh, S. P., et al. 2011, ApJ, 741, 70, doi: 10.1088/0004-637X/741/2/70

[41] Lidz, A., & Taylor, J. 2016, ApJ, 825, 143, doi: 10.3847/0004-637X/825/2/143

[42] Lidz, A., Zahn, O., Furlanetto, S. R., et al. 2009, ApJ, 690, 252, doi: 10.1088/0004-637X/690/1/252

[43] Liu, A., & Tegmark, M. 2012, MNRAS, 419, 3491, doi: `10.1111/j.1365-2966.2011.19989.x`

[44] Ly, C., Malkan, M. A., Kashikawa, N., et al. 2007, ApJ, 657, 738, doi: `10.1086/510828`

[45] Madau, P., Meiksin, A., & Rees, M. J. 1997, ApJ, 475, 429

[46] Mallat, S. G., & Zhang, Z. 1993, IEEE Transactions on Signal Processing, 41, 3397, doi: `10.1109/78.258082`

[47] Mashian, N., Sternberg, A., & Loeb, A. 2015, JCAP, 11, 028, doi: `10.1088/1475-7516/2015/11/028`

[48] Masui, K. W., Switzer, E. R., Banavar, N., et al. 2013, ApJL, 763, L20, doi: `10.1088/2041-8205/763/1/L20`

[49] Morales, M. F., Bowman, J. D., & Hewitt, J. N. 2006, ApJ, 648, 767, doi: `10.1086/506135`

[50] Moriwaki, K., Filippova, N., Shirasaki, M., & Yoshida, N. 2020, MNRAS, 496, L54, doi: `10.1093/mnrasl/slaa088`

[51] Moustakas, J., Kennicutt, Robert C., J., & Tremonti, C. A. 2006, ApJ, 642, 775, doi: `10.1086/500964`

[52] Murray, S. G., Power, C., & Robotham, A. S. G. 2013, Astronomy and Computing, 3, 23, doi: `10.1016/j.ascom.2013.11.001`

[53] Osterbrock, D. E., & Ferland, G. J. 2006, Astrophysics of gaseous nebulae and active galactic nuclei

[54] Parsons, A. R., Pober, J. C., Aguirre, J. E., et al. 2012, ApJ, 756, 165, doi: `10.1088/0004-637X/756/2/165`

[55] Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 594, A13, doi: `10.1051/0004-6361/201525830`

[56] Popping, G., van Kampen, E., Decarli, R., et al. 2016, MNRAS, 461, 93, doi: `10.1093/mnras/stw1323`

[57] Pullen, A. R., Chang, T.-C., Doré, O., & Lidz, A. 2013, ApJ, 768, 15, doi: `10.1088/0004-637X/768/1/15`

[58] Pullen, A. R., Doré, O., & Bock, J. 2014, ApJ, 786, 111, doi: `10.1088/0004-637X/786/2/111`

[59] Riechers, D. A., Pavesi, R., Sharon, C. E., et al. 2019, ApJ, 872, 7, doi: `10.3847/1538-4357/aafc27`

[60] Righi, M., Hernández-Monteagudo, C., & Sunyaev, R. A. 2008, A&A, 489, 489, doi: 10.1051/0004-6361:200810199

[61] Scott, D., & Rees, M. J. 1990, MNRAS, 247, 510

[62] Sheth, R. K., Mo, H. J., & Tormen, G. 2001, MNRAS, 323, 1, doi: 10.1046/j.1365-8711.2001.04006.x

[63] Silva, M., Santos, M. G., Cooray, A., & Gong, Y. 2015, ApJ, 806, 209, doi: 10.1088/0004-637X/806/2/209

[64] Silva, M. B., Santos, M. G., Gong, Y., Cooray, A., & Bock, J. 2013, ApJ, 763, 132, doi: 10.1088/0004-637X/763/2/132

[65] Stompor, R., Balbi, A., Borrill, J. D., et al. 2002, PhRvD, 65, 022003, doi: 10.1103/PhysRevD.65.022003

[66] Sun, G., Moncelsi, L., Viero, M. P., et al. 2018, ApJ, 856, 107, doi: 10.3847/1538-4357/aab3e3

[67] Switzer, E. R., Chang, T.-C., Masui, K. W., Pen, U.-L., & Voytek, T. C. 2015, ApJ, 815, 51, doi: 10.1088/0004-637X/815/1/51

[68] Uzgil, B. D., Aguirre, J. E., Bradford, C. M., & Lidz, A. 2014, ApJ, 793, 116, doi: 10.1088/0004-637X/793/2/116

[69] Visbal, E., & Loeb, A. 2010, JCAP, 11, 016, doi: 10.1088/1475-7516/2010/11/016

[70] Walter, F., Decarli, R., Sargent, M., et al. 2014, ApJ, 782, 79, doi: 10.1088/0004-637X/782/2/79

[71] Wyithe, J. S. B., & Loeb, A. 2008, MNRAS, 383, 606, doi: 10.1111/j.1365-2966.2007.12568.x

[72] Yue, B., Ferrara, A., Pallottini, A., Gallerani, S., & Vallini, L. 2015, MNRAS, 450, 3829, doi: 10.1093/mnras/stv933

Chapter 5

PROBING INTRA-HALO LIGHT WITH GALAXY STACKING IN CIBER IMAGES

Cheng, Y.-T., Arai, T., Bangale, P., et al. 2021, arXiv e-prints, arXiv:2103.03882. https://arxiv.org/abs/2103.03882

We study the stellar halos of $0.2 \lesssim z \lesssim 0.5$ galaxies with stellar masses spanning $M_* \sim 10^{10.5}$ to $10^{12} M_\odot$ (approximately L_* galaxies at this redshift) using imaging data from the Cosmic Infrared Background Experiment (CIBER). A previous CIBER fluctuation analysis suggested that intra-halo light (IHL) contributes a significant portion of the near-infrared extragalactic background light (EBL), the integrated emission from all sources throughout cosmic history. In this work, we carry out a stacking analysis with a sample of ~30,000 Sloan Digital Sky Survey (SDSS) photometric galaxies from CIBER images in two near-infrared bands (1.1 and 1.8 μm) to directly probe the IHL associated with these galaxies. We stack galaxies in five sub-samples split by brightness, and detect an extended galaxy profile, beyond the instrument point spread function (PSF), derived by stacking stars. We jointly fit a model for the inherent galaxy light profile, plus large-scale one- and two-halo clustering to measure the extended galaxy IHL. We detect non-linear one-halo clustering in the 1.8 μm band, at a level consistent with numerical simulations. By extrapolating the fraction of extended galaxy light we measure to all galaxy masses scales, we find ~ 30%/15% of the total galaxy light budget from galaxies is at radius $r > 10/20$ kpc, respectively. These results are new at near-infrared wavelengths at the L_* mass scale, and suggest that the IHL emission and one-halo clustering could have appreciable contributions to the amplitude of large-scale EBL background fluctuations.

5.1 Introduction

In the standard cosmological paradigm, galaxies grow hierarchically through merger and accretion. Galaxies accreting onto more massive systems become disrupted, and stars stripped away from their parent galaxies become redistributed in the merged dark matter halo. This results in extended stellar halos that are known to span tens or hundreds of kilo-parsecs. The stellar emission from this material is sometimes referred to as “intra-halo light” (IHL), or in massive galaxy clusters as “intra-cluster light” (ICL).

The properties of stellar halos across a wide range of mass scales have been extensively studied using analytical models [e.g., 77] and N-body simulations [e.g., 17, 23–26, 35, 80, 83]. Several observations have constrained the ICL content in galaxy clusters [e.g., 18, 43, 43, 44, 63], as well as stellar halos in lower mass systems by deeply imaging individual galaxies [e.g., 5, 51, 66, 92, 96] or through stacking [e.g., 34, 97, 102, 103].

An independent way to study the aggregate emission from diffuse sources like IHL is through measurements of the extragalactic background light (EBL), which encodes the integrated emission from all sources across cosmic history [27]. Absolute optical and near-infrared EBL photometry has proven challenging as measurements must tightly control systematic errors and carefully model and subtract local foregrounds [e.g., 55, 61, 69, 71, 99]. Several authors [13, 61, 62, 67, 71, 86, 87, 94, 99] have reported potential detections above the integrated galaxy light (IGL) derived from galaxy counts [32, 33, 48, 56, 59, 85], which may indicate the existence of extragalactic emission missed in source counting surveys.

Additionally, EBL fluctuation analyses have also consistently reported excess fluctuations over those expected from the IGL [29, 53, 54, 57, 68, 70, 73, 88, 93, 101]. One explanation is emission from the epoch of reionization [53, 54, 70, 73], while other studies suggest IHL contributes most of the excess fluctuations [29]. In particular, [101] interpret imaging data from the *Cosmic Infrared Background Experiment* (CIBER) as arising from an IHL intensity comparable to the IGL at near-infrared wavelengths. This result would imply that stars diffusely scattered in dark matter halos may account for a non-negligible fraction of the near-IR cosmic radiation budget. The absorption spectra from blazars constrain the EBL column density along the line of sight [e.g., 1, 2, 4, 6–10, 46, 65]. While IHL is generally produced at low redshifts, improving the uncertainties in its redshift history helps place IHL in the context of these constraints.

In this work, we further constrain the IHL using CIBER broad band imaging. Rather than studying EBL intensity fluctuations as in Zemcov et al. [101], we perform a stacking analysis to directly probe the stellar halos around galaxies. We stack a sample of $\sim 30,000$ Sloan Digital Sky Survey (SDSS) photometric galaxies at $z \sim 0.2 - 0.5$ across five 2×2 deg^2 fields. Our samples span a range of stellar masses at approximately L_* scales at this redshift [74]. Although we only study stellar halos around a subset of galaxies, rather than the aggregate population as probed by fluctuations, stacking provides a direct path to probe the IHL associated with this sample. Stacking complements fluctuation measurements by probing the relationship between individual galaxies and their stellar halos. Stacking also allows us to investigate how stellar halos depend on host galaxy properties, e.g, stellar mass, redshift, etc. A complementary fluctuation analysis of these same data is currently in progress.

This paper is organized as follows. First, we introduce CIBER in Sec. 5.2 and the data processing in Sec. 5.3. Sec. 5.4 and 5.5 describe the external data sets used in this work, including observed and simulated source catalogs. Sec. 5.6 details the stacking procedure, and Sec. 5.7 describes the point spread function (PSF) model. The stacking results are presented in Sec. 5.8. Sec. 5.9 introduces the theoretical model we use to fit the data, and the parameter fitting procedure. The results on model parameter constraints are given in Sec. 5.10, and further discussion is presented in Sec. 5.11. Sec. 5.12 summarizes the paper. Throughout this work, we assume a flat ΛCDM cosmology with $n_s = 0.97$, $\sigma_8 = 0.82$, $\Omega_m = 0.26$, $\Omega_b = 0.049$, $\Omega_\Lambda = 0.69$, and $h = 0.68$, consistent with the measurement from *Planck* [76]. All fluxes are quoted in the AB magnitude system.

5.2 CIBER Experiment

CIBER[1] [100] is a rocket-borne instrument designed to characterize the near-infrared EBL. CIBER consists of four instruments: two wide-field imagers [15], a narrow-band spectrometer [58], and a low-resolution spectrometer [95]. CIBER has flown four times in February 2009, July 2010, March 2012, and June 2013. The first three CIBER flights were launched at White Sands Missile Range, New Mexico on a Terrier-Black Brant IX rocket. These flights reached $\sim$ 330 km apogee with $\sim$ 240 sec of exposure time, and the payload was recovered for future flights. The fourth flight was a non-recovery flight launched 3:05 UTC 2013 June 6 from Wallops Flight Facility, Virginia on a four-stage Black Brant XII rocket. The payload reached 550 km altitude, much higher than the two-stage rocket used in the previous three flights. This gives more exposure time (335 sec) for observing more science fields with long integrations to achieve better sensitivity and systematics control.

This work presents the first science results from the CIBER fourth flight imager data. The data from previous flights have been studied with a fluctuation analysis,

---

[1] `https://ciberrocket.github.io/`

Table 5.1: CIBER Observing Fields

Field Name	R.A. (°)	Dec. (°)	Time After Launch (sec)	Number of Frames Used	Integration Time (sec)[a]
Elat10	191.50	8.25	387-436	24	42.72
Elat30	193.94	28.00	450-500	9	16.02
BootesB	218.11	33.18	513-569	29	51.62
BootesA	219.25	34.83	581-636	28	49.84
SWIRE(ELAIS-N1)	241.53	54.77	655-705	25	44.50

---

[a] We discard the beginning part of Elat30 field integration due to pointing instability.

published in [101]. With a large field of view and low sky background above the atmosphere, CIBER imaging provides fidelity on angular scales from $7''$ to 2°. For stacking, CIBER imaging can trace low surface brightness emission on degree angular scales providing a unique dataset compared with ground-based or small field-of-view space-borne studies. Each CIBER imager uses a 1024×1024 pixel HAWAII-1 HgCdTe detector. The two imagers are identical except for their $\lambda/\Delta\lambda \sim 2$ filters, which are centered at 1.05 and 1.79 μm[2].

During its fourth flight, CIBER observed eight science fields with ~ 50 sec integrations sampled at 1.78 sec intervals. We discard the first three fields in this analysis due to contamination from airglow that produces a strong non-uniform emission across the images that requires aggressive filtering which also significantly reduces our signal [101]. Table 5.1 summarizes the sky coordinates and the integration time of the five science fields used in this work. In the beginning of the Elat30 integration, the rocket's pointing was not stable which has the effect of smearing the PSF on the sky. As a result, we only use the last 16 sec of this integration in our analysis.

5.3 Data Processing

In this section, we describe the data reduction from the raw flight data to the final images used for stacking.

Raw Time Stream to Images

The raw imager data provides a time series for each pixel. We fit a slope to the time stream to obtain the photocurrent in each pixel, and convert the values from the raw analog-to-digital units (ADU) to e^{-} s^{-1} using known array gain factors.

The HAWAII-1 detector is linearly responsive to incoming flux over a certain dynamic range. For pixels pointing at bright sources, the detectors saturate and have a non-linear flux dependence, even for short integrations [15]. In any pixel that collects more than 5,000 ADU over the full integration only the first four frames are used in the photocurrent estimate. Hereafter, the term "raw image" refers to the photocurrent map after this linearity correction. Panel A of Fig. 5.1 and Fig. 5.2 show the raw images of the SWIRE field in the CIBER 1.1 and 1.8 μm bands, respectively.

---

[2] In the first and second CIBER flights, the longer wavelength band is centered at 1.56 μm, and thus it is named 1.6 μm band in previous CIBER publications [15, 101].

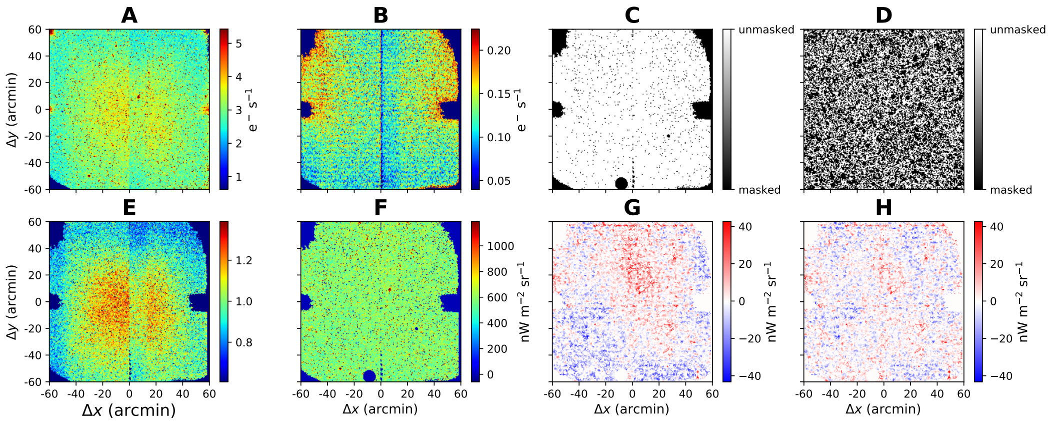

Figure 5.1: Images from the SWIRE field in the 1.1 μm band. A: the raw image of the photoccurent map. B: dark current template constructed from dark images before the flight. C: instrument mask encoding the pixels with fabrication defects, unusual photocurrents, and cosmic ray contamination. D: source mask for bright stars and galaxies in the 2MASS and Pan-STARRS catalogs. E: flat-field estimator from averaging the other four sky fields. F: raw image after dark current subtraction, flat field correction, and calibration. G: Image in Panel F after (constant) background removal and masking. This image is smoothed with a $\sigma = 35''$ Gaussian kernel to highlight large-scale fluctuations. H: Image in Panel G after subtracting a fitted 2-D polynomial, also shown smoothed with a $\sigma = 35''$ Gaussian kernel. Compared to Panel G, we see that the large-scale background fluctuations have been reduced after filtering. This is the final product of the data reduction pipeline.

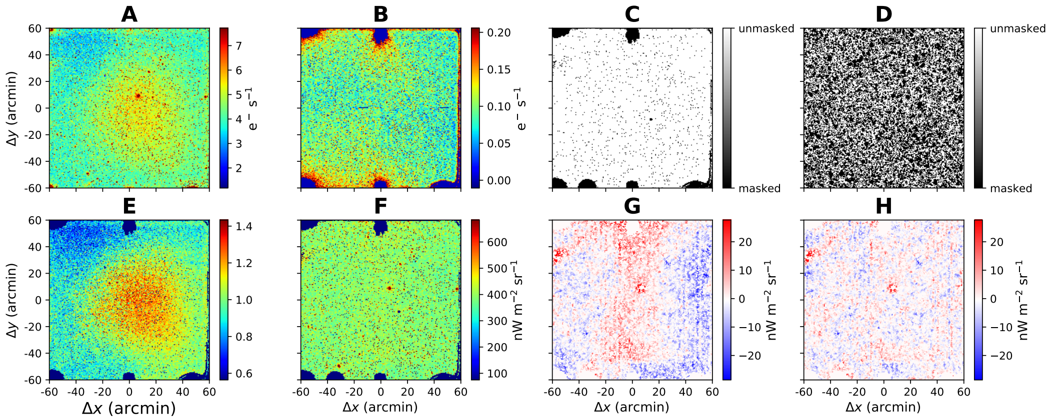

Figure 5.2: Same as Fig. 5.1 in the CIBER 1.8 μm band.

Dark Current

In the absence of incoming photons, the detectors have a nonzero response, commonly referred to as "dark current", due to thermally produced charge carriers and multiplexer glow. The detector dark current is measured before each flight with the telescopes' cold shutters closed. We obtain a dark current template for each detector by averaging 11 dark images and then subtracting each template from the corresponding raw images. The dark current level in CIBER imagers is $\sim$ 0.1 e^- s^{-1}, less than 10 % of the sky brightness. Panel B of Fig. 5.1 and Fig. 5.2 show the dark current maps of CIBER 1.1 and 1.8 μm bands, respectively.

Pixel Masks

We mask pixels that meet at least one of the following conditions: (1) a fabrication defect; (2) poor time-stream behavior; (3) abnormal photocurrents compared with other pixels; (4) a cosmic ray strike; or (5) being on or close to bright point sources on the sky. The pixels satisfying criteria (1)–(4) comprise the "instrument mask", and a "source mask" is composed of pixels with condition (5).

Instrument Mask

Pixels with fabrication defects and significant multiplexer glow are mostly distributed near the edges or corners of each quadrant on the detector arrays. They exhibit pathologies in their photocurrent response, and can be found by comparison to the population of normal pixels. We perform a 3-σ clipping on stacked dark images (the same dataset used for a dark current template in Sec. 5.3) to identify these pixels.

During integration, some cosmic ray events or electronic transients leave a step feature in the time stream. We use a 100-σ clip on each time stream to pick out pixels that show these abrupt changes during an integration. Sometimes cosmic ray events also leave a comet-like structure on the array, and these regions are also masked. The union of the pathological pixel, time-stream masks, and cosmic ray masks form the instrument mask. In total, $\sim$ 10% of pixels are removed by the instrument mask. Panel C of Fig. 5.1 and Fig. 5.2 show the instrument masks in the SWIRE field of 1.1 and 1.8 μm band, respectively.

Source Mask

To remove bright foreground stars and galaxies in our fields, we use position and brightness information from the Pan-STARRS and 2MASS catalogs (see Sec. 5.4 for details). We further derive source magnitudes in the two CIBER bands, $m_{1.1}$ and $m_{1.8}$, from these catalogs, as detailed in Sec. 5.4. We mask all point sources brighter than $m_{1.1} = 20$, choosing a masking radius for each source derived as follows. With the modeled instrument PSF (Sec. 5.7), the masking radius is chosen such that for each source, pixels with intensity brighter than νI_ν^{th} = 1 nW m^{-2} sr^{-1} in the 1.1 μm band are masked. This choice of threshold value removes ~ 50% of pixels in each field. We apply the same masking radius to 1.8 μm band sources. The same masking function is also applied to simulations to account for residual emission from bright sources outside the masks and the unmasked faint populations. Panel D of Fig. 5.1 and Fig. 5.2 show the SWIRE field source mask in the CIBER 1.1 and 1.8 μm bands, respectively.

The final mask we apply to the data is the union of the instrument mask and source mask. After applying these masks, we apply a final 3-σ pixel clipping mask to identify additional outliers not flagged through the other methods (e.g., from low-energy cosmic ray events or electronic tranisents).

Flat Fielding

CIBER images have a nonuniform response to a constant sky brightness across the detector array, known as the flat field response. For each CIBER field, the flat-field is estimated by averaging the dark-current-subtracted flight images of the other four sky fields.

A laboratory flat-field measurement was also taken before the flight using a field-filling integrating sphere, a uniform radiance source with a solar spectrum [described in 15]. Ideally, this is a better approach to measure the flat field since the one derived from stacking flight images contains fluctuations from the other fields that will not average down completely due to the small number of images. However, we found the flat field from the integrating sphere is not consistent with the flight data on large spatial scales [see 101], and therefore we do not use it in our analysis. The flat field estimator for the SWIRE field in CIBER 1.1 and 1.8 μm bands are shown in the Panel E of Fig. 5.1 and Fig. 5.2, respectively.

Surface Brightness Calibration

Throughout this work, we use nW m^{-2} sr^{-1} for the units of surface brightness (νI_ν). The calibration factor, C, that converts photocurrent (e$^-$ s^{-1}) to intensity (nW m^{-2} sr^{-1}) is derived in the following steps:

1. Take the raw images, subtract the dark current template, correct for the flat field, and apply the instrument and source masks;

2. Subtract the mean photocurrent in the unmasked region.
3. For each star in the Pan-STARRS catalog, calculate the flux νF_ν in CIBER bands from $m_{1.1}$ and $m_{1.8}$.
4. Sum the photocurrent in a 5×5 stamp centered on the source position[3].
5. Repeat step (3) and (4) for all the selected stars (see below) and take the average value of the flux ratio from (3) and (4) as the calibration factor C.

We select stars in the magnitude range $12.5 < m_{1.1} < 16$ for the 1.1 μm band, and $13.5 < m_{1.1} < 17$ for the 1.8 μm band. These magnitude ranges are chosen such that the brightest sources that saturate the detectors (even after non-linear correction) are excluded. Faint sources are not used because of their low signal-to-noise ratio. We use a different magnitude range for each band as they have different point source sensitivities. Panel F of Fig. 5.1 and Fig. 5.2 show the SWIRE field images masked by instrument masks at 1.1 and 1.8 μm, respectively, after flat fielding and calibration.

Background Removal

The total sky emission is composed of the EBL and various foreground components, including zodiacal light (ZL), diffuse galactic light (DGL), and integrated star light (ISL) from the Milky Way [71, 101]. ZL is the dominant foreground, approximately an order of magnitude brighter than the EBL [71]. Nevertheless, with its smooth spatial distribution on degree scales, the ZL can be mostly removed by subtracting the mean sky brightness in each field. Panel G of Fig. 5.1 and Fig. 5.2 show the mean-subtracted and masked SWIRE images at 1.1 and 1.8 μm, respectively. To highlight the large-scale fluctuations, we smooth the images with a $\sigma = 35''$ Gaussian kernel.

Image Filtering

Although the ZL signal is smooth, a flat-field estimation error may induce a nonuniform ZL residual that cannot be removed by mean subtraction. This residual may dominate over cosmological fluctuations on large scales. Therefore, after removing the mean value in the image, we filter the images by fitting and subtracting a 3rd/5th

---

[3] We have tested that using 3×3, 5×5, or 7×7 stamp size gives consistent results. Our beam size is approximately twice of the pixel size, so a 3×3 stamp already has enclosed most of the flux from a point source.

order 2-D polynomial function for the 1.1/1.8 μm images to filter out any residual large-scale variations (Panel H of Fig. 5.1 and Fig. 5.2). The filtering will also suppress large-scale cosmological signals, and therefore the choice of polynomial order used for filtering is determined by optimizing the trade-off between the reduction of background fluctuations and the large-scale two-halo signal. The effect of filtering on the detected one-halo and galaxy extension terms is small, as our filtering removes fluctuations at a much larger scales than these signals, and the signal filtering is accounted for in simulations (see Sec. 5.9).

5.4 External Catalogs

Throughout this work, we used several external source catalogs for (1) masking bright foreground sources (Sec. 5.3); (2) calibration (Sec. 5.3); (3) modeling the PSF by stacking bright stars in the fields (Sec. 5.7); and (4) selecting galaxies for stacking (Sec. 5.8).

To match the catalog sources to our data, we fit the astrometry coordinates of our images with the online software `nova.astrometry.net` [60]. For each image, we solve for the astrometry in four quadrants separately to mitigate the effect of image distortion. Since there is a fixed $\sim 50''$ misalignment between the 1.1 and 1.8 μm images as they are produced by different telescopes, their astrometry is solved separately.

Pan-STARRS

We use the Pan-STARRS catalog [22] for masking. Pan-STARRS covers all of the CIBER fields with a depth of $m \sim 20$ in the g, r, i, z, y bands. We query the source positions and magnitudes in all five Pan-STARRS bands from their DR1 `MeanObject` table, and derive $m_{1.1}$ and $m_{1.8}$ with the *LePhare* SED fitting software [12, 52]. We use sources that have a y band measurement and a quality flag (`qualityFlag` in `ObjectThin` table) that equals to 8 or 16 for masking.

2MASS

Some bright stars are not included in the Pan-STARRS catalog, and thus we use the 2MASS [89] `Point Source Catalog (PSC)` to get the complete point source list. For 2MASS sources, $m_{1.1}$ ($m_{1.8}$) is derived by linear extrapolation with the 2MASS photometric fluxes in J and H (H and K_s) bands, respectively. We also use bright stars in 2MASS for modeling the PSF (see Sec. 5.7).

SDSS

We use the Sloan Digital Sky Survey (SDSS) DR13 [14] `PhotoObj` catalog to get the star/galaxy classification ("type" attribute 6–stars, 3–galaxies) and the galaxy photometric redshift ("Photoz" attribute) for sources in our fields. This information is essential for selecting target galaxies for stacking and inferring their redshift distribution (Sec. 5.8), as well as selecting stars for stacking to model the PSF (Sec. 5.7).

SWIRE Photometric Redshift Catalog

Rowan-Robinson et al. [81, 82] performed SED fitting on $\sim 10^6$ sources in the SWIRE field, based on optical and infrared photometric data from multiple surveys. This provides information on the stellar masses of our stacked galaxies for our analysis (see Sec. 5.8).

Gaia

Gaia DR2 [40, 41] provides high-precision astrometry for stars in the Milky Way, which gives high-purity star samples used for both validating the PSF model (Sec. 5.7) and cleaning residual stars in the galaxy sample selected by SDSS (Sec. 5.8).

Nearby Cluster Catalog

Nearby galaxy clusters along the line of sight introduce extended emission in stacking, so we exclude galaxies that are close to nearby clusters (Sec. 5.8). We use the cluster catalog from Wen et al. [98], which compiles $0.05 \leqslant z < 0.8$ galaxy clusters detected in SDSS-III [11]. We also use the Abell cluster samples [3] for local galaxy clusters. There are 7 Abell clusters and ~ 200 clusters from Wen et al. [98] over the five CIBER fields.

5.5 Simulation Catalog—MICECAT

In addition to the observed source catalogs, we make use of the MICECAT simulated galaxy catalog [38, 39, 50] to estimate the signal from galaxy clustering. MICECAT is a product of the N-body cosmological simulation MICE Grand Challenge run (MICE-GC), which has 70 billion dark matter particles in a 3072^3 $\mathrm{Mpc}^3\mathrm{h}^{-3}$ cubic co-moving box. The dark matter halos are resolved down to $\sim 3 \times 10^{10} M_\odot h^{-1}$.

MICECAT is a mock catalog that simulates ideal observations of a 5000 deg^2 light cone covering $0 < z < 1.4$. MICECAT builds on MICE-GC by combining a halo occupation distribution (HOD) with subhalo abundance matching (SHAM) to calibrate to observed luminosity functions and clustering [20]. MICECAT simulates a mass-limited sample complete to $m_i \sim 22$ and $m_i \sim 24$ at $z \simeq 0.5$ and $z \simeq 0.9$, respectively [31]. The MICECAT mocks are large enough to permit us to generate up to $\sim 10^3$ independent CIBER field-sized (2×2 deg^2) mock catalogs. We use modeled magnitudes from MICECAT in Euclid NISP Y and H bands for CIBER $m_{1.1}$ and $m_{1.8}$, respectively, since the NISP filters are similar to the CIBER imager bands.

MICECAT simulates both central and satellite galaxies generated with its HOD+SHAM model, which allows us to model the linear (two-halo) and non-linear (one-halo) clustering in the stacking signal separately. We use the radial shapes derived from MICECAT stacking to fit the one-halo and two-halo amplitudes in our stacking data. Details on modeling galaxy clustering in the stacking signals are further described in Sec. 5.9.

5.6 Stacking

Sub-pixel Stacking

CIBER imager pixels under-sample the PSF, and therefore the surface brightness profile of individual sources is poorly resolved. However, given external source catalogs with high astrometric accuracy, we can stack on a sub-pixel basis and reconstruct the average source profile at scales finer than the native pixel size. This "sub-pixel stacking" technique has been used in previous CIBER imager analyses [15, 101], and further investigated recently in the context of optimal photometry [90]. We summarize the sub-pixel stacking procedure as follows:

1. Select a list of stacking target sources from external catalogs.
2. Re-grid each pixel into $N_{\mathrm{sub}} \times N_{\mathrm{sub}}$ sub-pixels (we use $N_{\mathrm{sub}} = 10$ in this work). The intensities of all sub-pixels are assigned to the same value as the native pixel without interpolation.
3. For each source, unmask pixels associated with its source mask. Pixels masked due to nearby sources or from the instrument mask remain masked.
4. Crop an $N_{\mathrm{size}} \times N_{\mathrm{size}}$ (at sub-pixel resolution) stamp centered on the target source. We choose $N_{\mathrm{size}} = 2401$ in this work, which corresponds to a $28' \times 28'$ stamp.

5. Repeat steps 3 and 4 for all target sources, average the stamps, and return the final stacked 2-D image $\Sigma_{\rm stack}(\mathbf{r})$.

The stacked profile $\Sigma_{\rm stack}$ is a convolution of the intrinsic source profile, $\Sigma_{\rm src}$, the instrument PSF ($PSF_{\rm instr}$)[4], and the pixel function $PSF_{\rm pix}$:

$$\begin{aligned}\Sigma_{\rm stack}(\mathbf{r}) &= [\Sigma_{\rm src}(\mathbf{r}) \circledast PSF_{\rm instr}(\mathbf{r})] \circledast PSF_{\rm pix}(\mathbf{r}) \\ &=\Sigma_{\rm src}(\mathbf{r}) \circledast PSF_{\rm stack}(\mathbf{r}),\end{aligned} \tag{5.1}$$

where $\mathbf{r} = (x, y)$ is a two-dimensional sub-pixel coordinate system with its origin at the stack center. We define the effective PSF as $PSF_{\rm stack}(\mathbf{r}) \equiv PSF_{\rm instr}(\mathbf{r}) \circledast PSF_{\rm pix}(\mathbf{r})$. The pixel function accounts for the fact that sub-pixels retain the value of the original pixels, which is a convolution effect. The pixel function is a matrix with each element proportional to the counts where the sub-pixel and the center sub-pixel that contains the source are within the same native pixel. The position of the center sub-pixel within the native pixel is a uniform probability distribution, and therefore when stacking on a large number of sources, the pixel function converges to the analytic form [90]:

$$PSF_{\rm pix}(\mathbf{r}) = \begin{cases} (N_{\rm sub} - x)(N_{\rm sub} - y) \\ \qquad\qquad \text{if } |x|, |y| < N_{\rm sub} \\ 0 \qquad \text{otherwise} \end{cases} \tag{5.2}$$

As a practical matter, $PSF_{\rm pix}$ can be determined through simulations. $PSF_{\rm stack}(\mathbf{r})$ can be measured by stacking stars in the field, where $\Sigma_{\rm src}(\mathbf{r})$ is a delta function, so $\Sigma_{\rm stack}(\mathbf{r}) = PSF_{\rm stack}(\mathbf{r})$. Note that the expression in the second line of Eq. 5.1 implies that the intrinsic profile $\Sigma_{\rm src}(\mathbf{r})$ can be obtained from the stacked profile $\Sigma_{\rm stack}(\mathbf{r})$ with the knowledge of $PSF_{\rm stack}(\mathbf{r})$, instead of determining $PSF_{\rm instr}(\mathbf{r})$.

We perform stacking and PSF modeling separately for each field, since $PSF_{\rm instr}$ is slightly different across the fields due to the varying pointing performance of the altitude control system during each integration (c.f. top panel of Fig. 5.5). After obtaining the 2-D stacked images, we bin them into 25 logarithmically-spaced 1-D radial bins. Within each bin, the number of stacked images on each sub-pixel is used for weighting when calculating the average profile in each radial bin. Note that the weight is not the same across sub-pixels since the masks are different for each stacked image.

---

[4] Instrument PSF includes all effects from the optics, detector array, and pointing jitter during the integration.

Covariance Matrix of Stacking Profile

The covariance matrix of the binned 1-D radial stacked profile is calculated with a jackknife resampling technique. For each stack, we split sources into $N_{\rm J} = 64$ sub-groups based on their spatial coordinates in the image. The CIBER imager arrays have 1024×1024 pixels, and thus each sub-group corresponds to sources in a 128×128 pixel sub-region on the array. The radial profile of the k-th jackknife sample, $\Sigma^k_{\rm stack}$, is obtained from stacking on sources in all the other sub-regions, and then the covariance matrix between radial bin (r_i, r_j) is given by

$$\begin{aligned}
&C_{\rm stack}(r_i, r_j) = \frac{N_{\rm J}-1}{N_{\rm J}} \sum_{k=1}^{N_{\rm J}} \left[\Delta\Sigma^k_{\rm stack}(r_i) \cdot \Delta\Sigma^k_{\rm stack}(r_j)\right] \\
&\Delta\Sigma^k_{\rm stack}(r_i) \equiv \Sigma^k_{\rm stack}(r_i) - \Sigma_{\rm stack}(r_i) \\
&\Delta\Sigma^k_{\rm stack}(r_j) \equiv \Sigma^k_{\rm stack}(r_j) - \Sigma_{\rm stack}(r_j),
\end{aligned} \tag{5.3}$$

where $\Sigma_{\rm stack}$ is the average stacked profile of all of the sub-regions.

One of our galaxy stacking samples (mag bin # 1 in Sec. 5.8) has a small number of sources ($\ll 64$ for each field), which makes the covariance estimation from the jackknife method unstable. Therefore we perform bootstrap resampling with $N_B = 1000$ realizations to calculate the covariance for this case. In this bootstrap, we obtain the radial profile of the k-th bootstrap sample, $\Sigma^k_{\rm stack}$, by stacking the same number of sources as the original sample, but the sources are randomly selected from the original sample with replacement. The covariance matrix is then given by

$$C_{\rm stack}(r_i, r_j) = \frac{1}{N_{\rm B}-1} \sum_{k=1}^{N_{\rm B}} \left[\Delta\Sigma^k_{\rm stack}(r_i) \cdot \Delta\Sigma^k_{\rm stack}(r_j)\right]. \tag{5.4}$$

In all the other cases, the covariance is derived from jackknife instead of bootstrap resampling since it is numerically expensive to perform a sufficient number of bootstrap realizations given that we have hundreds or thousands of galaxies per field in each stack. We assign galaxies to sub-groups by their spatial positions instead of randomly grouping them to account for large-scale spatial fluctuations.

The first few radial bins within the CIBER 7″ native pixel are highly correlated since all the sub-pixels are assigned to the same value as the native pixel. We also find a high correlation on large angular scales, as the stacking signal is dominated by large-scale spatial variations.

5.7 PSF Modeling

An accurate model for the PSF is essential for quantifying the galaxy extension from stacking images. As stars are point sources on the sky, we measure the PSF of each field by stacking stars in the same CIBER field. The radial profile of star stacks gives $PSF_{\rm stack}$ (Eq. 5.1), which accounts for all effects that distribute the light from a point source to the stacked profile, including spreading by the instrument optical system and detectors, pointing instability during integration, astrometry uncertainties, and the pixel function $PSF_{\rm pix}$. Since we use bright stars in the CIBER fields to model the PSF, the uncertainty on the PSF is subdominant to our galaxy stacked profiles.

Modeling $PSF_{\rm stack}$

Infrared detectors have a brightness-dependent PSF, the so-called "brighter-fatter effect" [49]. This nonlinearity makes brighter point sources appear broader on the detector array than fainter ones. To model $PSF_{\rm stack}$ robustly on both small and large scales, we construct an overall star profile from three brightness bins. For the core region ($r < 22''$), we stack $13 < m_{1.1} < 14$ sources in the field; for intermediate scales, $22'' < r < 40''$, we fit a slope to the stacking profile of $9 < m_J^{\rm 2MASS} < 10$ sources; for outer radii, we fit another slope to the stacking profile of the brightest $4 < m_J^{\rm 2MASS} < 9$ sources, and connect the two slopes at $r = 40''$ ($m_J^{\rm 2MASS}$ is the 2MASS J-band Vega magnitude). The choice of magnitude bins and transition radii minimizes the error on all scales. At small radii, using faint stars avoids detector nonlinearity, and at large radii, bright stars provide better sensitivity to the extended PSF. For the intermediate scales, we check that the fitted slope from the three star stacking profiles ($4 < m_J^{\rm 2MASS} < 9$, $9 < m_J^{\rm 2MASS} < 10$, $13 < m_{1.1} < 14$) are statistically consistent. The top panel of Fig. 5.3 shows $PSF_{\rm stack}$ from the SWIRE field in the 1.1 μm band. The top panel of Fig. 5.5 shows $PSF_{\rm stack}$ in all five fields in both bands. The slight variation across fields is due to the difference in the pointing stability during each integration, but such motion is common to all sources within an integration.

Validating $PSF_{\rm stack}$

To validate that our PSF model is applicable to the fainter sources of interest, we perform a consistency test by stacking on stars in the Gaia catalog within the same magnitude range as our stacked galaxy samples ($16 < m_{1.1} < 20$), and compare these star stacking profiles with our $PSF_{\rm stack}$ model.

To get a clean star sample free of galaxies, we apply the following criteria for selecting stars from Gaia:

1. The source has a parallax measurement $> 2 \times 10^{-4}$ mas (i.e., distance < 5 kpc).
2. No astrometric excess noise is reported in the Gaia catalog (`astrometric_excess_noise` = 0). Large astrometric excess noise implies the source might be extended rather than a point source.
3. No SDSS galaxies within $0.7''$ (sub-pixel grid size) radius around the source.
4. We classify SDSS stars and galaxies using 10 pairs of magnitude differences between the five Pan-STARRS photometric magnitudes (g, r, i, z, and y bands), rejecting sources if they are classified as galaxies by our trained model.

After selecting stars with the above conditions from the the Gaia catalog, we stack them in four equally-spaced magnitude bins between $16 < m_{1.1} < 20$, and compare their stacking profile with the $PSF_{\rm stack}$ model. These stars span the same brightness range used for galaxy stacking. We down-sample original 25 radial bins to 15 bins (7 bins for $16 < m_{1.1} < 17$ case), following the same binning used for the galaxy stacking profile (Sec. 5.8). The results in the 1.1 μm band SWIRE field are shown on the bottom panel of Fig. 5.3. In Fig. 5.4 we show the difference of Gaia star stacks and the $PSF_{\rm stack}$ model. The errors are propagated from the covariance of the $PSF_{\rm stack}$ model and Gaia star stacks. We also show the χ^2 values and the corresponding probability to exceed (PTE) on all five CIBER fields in both bands. The PSF model shows excellent agreement with the star stacks.

Modeling $PSF_{\rm instr}$

Although knowledge of the instrument PSF is not required for reconstructing the source profile $\Sigma_{\rm src}$ from the stacking profile $\Sigma_{\rm stack}$, $PSF_{\rm instr}$ is still needed when we model the clustering signal from a simulated catalog (Sec. 5.9), where we make mock galaxy images using the CIBER PSF and pixel gridding. $PSF_{\rm instr}$ is also useful for determining the masking radius for bright sources (Sec. 5.3).

$PSF_{\rm instr}$ is modeled as follows: first, we deconvolve $PSF_{\rm pix}(\mathbf{r})$ (Eq. 5.2) from the $PSF_{\rm stack}(\mathbf{r})$ model with 10 iterations of the Richardson-Lucy deconvolution algorithm [64, 79]. The deconvolution is unstable at large radii due to noise fluctuations. To get a smooth model for $PSF_{\rm instr}$, we fit a β model [21] to the 1-D profile of the

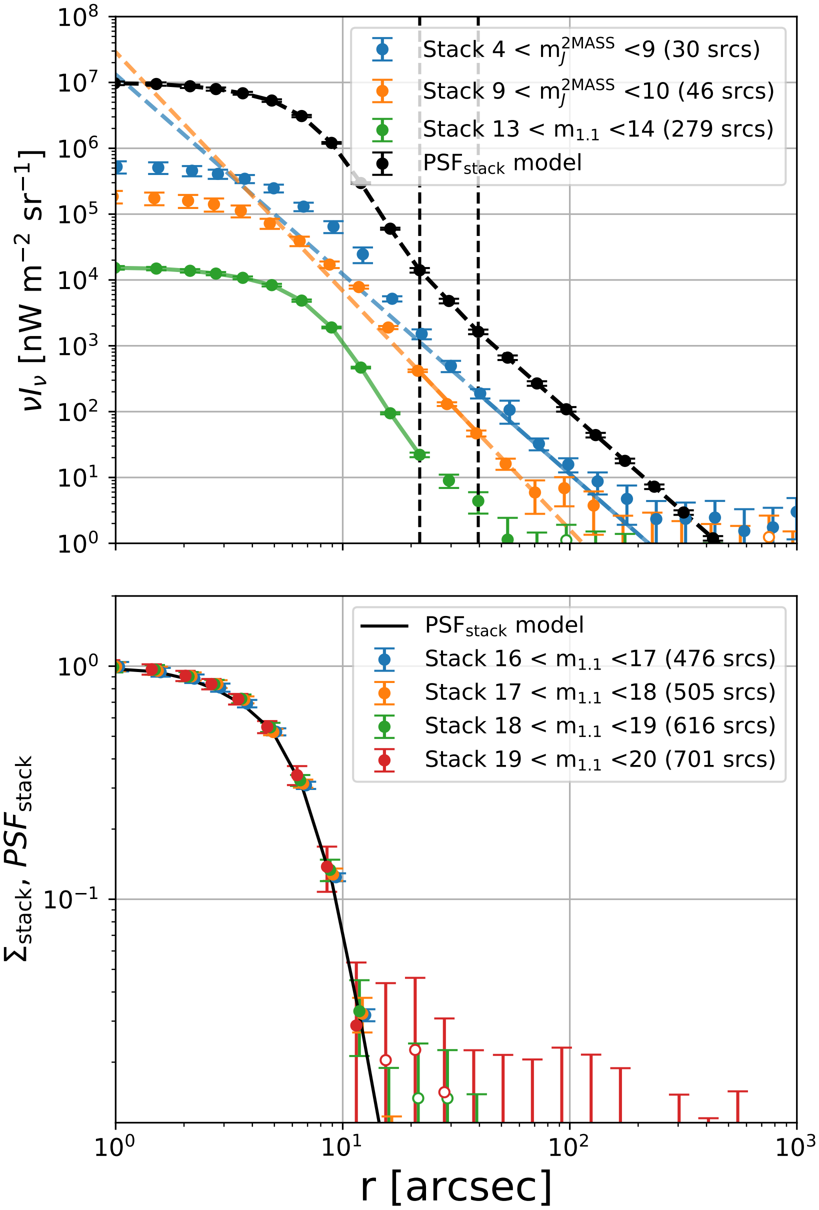

Figure 5.3: We illustrate the process of constructing and validating the $PSF_{\rm stack}(r)$ model, in the 1.1 μm band SWIRE field. Top: star stacking profile in three different brightness bins (blue, orange, and green), and the combined $PSF_{\rm stack}(r)$ model (black dashed curve) derived from splicing these three stacking profiles together at the radii marked by the black vertical dashed lines. The black data points show the binned $PSF_{\rm stack}(r)$ and the error bars propagated from their original star stacks. The filled data points and the three colored solid curves are the data used in the $PSF_{\rm stack}(r)$ model. Bottom: comparison of the $PSF_{\rm stack}(r)$ model with the stacking profiles from fainter stars selected from Gaia. The four chosen brightness bins match the ones used in galaxy stacking. The $PSF_{\rm stack}(r)$ model agree closely with the star stacking profiles, as shown in Fig. 5.4.

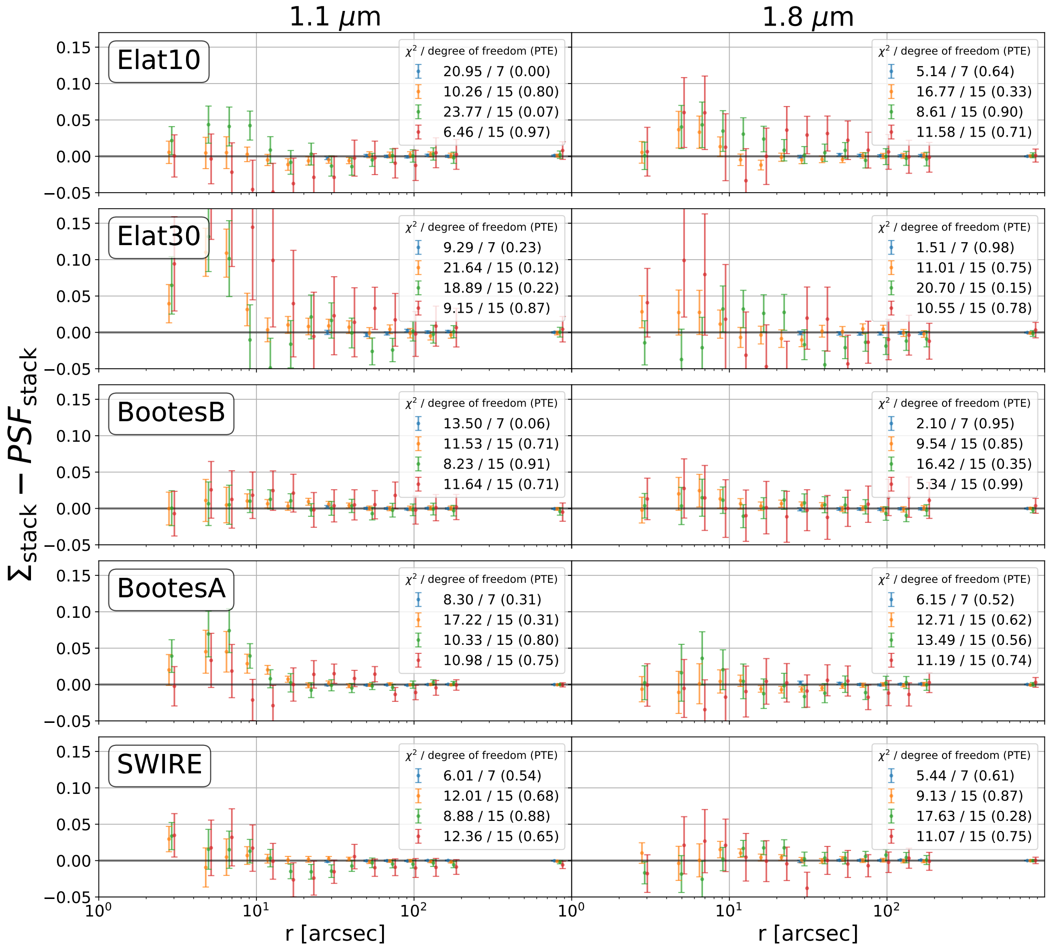

Figure 5.4: The difference of the $PSF_{\rm stack}(r)$ model and the star stacking profiles in all five CIBER fields in the 1.1 μm (left) and 1.8 (right) μm bands ($16 < m_{1.1} < 17$ (blue), $17 < m_{1.1} < 18$ (orange), $18 < m_{1.1} < 19$ (green), and $19 < m_{1.1} < 20$ (red)). The χ^2 values and their corresponding PTE given in the legend are consistent with the model. The degrees of freedom for each case is simply the number of radial bins. Open circles in the top and middle panels represent negative data points.

deconvolved image:

$$PSF_{\rm instr}(r) = \left(1 + \left(\frac{r}{r_c}\right)^2\right)^{-3\beta/2}. \tag{5.5}$$

Though not physically motivated, we find β model is a good empirical description of the extended PSF, and requires only two free parameters to achieve acceptable goodness of fit for every $PSF_{\rm stack}$.

The bottom panel of Fig. 5.5 illustrates this procedure in the 1.1 μm band of the SWIRE field. The $PSF_{\rm stack}$ model, obtained from star stacks in three different brightness bins, matches the β model of $PSF_{\rm instr}$ convolved with the pixel function

$PSF_{\rm pix}$ (Eq. 5.2). Our instrument PSF has comparable size to a pixel (FWHM $\sim 7''$).

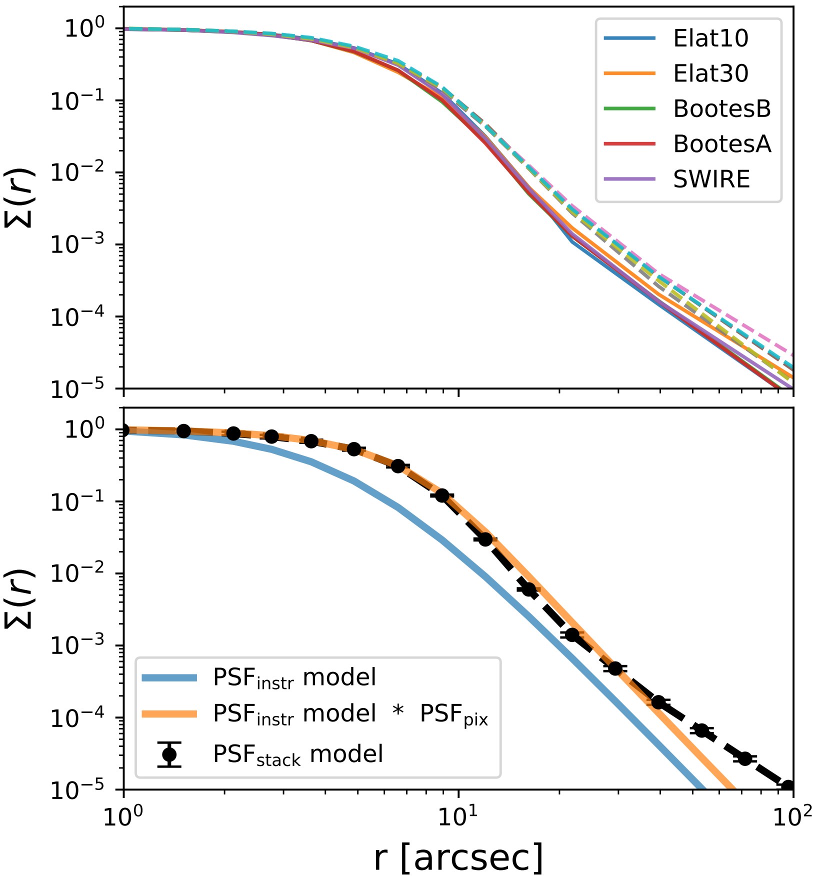

Figure 5.5: Top: $PSF_{\rm stack}$ model for each of the five fields in the 1.1 μm (solid) and 1.8 μm (dashed) bands. The variation across fields is due to the difference in pointing stability. Bottom: demonstration of the $PSF_{\rm instr}$ reconstruction process. Black data points show the $PSF_{\rm stack}$ model in the 1.1 μm band SWIRE field, derived from splicing the star stacking profile in three different brightness bins (c.f. Fig. 5.3 top panel). The blue line is the $PSF_{\rm instr}$ model derived from fitting a β model to $PSF_{\rm stack}$ after deconvolving $PSF_{\rm pix}$ with the Richardson-Lucy deconvolution algorithm. The orange line shows the convolution of $PSF_{\rm instr}$ with $PSF_{\rm pix}$ matching the $PSF_{\rm stack}$ model, as a consistency check. Our model for $PSF_{\rm instr}$ is in agreement with data for $r \lesssim 30''$. Our analysis is not susceptible to the moderate error at larger radii, as $PSF_{\rm instr}$ is only used for characterizing the clustering signal from nearby galaxies.

5.8 Galaxy Stacking

We stack galaxies within magnitude ranges $16 < m_{1.1} < 20$, divided into several sub-samples spanning $\Delta m_{1.1} = 1$. Our choice of magnitude bins optimizes the SNR on the stacks, giving sufficient sample sizes for each source brightness.

Source Selection Criteria

The stacking galaxy samples are selected from the SDSS catalog in the CIBER fields. To mitigate systematic effects from confusion, nearby clusters, or misclassified stars in the sample, we reject sources if they meet any of the following criteria:

- Sources are not labeled as galaxies in the SDSS catalog, i.e., the "type" attribute in the SDSS `PhotoObj` table is not equal to 3.
- Sources are located in the instrument mask.
- Other Pan-STARRS sources exist in the same CIBER pixel.
- The SDSS photometric redshift is less than 0.15. This criteria prevents nearby galaxies from introducing substantial power on large angular scales that would otherwise mimic the clustering signal.
- Sources have nearby Gaia counterparts within $0.7''$, i.e., the size of the sub-pixel used in our stacking. These sources are likely to be stars that are misclassified as galaxies in the SDSS catalog.
- Sources are within (1) a 500" radius of any galaxy cluster in Abell [3] (Sec 5.4); or (2) R_{200} of any galaxy cluster with halo mass $M_h > 10^{14} M_\odot$ or redshift $z < 0.15$ in the SDSS cluster catalog (Wen et al. [98], Sec 5.4). Approximately 10% of the sky area in each field is excluded by this condition.

The last condition mitigates contamination from nearby clusters along the line of sight, since they have structures spanning large angular scales, which will produce spurious large-scale extended signals in the stack. Furthermore, as we do not have information on whether a galaxy in SDSS is a member of a large galaxy cluster, the criteria also excludes cluster members from our stacking sample. Stacking on cluster members introduces extra non-linear one-halo clustering that can overwhelm the linear two-halo clustering signal on large scales.

To quantify the effect of applying this condition, we generate a mock CIBER map from the MICECAT catalog, implementing the same strategies described above to select sources, and stacking on the mock maps to measure the one- and two-halo clustering signals (see Sec. 5.9 for a detailed description of stacking with MICECAT-generated maps). We tested over a range of halo mass and redshift for selecting clusters, and found that excluding sources around clusters with $M_h > 10^{14} M_\odot$

(or redshift $z < 0.15$) can effectively reduce the one-halo clustering signal on large scales without losing a significant number of sources. For example, for the magnitude range of interest in this work (see Sec. 5.8), we can reduce the one-halo power by $\sim 3 - 5\times$ at 100 arcsec radius just by excluding galaxies near clusters following our criteria.

Stacking Sub-samples

For the SDSS galaxies within $16 < m_{1.1} < 20$ that survive all the selection criteria above, we split the sources into two sets. The first set is based on 1.1 μm flux in four bins: $16 < m_{1.1} < 17$, $17 < m_{1.1} < 18$, $18 < m_{1.1} < 19$, and $19 < m_{1.1} < 20$. Hereafter, these four bins are named “mag bin # 1”, “mag bin # 2”, “mag bin # 3”, and “mag bin # 4”, respectively. In addition, we also define a “total stack” with all $17 < m_{1.1} < 20$ sources to achieve better large-scale sensitivity.

The second set is defined by both the 1.1 μm apparent magnitude $m_{1.1}$ and the absolute magnitude $M_{1.1}$: $M_{1.1} = m_{1.1} - DM(z) + 2.5\log_{10}(1+z)$, where DM is the distance modulus, using SDSS photometric redshifts. The absolute flux serves as a proxy for galaxy size. Galaxies with comparable absolute flux have similar bolometric luminosity, which is correlated with stellar mass, star formation rate, etc. We use these samples to explore the dependence of our results on different galaxy properties. Since the sets approximately correspond to three higher and two lower stellar mass populations, with different redshift distributions, we call them “high-M/low-z”, “high-M/med-z”, “high-M/high-z”, “low-M/low-z”, and “low-M/med-z”.

In the SWIRE field, we have additional information from a photometric redshift catalog [81] based on an SED fit to each galaxy. As the stacked samples from each field are selected with the same criteria, we can assume the galaxy property distributions in the SWIRE field are the same as other fields, and thus infer the stellar mass distribution over all five fields. The $\log M_*$ column in Table 5.2 lists the median and 68% interval stellar mass in the SWIRE field samples from the Rowan-Robinson et al. [81] catalog. The stellar masses of our samples span from $\sim 10^{10.5}$ to $10^{12} M_\odot$, i.e., $\sim L_*$ galaxies at this redshift [74]. In addition, with the stellar mass distribution, we infer the host halo mass of our samples using the mean stellar-to-halo mass relation given by Zu & Mandelbaum [104], which connects the halo mass to stellar mass with galaxy clustering and lensing measurements. We also derive the corresponding virial radius, R_{200} (in physical and angular units), in Table 5.2. The virial radius is calculated from $R_{200} = [3M_h/(4\pi \cdot 200\rho_c)]^{1/3}$, where

ρ_c is the critical density.

We note that by selecting galaxies based on absolute or apparent fluxes, our samples will include both central and satellite galaxies. We infer the fraction of central galaxies, f_{cen}, in each sub-sample from MICECAT by applying the same selection criteria from a MICECAT simulation (i.e., observed magnitude, absolute magnitude and redshift cuts, and excluding sources close to nearby clusters). The distribution of redshift, stellar mass, halo mass, virial radius, and f_{cen} of our sub-samples are summarized in Fig. 5.6 and Table 5.2.

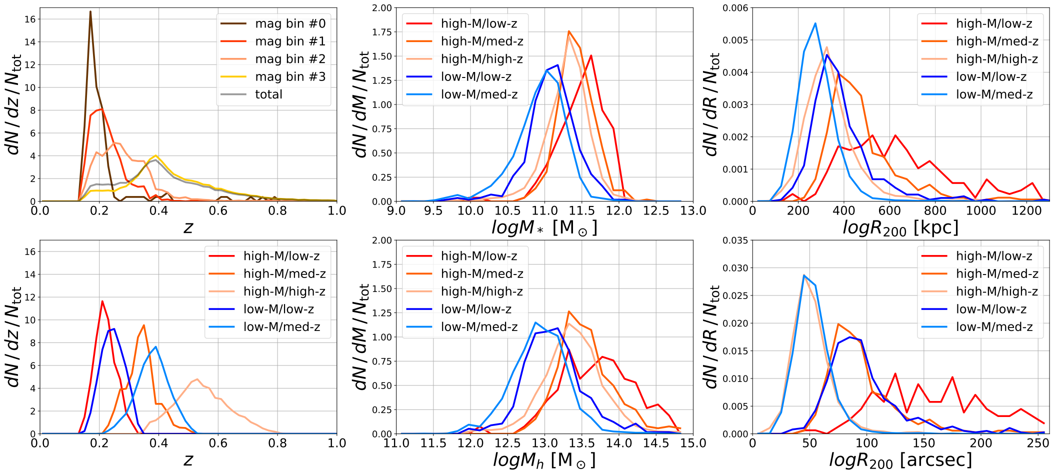

Figure 5.6: Left: redshift distributions of the 10 galaxy sub-samples used for stacking. The redshifts are derived from SDSS photometry. Middle-top: stellar mass distributions for the 5 apparent and absolute magnitude selected bins. The stellar masses are inferred from Rowan-Robinson et al. [81] for the SWIRE field. Middle-bottom: halo mass distributions in 5 apparent and absolute magnitude selected bins, modeled by applying stellar-to-halo mass relation from Zu & Mandelbaum [104]. Right: distributions of virial radius in physical (top) and observed angular (bottom) units. For visualization purposes, all curves are normalized by the total number of sources in each sub-sample (N_{tot}).

Galaxy Stacking Profile

We calculate 1-D radial profiles from galaxy stacks by averaging pixels in concentric annuli, as shown in Fig. 5.7 and Fig. 5.8. For comparison, we also plot the expected profile of stacked point sources, PSF_{stack}, scaled to match the first radial bin of the stacked galaxy profile. In all cases, the galaxy profiles are clearly broader than the PSF_{stack} profile.

Table 5.2: Summary of the properties on each stacked galaxy sub-sample with the +/− values indicating the 68% interval ranges.

Name	Selection Criteria	$N_{\rm gal}$	z	$\log M_*$ $[M_\odot]$	$\log M_h$ $[{\rm M}_\odot]$	R_{200} [kpc]	R_{200} [arcsec]	$f_{\rm cen}$[a]
mag bin #1	$16 < m_{1.1} < 17$	129	$0.18^{+0.04}_{-0.02}$	$11.6^{+0.3}_{-0.3}$	$13.8^{+0.5}_{-0.4}$	679^{+325}_{-181}	215^{+103}_{-57}	0.65
mag bin #2	$17 < m_{1.1} < 18$	1173	$0.21^{+0.07}_{-0.04}$	$11.5^{+0.3}_{-0.4}$	$13.7^{+0.6}_{-0.6}$	584^{+357}_{-215}	163^{+100}_{-60}	0.67
mag bin #3	$18 < m_{1.1} < 19$	3465	$0.27^{+0.09}_{-0.07}$	$11.2^{+0.4}_{-0.3}$	$13.3^{+0.5}_{-0.4}$	401^{+178}_{-116}	94^{+42}_{-27}	0.62
mag bin #4	$19 < m_{1.1} < 20$	31157	$0.42^{+0.17}_{-0.11}$	$11.1^{+0.3}_{-0.5}$	$13.0^{+0.5}_{-0.5}$	285^{+127}_{-86}	50^{+22}_{-15}	0.63
total	$17 < m_{1.1} < 20$	35795	$0.40^{+0.17}_{-0.14}$	$11.1^{+0.3}_{-0.4}$	$13.1^{+0.5}_{-0.5}$	302^{+135}_{-93}	55^{+24}_{-17}	0.63
high-M/low-z	$17 < m_{1.1} < 18, -23 < M_{1.1} < -22$	743	$0.22^{+0.04}_{-0.03}$	$11.6^{+0.2}_{-0.4}$	$13.7^{+0.5}_{-0.5}$	608^{+266}_{-201}	168^{+73}_{-55}	0.66
high-M/med-z	$18 < m_{1.1} < 19, -23 < M_{1.1} < -22$	1274	$0.34^{+0.05}_{-0.05}$	$11.4^{+0.3}_{-0.2}$	$13.5^{+0.4}_{-0.3}$	447^{+157}_{-94}	89^{+31}_{-19}	0.62
high-M/high-z	$19 < m_{1.1} < 20, -23 < M_{1.1} < -22$	10916	$0.54^{+0.10}_{-0.09}$	$11.3^{+0.3}_{-0.3}$	$13.4^{+0.3}_{-0.4}$	325^{+100}_{-82}	50^{+15}_{-13}	0.66
low-M/low-z	$18 < m_{1.1} < 19, -22 < M_{1.1} < -21$	1645	$0.24^{+0.05}_{-0.03}$	$11.1^{+0.3}_{-0.2}$	$13.1^{+0.4}_{-0.3}$	359^{+129}_{-78}	90^{+33}_{-20}	0.57
low-M/med-z	$19 < m_{1.1} < 20, -22 < M_{1.1} < -21$	14730	$0.38^{+0.05}_{-0.05}$	$11.0^{+0.2}_{-0.4}$	$12.9^{+0.3}_{-0.4}$	275^{+78}_{-67}	51^{+15}_{-13}	0.58

---

[a] $N_{\rm gal}$ is the total number of galaxies across five CIBER fields in each sub-sample, and the redshifts z are derived from SDSS photometry. The quantities on the left side of the double vertical line are derived from a partial set of samples or external catalogs for the sources used in stacks. We infer M_* by matching SWIRE field sources to the catalog from [81], assuming the same M_* distribution applies to the other four fields. The halo mass and the virial radius are derived with the stellar-to-halo mass relation from Zu & Mandelbaum [104]. The fraction of central galaxies ($f_{\rm cen}$) is derived by applying the same cuts to a simulated catalog from MICECAT.

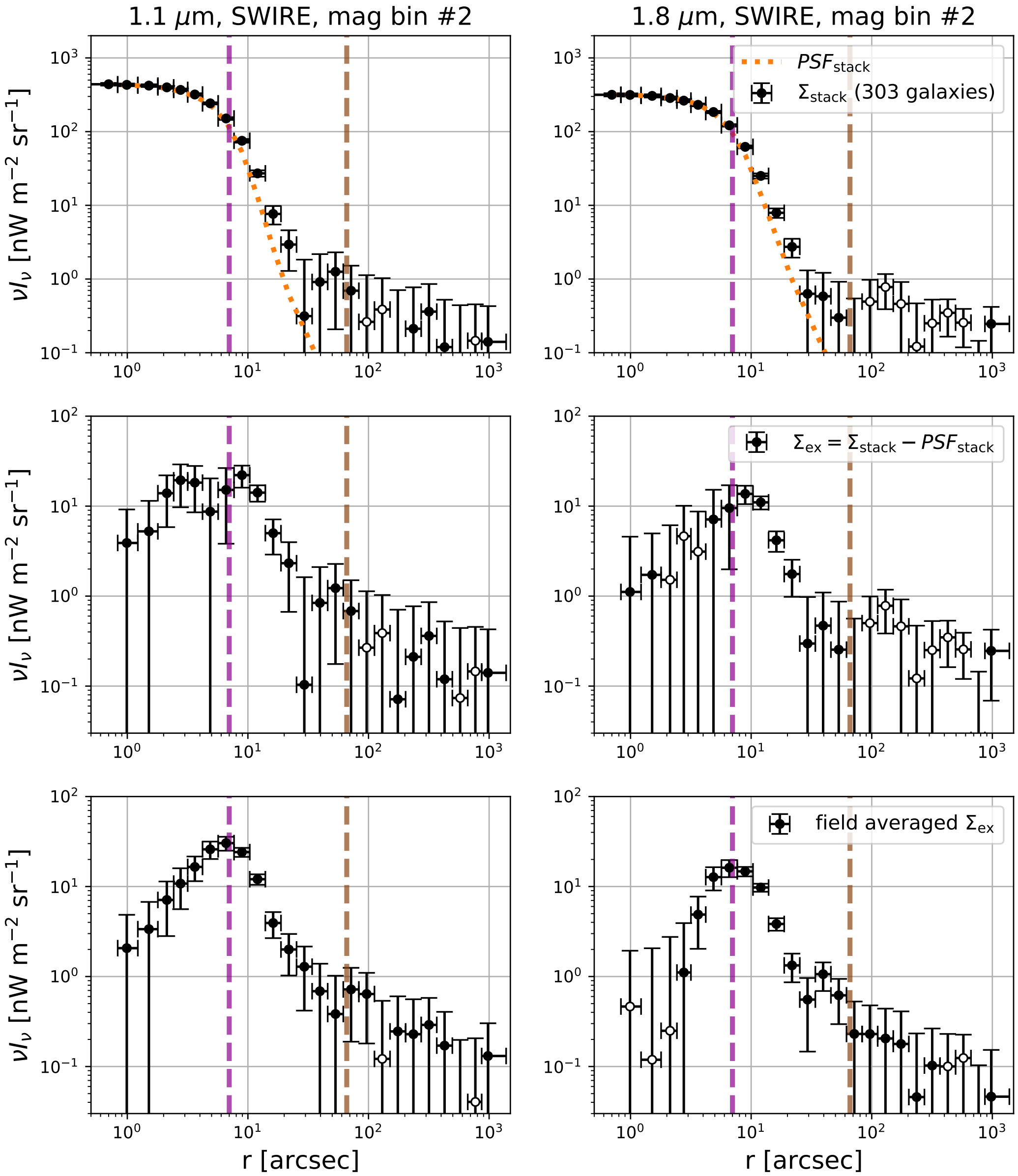

Figure 5.7: The stacked galaxy radial profile from the SWIRE field mag bin #2 in the 1.1 (left) and 1.8 μm (right) bands. Top: galaxy stacked profile Σ_{stack} (black) and PSF_{stack} model (orange dashed), scaled to match the innermost radial bin of Σ_{stack}. The error bars give the diagonal element of the covariance matrix derived by the Jackknife method (described in Sec. 5.3). Middle: the excess profile (Σ_{ex}, Eq. 5.6) for the case shown in the top row. The excess is defined as the difference between the galaxy stacked profile and the PSF_{stack} model, i.e., the difference of the black data from the orange curve in the top row. Bottom: the field-averaged excess profile Σ_{ex} for mag bin #2, derived from the weighted average of the excess profile in the five individual fields. The improved sensitivity from combining fields can be seen compared to the middle row. The purple and brown dashed lines mark the pixel size and the median R_{200} values inferred from MICECAT, respectively. Open circles in all the plots represent negative values.

Excess Profile

We define an "excess profile" $\Sigma_{\rm ex}(r)$ as follows:

$$\Sigma_{\rm ex}(r) = \Sigma_{\rm stack}(r) - A \cdot PSF_{\rm stack}(r), \tag{5.6}$$

where the normalization factor A is chosen such that $PSF_{\rm stack}$ matches $\Sigma_{\rm stack}$ at the innermost radial bin r_1, and thus by construction, $\Sigma_{\rm ex}(r_1) = 0$, and $A \equiv \Sigma_{\rm stack}(r_1)/PSF_{\rm stack}(r_1)$.

Since the excess profile is fixed at r_1, the uncertainties on the galaxy profile and the PSF profile at r_1 have to be accounted for by propagating this error to the other radial bins, and thus the excess profile covariance is given by

$$C_{\rm ex} = \Sigma_{\rm stack}(r_1)^2 \left[C_{\rm norm}\left(C_{\rm stack}\right) + C_{\rm norm}\left(C_{\rm PSF}\right)\right], \tag{5.7}$$

where $C_{\rm PSF}$ and $C_{\rm stack}$ are the covariance of $PSF_{\rm stack}$ and $\Sigma_{\rm stack}$, respectively, and

$$\begin{aligned} &C_{\rm norm}\left(C, \frac{\Sigma_{\rm stack}(r_i)}{\Sigma_{\rm stack}(r_1)}, \frac{\Sigma_{\rm stack}(r_j)}{\Sigma_{\rm stack}(r_1)}\right) \\ &\quad= \frac{\Sigma_{\rm stack}(r_i)\Sigma_{\rm stack}(r_j)}{\Sigma_{\rm stack}(r_1)^2} \cdot \\ &\quad\left[\frac{C(r_i, r_j)}{\Sigma_{\rm stack}(r_i)\Sigma_{\rm stack}(r_j)} - \frac{C(r_i, r_1)}{\Sigma_{\rm stack}(r_i)\Sigma_{\rm stack}(r_1)}\right. \\ &\quad\left. - \frac{C(r_j, r_1)}{\Sigma_{\rm stack}(r_j)\Sigma_{\rm stack}(r_1)} + \frac{C(r_1, r_1)}{\Sigma_{\rm stack}(r_1)^2}\right] \end{aligned} \tag{5.8}$$

is the covariance for the normalized profile that follows from the product rule for derivatives.

To fit a model to the measured $\Sigma_{\rm ex}$, we also need the inverse of $C_{\rm ex}$. However, $C_{\rm ex}$ is close to singular since our radial bins are highly correlated. Therefore, we reduce the original 25 radial bins to 15 bins by combining highly correlated bins in the inner and outer regions[5]. After this down-sampling, we derive the inverse covariance estimator by

$$C_{\rm ex}^{-1} = \frac{N_{\rm J} - N_{\rm bin} - 2}{N_{\rm J} - 1} C_{\rm ex}^{*-1}, \tag{5.9}$$

where $N_{\rm J} = 64$, the number of sub-groups used for estimating covariance, and the number of bins $N_{\rm bin} = 15$. $C_{\rm ex}^{*-1}$ is the direct inverse of the $C_{\rm ex}$ matrix, and the

---

[5] Mag bin # 1 is down-sampled to 7 radial bins as its degree of freedom is limited by the small number of stacked sources.

pre-factor in Eq. 5.9 de-biases the inverse covariance estimator, as our covariance matrix is derived from our data [47][6].

While we have high sensitivity on the small radial bins of both the galaxy stacked profiles and the PSF model, the A value has minimal dependency on the radius chosen for normalization, and the uncertainty of normalization has been accounted by the covariance (Eq. 5.7), and thhus our model parameter inference (Sec. 5.9) does not depend on the definition of the excess profile.

We present field-averaged excess profiles in Fig. 5.9. Note that the field-averaged excess profile is only plotted for visualization purposes, since the field-to-field PSF variation must be explicitly accounted in parameter fitting.

5.9 Modeling the Galaxy Profiles

We model the galaxy profile with three components as follows. We start by decomposing the stacked profile in image space (Sec. 5.9), define fitted profiles (Sec. 5.9), and introduce our model for each component of the stack (Sec. 5.9). Finally, we describe the model fitting procedure in Sec. 5.9.

Components in Image Space

The raw CIBER image, $I_{\rm raw}$, can be expressed as[7]

$$\begin{aligned} I_{\rm raw}(\mathbf{x}) = &\left[I_{\rm sig}(\mathbf{x}) + I_{\rm LoS}(\mathbf{x})\right] \circledast PSF_{\rm instr}(\mathbf{r}) \cdot FF(\mathbf{x}) \\ &+ I_{\rm DC}(\mathbf{x}) + I_{\rm n}(\mathbf{x}), \end{aligned} \tag{5.10}$$

where $\mathbf{x}$ is the 2-D pixel coordinate, FF is the flat-field gain, $I_{\rm DC}$ is the dark current map, and $I_{\rm n}$ is the read noise plus photon noise. The sky emission is decomposed into $I_{\rm sig}$ and $I_{\rm LoS}$ terms, where the first term accounts for the signal associated with stacked galaxies, and $I_{\rm LoS}$ represents uncorrelated emission from all other sources along the line of sight, including Galactic foregrounds.

After dark current subtraction and flat-field correction, we retrive $I'_{\rm raw}$:

$$I'_{\rm raw}(\mathbf{x}) = \left[I_{\rm sig}(\mathbf{x}) + I_{\rm LoS}(\mathbf{x})\right] \circledast PSF_{\rm instr}(\mathbf{r}) + I'_{\rm n}(\mathbf{x}), \tag{5.11}$$

where $I'_{\rm n}(\mathbf{x}) = I_{\rm n}(\mathbf{x})/FF(\mathbf{x})$, the instrument noise divided by the flat-field response. For simplicity, we ignore the error in the flat-field estimator in Eq. 5.11. In practice,

---

[6] For mag bin # 1, $N_{\rm bin} = 7$, and $N_{\rm J} = 64$ is replaced by $N_{\rm B} = 1000$ since we use bootstrap resampling method in this case.

[7] For clarification, $\mathbf{x}$ denotes 2-D coordinate on CIBER images, and $\mathbf{r}$ represents the coordinate that has origin at the source center, which is used in $PSF_{\rm instr}$ and stacked maps. Since we only consider 1-D radially averaged profile, $\mathbf{r}$ is replaced by 1-D variable "r".

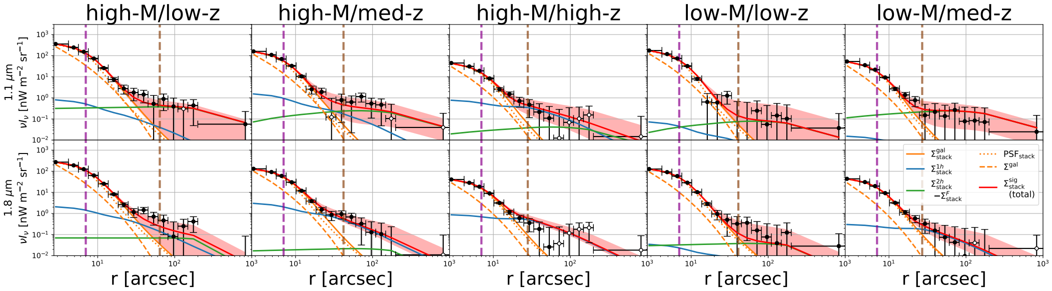

Figure 5.8: The stacked profile (black data) of each sub-sample stack averaged over five CIBER fields in the 1.1 μm (top) and 1.8 μm (bottom) bands. Red lines and shaded regions indicate the median and 68% confidence interval of the joint fit constrained through MCMC, respectively. The blue, green, and orange solid lines show the best-fit model of the stacked one-halo, two-halo, and galaxy profile term from MCMC. The orange dashed and dotted lines show the best-fit intrinsic galaxy profile $\Sigma^{\rm gal}$ and the $PSF_{\rm stack}$ model. The purple and brown dashed lines mark the pixel size ($7''$) and R_{200} value inferred from MICECAT. Open circles represent negative data points.

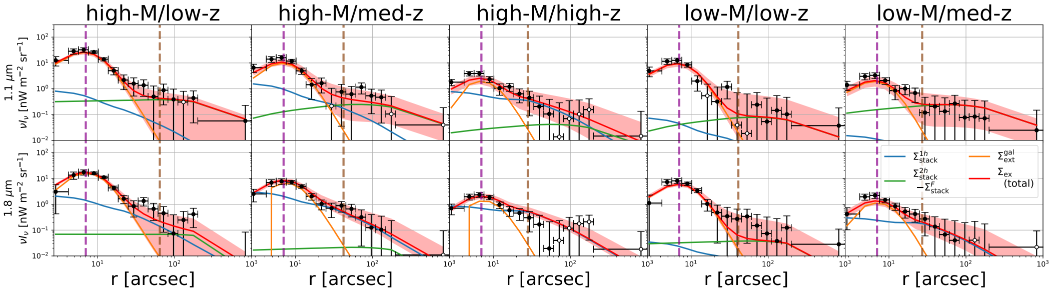

Figure 5.9: The measured (black data) and modeled (red) excess profile $\Sigma_{\rm ex}$ (black data) of each case shown in Fig. 5.8. Note the excess profile is defined by the difference of the stacked profile and $PSF_{\rm stacked}$ model (orange dotted line). other lines are same as the ones shown in Fig. 5.8.

the flat-field estimation uncertainties will not bias the stacking results as they are not correlated with individual stacked sources, and the effect on the covariance is accounted by the Jackknife method (see Sec. 5.6). We define the mask $M(\mathbf{x})$ as a binary function set to zero at masked pixels, and one otherwise. The filtered map is expressed with $\mathcal{F}\left[I'_{\rm raw}(\mathbf{x}), M(\mathbf{x})\right]$, which is a function of the input map $I'_{\rm raw}(\mathbf{x})$ and mask $M(\mathbf{x})$. As described in Sec. 5.3, we choose $\mathcal{F}$ to be a 3rd (1.1 μm)/5th (1.8 μm) order 2-D polynomial function fitted to the masked $I'_{\rm raw}$ map[8]. The image used

---

[8] Note that the filter map $\mathcal{F}$ can be decomposed into the sum of three filter maps because the polynomial fitting is a linear operation, i.e., given two maps $A(\mathbf{x})$ and $B(\mathbf{x})$, and a mask $M(\mathbf{x})$, $\mathcal{F}\left[A(\mathbf{x}) + B(\mathbf{x}), M(\mathbf{x})\right] = \mathcal{F}\left[A(\mathbf{x}), M(\mathbf{x})\right] + \mathcal{F}\left[B(\mathbf{x}), M(\mathbf{x})\right]$.

for stacking $I_{\rm map}$ can thus be written as

$$\begin{aligned} I_{\rm map}(\mathbf{x}) &= I'_{\rm raw}(\mathbf{x})M(\mathbf{x}) - \mathcal{F}\left[I'_{\rm raw}(\mathbf{x}), M(\mathbf{x})\right] M(\mathbf{x}) \\ &= I^{\rm sig}_{\rm map}(\mathbf{x}) + I^{\rm LoS}_{\rm map}(\mathbf{x}) + I^{\rm n'}_{\rm map}(\mathbf{x}), \end{aligned} \tag{5.12}$$

where

$$\begin{aligned} I^{\rm sig}_{\rm map}(\mathbf{x}) &= \left[I_{\rm sig}(\mathbf{x}) \circledast PSF_{\rm instr}(\mathbf{r})\right] M(\mathbf{x}) \\ &- \mathcal{F}\left[I_{\rm sig}(\mathbf{x}) \circledast PSF_{\rm instr}(\mathbf{r}), M(\mathbf{x})\right] M(\mathbf{x}), \end{aligned} \tag{5.13}$$

$$\begin{aligned} I^{\rm LoS}_{\rm map}(\mathbf{x}) &= \left[I_{\rm LoS}(\mathbf{x}) \circledast PSF_{\rm instr}(\mathbf{r})\right] M(\mathbf{x}) \\ &- \mathcal{F}\left[I_{\rm LoS}(\mathbf{x}) \circledast PSF_{\rm instr}(\mathbf{r}), M(\mathbf{x})\right] M(\mathbf{x}), \end{aligned} \tag{5.14}$$

and

$$\begin{aligned} I^{\rm n'}_{\rm map}(\mathbf{x}) &= I'_{\rm n}(\mathbf{x})M(\mathbf{x}) \\ &- \mathcal{F}\left[I'_{\rm n}(\mathbf{x}), M(\mathbf{x})\right] M(\mathbf{x}). \end{aligned} \tag{5.15}$$

Components in the Stack

The stacked profile $\Sigma_{\rm stack}$ can be expressed as the sum of stacks on the three maps in Eq. 5.12:

$$\Sigma_{\rm stack}(r) = \Sigma^{\rm sig}_{\rm stack}(r) + \Sigma^{\rm LoS}_{\rm stack}(r) + \Sigma^{\rm n}_{\rm stack}(r). \tag{5.16}$$

The last two terms can be ignored in modeling since they are uncorrelated with the stacked sources, so $\left\langle \Sigma^{\rm LoS}_{\rm stack}(r) \right\rangle = \left\langle \Sigma^{\rm n}_{\rm stack}(r) \right\rangle = 0$.

We model the stacked galaxy profile as

$$\begin{aligned} \Sigma^{\rm sig}_{\rm stack}(r) = [\ \Sigma^{\rm gal}_{\rm stack}(r) + \Sigma^{\rm 1h}_{\rm stack}(r) + \Sigma^{\rm 2h}_{\rm stack}(r)\ ] \\ - \Sigma^{\mathcal{F}}_{\rm stack}(r), \end{aligned} \tag{5.17}$$

where the first three terms are the signal terms, and the last term is the filtered signal map in Eq. 5.13. The galaxy profile term, $\Sigma^{\rm gal}_{\rm stack}$, represents the intrinsic galaxy profile, which includes the galaxy shape and the extended stellar halo. We decompose the galaxy profile term, $\Sigma^{\rm gal}_{\rm stack}$, into "core" and "extended" parts:

$$\Sigma^{\rm gal}_{\rm stack}(r) = \Sigma^{\rm gal}_{\rm core}(r) + \Sigma^{\rm gal}_{\rm ext}(r), \tag{5.18}$$

where the core component is the integrated emission of the $PSF_{\rm stack}$ fitted to the stacking profile, i.e., the $A \cdot PSF_{\rm stack}$ term in Eq. 5.6, and the extended component is the rest of the galaxy emission:

$$\begin{aligned}\Sigma^{\rm gal}_{\rm core}(r) =&\Sigma_{\rm stack}(r) - \Sigma_{\rm ex}(r),\\ \Sigma^{\rm gal}_{\rm ext}(r) =&\Sigma_{\rm ex}(r)\\ &- \left[\Sigma^{\rm 1h}_{\rm stack}(r) + \Sigma^{\rm 2h}_{\rm stack}(r) - \Sigma^{\mathcal{F}}_{\rm stack}(r)\right].\end{aligned} \tag{5.19}$$

In addition, galaxy clustering will also contribute to the stacked profile, primarily on large scales. We model clustering with the halo model framework [28], where large-scale clustering is described by the correlation within (one-halo) and between (two-halo) dark matter halos. $\Sigma^{\rm 1h}_{\rm stack}$ and $\Sigma^{\rm 2h}_{\rm stack}$ represent the profile for one- and two-halo clustering, respectively.

In practice, there is no well-defined boundary between the stellar halo of a galaxy and unbound stars in the dark matter halo, and the definition of IHL (or ICL) varies in the literature. To some degree, the galaxy extension term and the one-halo term each partially comprise stars not bound to individual galaxies in the halo. Since there are different definitions of IHL (or ICL) and the one-halo term in the literature, here we describe how our modelled components are defined.

In our definition, the galaxy extension describes emission associated with each galaxy, whereas the one-halo term accounts for other galaxies, their extensions, and diffuse stars in the same halo, as illustrated in Fig. 5.10. When we stack on a central galaxy, the galaxy extension term accounts for the extended emission around the stacked galaxy, and the one-halo term describes diffuse stars, undetected galaxies, and extension around all the satellite galaxies beyond masking limit in the same halo. Whereas, when we stack on a satellite galaxy, the galaxy extension term only includes the extended halo around that satellite galaxy, and all the other components are described by the one-halo term. In our sample, we estimate that $\sim 60\%$ of stacked galaxies are central galaxies, and $\sim 40\%$ are satellite galaxies (See Table 5.2).

Modeling the Stacked Galaxy Profile

The stacked galaxy profile $\Sigma^{\rm gal}_{\rm stack}(r) = \Sigma^{\rm gal}(r) \circledast PSF_{\rm stack}(r)$, is the intrinsic galaxy profile $\Sigma^{\rm gal}$, including the galaxy shape and the extended stellar halo, convolved with $PSF_{\rm stack}$. Following [97], we model $\Sigma^{\rm gal}$ with a double Sersic function:

$$\begin{aligned}\Sigma^{\rm gal}(r) = A^{\rm gal}\Big(&10^{I_{e,1}}\exp\left\{-b_{n1}\left[\left(r/R_{e,1}\right)^{1/n_1} - 1\right]\right\}\\ +\ &10^{I_{e,2}}\exp\left\{-b_{n2}\left[\left(r/\left(R_{e,1}+R_{e,2}\right)\right)^{1/n_2} - 1\right]\right\}\Big).\end{aligned} \tag{5.20}$$

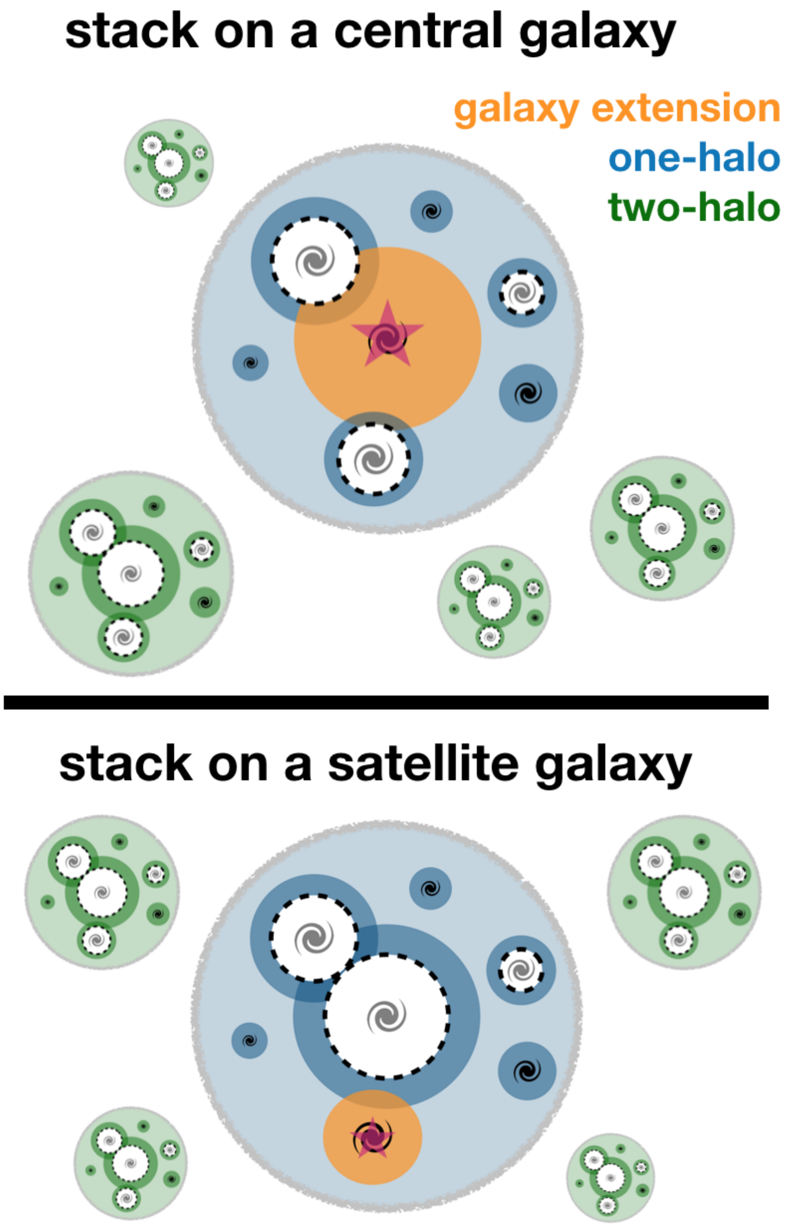

Figure 5.10: Illustration of the components in our model when stacking on a central (top) or a satellite (bottom) galaxy. The dark regions show the galaxy extensions associated with each galaxy, and the light blue and green regions show diffuse stars in the halos that are not tightly bound to any galaxy. The white parts with black dashed boundaries show the masked regions. The smaller galaxies without masks are fainter than the masking cutoff. The magenta stars and the orange regions show the stacked galaxy and its extension. The blue regions represent the one-halo term, and the green regions show the two-halo term contributed by emission from other halos. When stacking on a central galaxy, the one-halo term includes the satellite galaxy extensions beyond the masking radius, as well as faint satellite galaxies and their stellar halos. When stacking on a satellite galaxy, the one-halo term includes the extensions of both the central and the satellite galaxies beyond their masks, as well as the fainter satellite galaxies.

[97] performed a stacking analysis on isolated galaxies from Hyper Suprime-Cam (HSC) images, and fitted the stacked profile of their high-concentration samples with this model. The first term captures the galaxy shape, and the second term models the extended emission. Due to the lack of angular resolution in CIBER data, we are sensitive to the extended profile, and therefore we only vary $R_{e,2}$ to fit our stacked profile. We fix all of the other parameters to the best fit values given by Table 3 of [97], although when convolved, the total closely follows the PSF[9].

Our one- and two-halo clustering models, $\Sigma^{1h}_{\text{stack}}$ and $\Sigma^{2h}_{\text{stack}}$, and the filtered signal $\Sigma^{\mathcal{F}}_{\text{stack}}$, are constructed from the MICECAT simulation. MICECAT includes central and satellite galaxies of each halo, and each galaxy has a halo ID, enabling us to decouple the one-halo and two-halo contribution in the stacked signal, and thus to take into account the complication that we have both central and satellite galaxies in our samples. We model the one-halo term $\Sigma^{1h}_{\text{stack}}$ from MICECAT using the following steps:

1. Select the stacked target in the catalog using the same selection criteria.
2. For each target galaxy, generate a source map (using PSF_{instr}) for all galaxies residing in the same halo except for the target galaxy.
3. Generate a source mask using the same prescription as our data.
4. Stack on the target source position.
5. Iterate steps (2)-(4) for all target sources.

The derived stacked profile provides our template for the one-halo term, T^{1h}_{stack}. The filtered signal term $\Sigma^{\mathcal{F}}_{\text{stack}}$ accounts for the loss of clustering signal from filtering. $\Sigma^{\mathcal{F}}_{\text{stack}}$ is the stacked profile on the 2-D polynomial filtered map (the second term of Eq. 5.13), which can be modeled by filtering the simulated map from MICECAT. We model the two-halo term $\Sigma^{2h}_{\text{stack}} - \Sigma^{\mathcal{F}}_{\text{stack}}$ after filtering with the following process:

1. Make a CIBER-sized mock image from all the catalog sources with the model PSF_{instr}, and mask it with a source mask generated using the same masking process applied to the data.

---

[9] In [97], the values of $R_{e,1}$ and $R_{e,2}$ are reported in terms of $x_{e,1} = R_{e,1}/R_{200}$ and $x_{e,2} = R_{e,2}/R_{200}$. R_{200} is the projected virial radius of the host dark matter halo in angular units, and its value for each sub-samples is given in Table 5.2.

2. Fit and subtract a 2-D polynomial map to the image.
3. Select the stacked target in the catalog using the same selection criteria as the real sources.
4. Perform stacking with the target source, subtracting all galaxies within the same halo to remove the target galaxy and the one-halo contribution.
5. Iterate on step (4) to derive a stacked profile of the filtered two-halo signal.

The resulting stacked profile, $T_{\rm stack}^{\rm 2h-\mathcal{F}}$, is a model for $\Sigma_{\rm stack}^{\rm 2h} - \Sigma_{\rm stack}^{\mathcal{F}}$, which provides our template for the two-halo term. This process was performed on 400 realizations with CIBER-sized mock images from MICECAT, and we take the average stacked profile as the one-halo and filtered two-halo templates. As diffuse stars and faint galaxies below the resolution limit of MICECAT will not be accounted for, we assign free amplitudes to the one-halo and two-halo templates, which are then fit to the observed stacked data. Therefore, our three-parameter ($R_{\rm e,2}, A_{\rm 1h}, A_{\rm 2h}$) model can be written as

$$\begin{aligned}&\Sigma_{\rm stack}(r, \left\{R_{\rm e,2}, A_{\rm 1h}, A_{\rm 2h}\right\})\\ &\qquad = \Sigma^{\rm gal}(r, \left\{R_{\rm e,2}\right\}) \circledast PSF_{\rm stack}(r)\\ &\qquad + A_{\rm 1h}T_{\rm stack}^{\rm 1h}(r) + A_{\rm 2h}T_{\rm stack}^{\rm 2h-\mathcal{F}}(r).\end{aligned} \tag{5.21}$$

We note that the one- and two-halo profiles already include the PSF convolution in our model.

Model Fitting

For each CIBER field and band, we fit the excess profile Eq. 5.6, to a three-parameter model $\Sigma_{\rm ex}^{\rm m}(r, \{R_{\rm e,2}, A_{\rm 1h}, A_{\rm 2h}\})$ (Eq. 5.21) using a Markov Chain Monte Carlo (MCMC). We assume a Gaussian likelihood, which is given by

$$\begin{aligned}\chi^2 &= \left(\Sigma_{\rm ex}^{\rm d} - \Sigma_{\rm ex}^{\rm m}\right)^T C_{\rm ex}^{-1}\left(\Sigma_{\rm ex}^{\rm d} - \Sigma_{\rm ex}^{\rm m}\right)\\ \ln\mathcal{L} &= -\frac{1}{2}\chi^2 - \frac{1}{2}\ln|C_{\rm ex}| + \text{constant},\end{aligned} \tag{5.22}$$

where the inverse covariance $C_{\rm ex}^{-1}$ is given by Eq. 5.9.

We use the fit from individual fields for a consistency check. To provide a best estimate using the combination of all the fields that were observed at once, we also fit to the five CIBER fields using the joint likelihood:

$$\ln\mathcal{L} = \sum_{i=1}^{N_{\rm field}} \ln\mathcal{L}_i \tag{5.23}$$

where $N_{\text{field}} = 5$. Note that the PSF model is different for each field, so the information from different fields is combined in the likelihood.

We use the affine-invariant MCMC sampler `emcee` [37] to sample from the posterior distribution. We set flat priors for $R_{e,2}$, A_{1h}, and A_{2h} in the range of $[10^{-4}R_{200}, R_{200}]$, [0, 50], and [0, 200], respectively. We use an ensemble of 100 walkers taking 1000 steps with 150 burn-in steps. We checked that the chains show good convergence by computing the Gelman-Rubin statistic R [42]. For all three parameters in all cases, we find $R < 1.1$.

5.10 Results

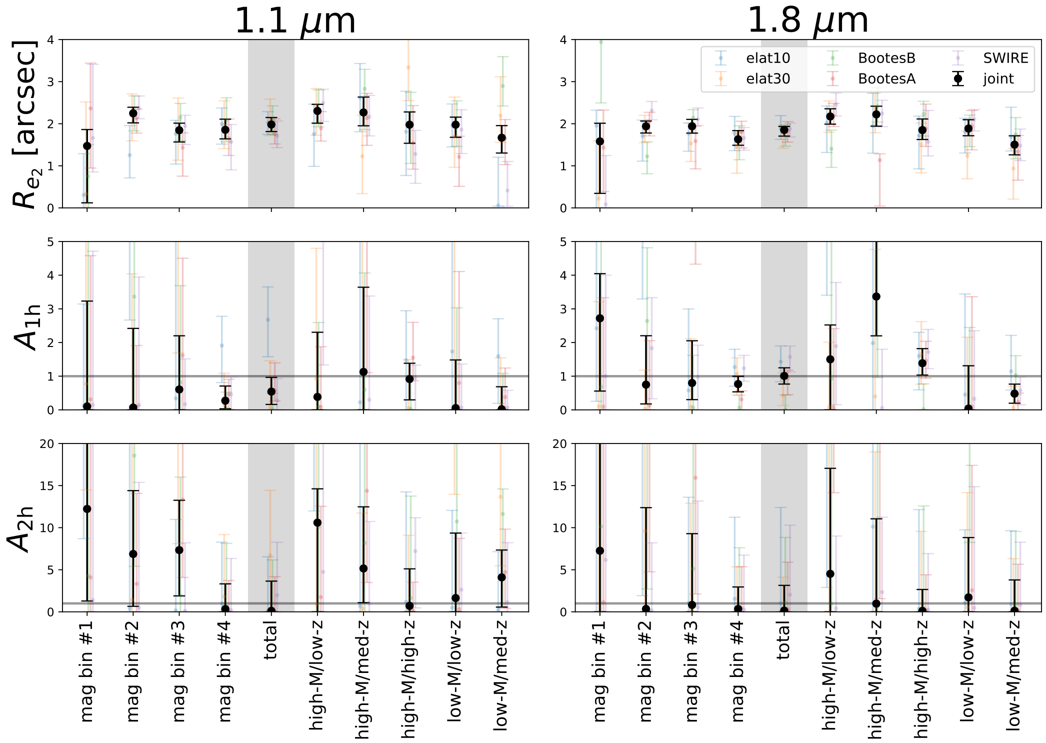

Figure 5.11: Marginalized parameter constraints from MCMC for each case listed in Table 5.2. The data points and error bars are the median and 68% confidence intervals from MCMC. Black data points show the joint fit from all five fields, with colored points for the individual fields. The gray horizontal lines in the middle and bottom panels mark $A_{1h} = 1$ and $A_{2h} = 1$, which are the clustering amplitudes given by MICECAT. The shaded regions show the total stack over all $17 < m_{1.1} < 20$ galaxies.

We show the MCMC results in Fig. 5.11 and Table 5.3, for all cases listed in Table 5.2. As a sanity check, we calculate the χ^2 value between the results from individual fields and the joint fit using 100 data points for each of the three parameters (5 fields × 10

Table 5.3: Summary of parameter constraints from the joint fit in each case listed in Table 5.2. For the cases with less than a 2σ detection (95% confidence interval), we quote the 2σ upper bound. For detections, the +/− values enclose the 68% confidence interval.

Name	1.1 μm $R_{e,2}$ [arcsec]	1.1 μm A_{1h}	1.1 μm A_{2h}	1.8 μm $R_{e,2}$ [arcsec]	1.8 μm A_{1h}	1.8 μm [a] A_{2h}
mag bin #1	< 2.76	< 6.06	< 48.91	< 2.53	< 5.72	< 58.05
mag bin #2	$2.25^{+0.14}_{-0.23}$	< 4.70	< 24.22	$1.94^{+0.12}_{-0.16}$	< 3.44	< 24.76
mag bin #3	$1.85^{+0.17}_{-0.28}$	< 4.18	< 18.94	$1.94^{+0.16}_{-0.16}$	< 2.96	< 18.30
mag bin #4	$1.85^{+0.25}_{-0.21}$	< 1.16	< 6.87	$1.63^{+0.21}_{-0.14}$	$0.77^{+0.23}_{-0.23}$	< 6.59
total	$1.98^{+0.17}_{-0.17}$	< 1.41*	< 7.30	$1.85^{+0.08}_{-0.15}$	$1.01^{+0.24}_{-0.24}$	< 6.86
high-M/low-z	$2.30^{+0.16}_{-0.29}$	< 4.76	< 25.58	$2.17^{+0.18}_{-0.18}$	< 4.2	< 33.10
high-M/med-z	$2.27^{+0.37}_{-0.32}$	< 6.42	< 19.53	$2.22^{+0.19}_{-0.28}$	$3.37^{+1.99}_{-1.17}$	< 22.76
high-M/high-z	$1.98^{+0.30}_{-0.44}$	< 1.88	< 9.08	$1.85^{+0.26}_{-0.22}$	$1.39^{+0.43}_{-0.35}$	< 6.19
low-M/low-z	$1.98^{+0.18}_{-0.30}$	< 3.18	< 16.38	$1.89^{+0.21}_{-0.17}$	< 2.77	< 17.65
low-M/med-z	$1.67^{+0.29}_{-0.36}$	< 1.30	< 11.30	$1.50^{+0.21}_{-0.24}$	< 1.01	< 7.58

---

[a] In 1.1 μm "total" bin, the 68% confidence interval of one-halo amplitude A_{1h} is $0.54^{+0.42}_{-0.38}$, approximately an 1σ detection.

mag bins × 2 bands). The resulting χ^2 values indicate our fit is internally consistent across the 5 CIBER fields. In Fig. 5.8 and Fig. 5.9, we show the stacked and excess profile data averaged over five fields, respectively, along with the marginalized one-halo, two-halo, and galaxy profile model from the joint fit. Fig. 5.12 shows the fitted intrinsic galaxy profile $\Sigma^{\rm gal}$ (Eq. 5.20) and the one- and two- halo terms in the "total" magnitude bin, also averaged over five fields. The field-averaged profiles are only shown for visualization purposes; when we fit the data with MCMC, the information is combined in the likelihood function rather than in data space.

5.11 Discussion

Missing Light in Galaxy Photometry

Given the best-fitting extended galaxy profile, we can calculate the fraction of flux missed in photometric galaxy surveys using a limited aperture. From our model, the fraction of flux within a photometric aperture can be approximated by $f_{\rm core} \equiv L_{\rm core}/(L_{\rm core} + L_{\rm ext})$, where $L_{\rm core}$ and $L_{\rm ext}$ are the total flux in the core and extension profile (Eq. 5.19), respectively. In practice, there are various ways to perform photometry. The Petrosian flux [75] is derived from aperture photometry and thus it is the most straightforward method to compare to our results. The Petrosian flux is defined by the total flux within a multiplicative factor of the Perosian radius of sources. We obtain the Petrosian radius and Petrosian flux from the SDSS catalog of each stacked galaxy in our sample. In SDSS, the Petrosian flux is calculated by integrating the emission within twice the Petrosian radius[10]. With our galaxy profile, we can calculate the fraction of flux within the same radius ($f_{\rm petro}$). The results are summarized in Table 5.4.

We also estimate the missing light fraction with the 'model magnitude' given in SDSS ($f_{\rm model}$). Rather than integrating within a certain aperture size, the model magnitude is derived by fitting the galaxy profile with an exponential or de Vaucouleurs functional form, choosing the one with the higher likelihood in the fitting[11]. While it is difficult to apply the same fitting procedure to the sources in CIBER images, we can calculate the ratio between the model flux and the Petrosian flux of each source in the SDSS catalog, and thus infer the fraction of missing light in the model flux. We find that both the Petrosian flux, which measures source emission within a limited aperture size, and the model flux derived from fitting a light profile

---

[10] `https://www.sdss.org/dr12/algorithms/magnitudes/#mag_petro`

[11] See `https://www.sdss.org/dr12/algorithms/magnitudes/` for the detailed descriptions on model magnitude.

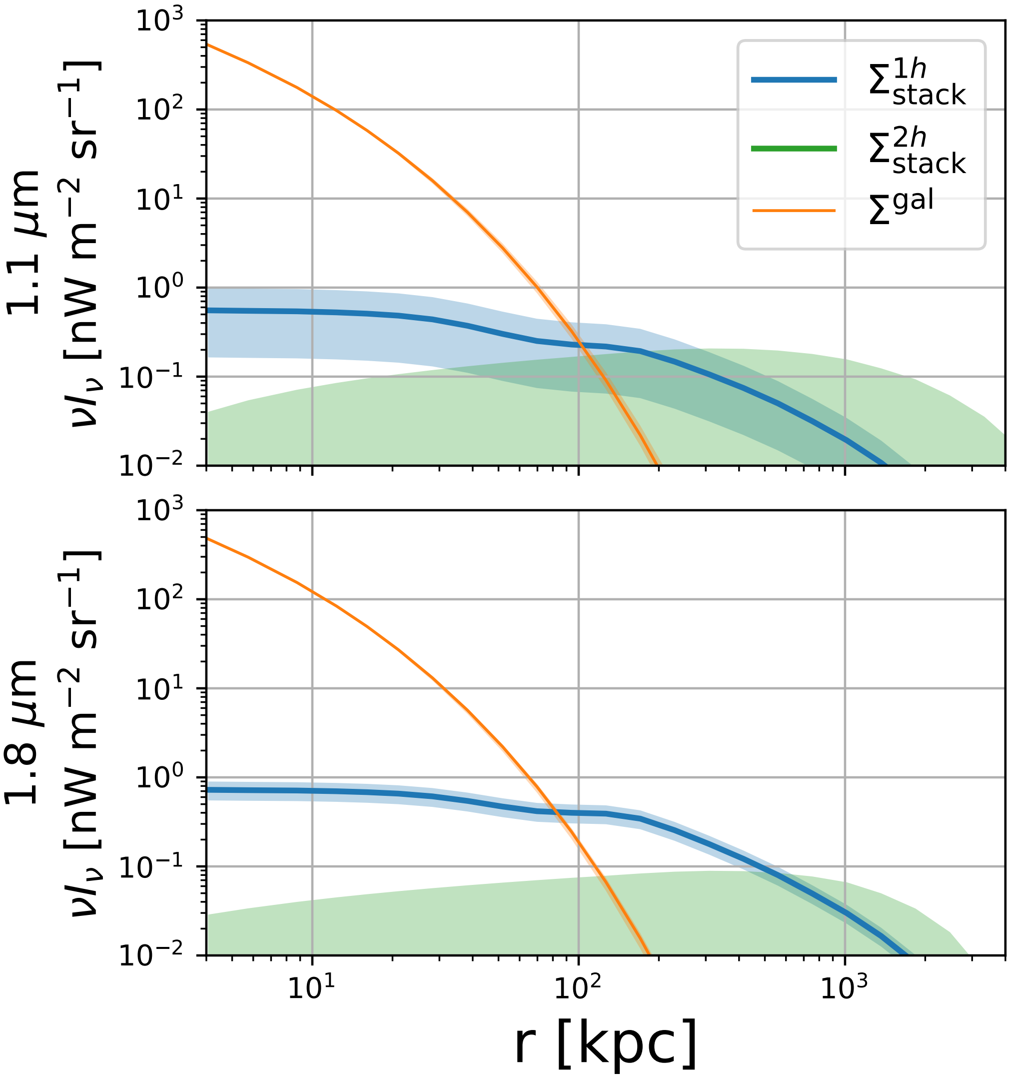

Figure 5.12: The fitted intrinsic galaxy profile Σ^{gal} (Eq. 5.20) (orange), stacked one-halo (blue) and two-halo (green) profiles in the "total" magnitude bin averaged over five CIBER fields in the 1.1 μm (top) and 1.8 μm (bottom) bands. We convert the angular scale to physical units (kpc) using the median conversion factor inferred from MICECAT (Table 5.2). Solid lines and shaded regions indicate the median and 68% confidence interval of the joint fit constrained through MCMC, respectively.

to the small-radii regions of the galaxy, miss $\sim 20\%$ of the total galaxy light, a deficit detected at $\sim 7\sigma$ ($\sim 4\sigma$) level for Petrosian (model) flux when combing constraints from all five sub-samples. This value is slightly larger than the light fraction in our galaxy extension term ($\sim$ 10 to 20 %). Our results on the missing light fraction in the Petrosian flux is in agreement with previous analytical calculation [45]. Interestingly, Tal & van Dokkum [91] probed the radial profile of $z \sim 0.34$ luminous red

Table 5.4: Fraction of flux in core component compared to flux captured in Petrosian and SDSS model flux, assuming the galaxy light profile follows the stacking results in this work. The total row shows the weighted average of the five listed sub-samples.

Name	1.1 μm f_{core}	1.1 μm f_{petro}	1.1 μm f_{model}	1.8 μm f_{core}	1.8 μm f_{petro}	1.8 μm f_{model}
high-M/low-z	$0.79^{+0.04}_{-0.02}$	$0.78^{+0.08}_{-0.10}$	$0.84^{+0.11}_{-0.12}$	$0.81^{+0.02}_{-0.02}$	$0.80^{+0.07}_{-0.10}$	$0.85^{+0.10}_{-0.12}$
high-M/med-z	$0.81^{+0.04}_{-0.05}$	$0.74^{+0.07}_{-0.13}$	$0.78^{+0.08}_{-0.15}$	$0.83^{+0.04}_{-0.03}$	$0.75^{+0.08}_{-0.11}$	$0.78^{+0.10}_{-0.13}$
high-M/high-z	$0.86^{+0.06}_{-0.04}$	$0.73^{+0.07}_{-0.16}$	$0.77^{+0.15}_{-0.19}$	$0.89^{+0.03}_{-0.04}$	$0.75^{+0.07}_{-0.16}$	$0.79^{+0.16}_{-0.18}$
low-M/low-z	$0.84^{+0.04}_{-0.02}$	$0.78^{+0.05}_{-0.11}$	$0.80^{+0.10}_{-0.12}$	$0.85^{+0.02}_{-0.03}$	$0.79^{+0.05}_{-0.11}$	$0.81^{+0.10}_{-0.12}$
low-M/med-z	$0.89^{+0.04}_{-0.04}$	$0.78^{+0.06}_{-0.16}$	$0.80^{+0.09}_{-0.16}$	$0.92^{+0.03}_{-0.03}$	$0.80^{+0.06}_{-0.15}$	$0.83^{+0.10}_{-0.14}$
total	$0.83^{+0.02}_{-0.01}$	$0.77^{+0.03}_{-0.06}$	$0.80^{+0.05}_{-0.06}$	$0.86^{+0.01}_{-0.01}$	$0.78^{+0.03}_{-0.05}$	$0.81^{+0.05}_{-0.06}$

galaxies (LRGs) in SDSS with a stacking analysis, and they also found ~20% of the total light missing at large radii when fitting a Sersic model to individual galaxies. Although their galaxy samples are at somewhat higher mass ($M_* \sim 10^{11} - 10^{12} M_\odot$), and model magnitudes are fitted with a different functional form, we arrive at a similar fraction of missing flux.

Extended Stellar Halo

The Illustris simulation [80] traces the dynamics and merger history of stellar particles and estimates the "ex-situ" population of stars that formed in other galaxies, and were later stripped and accreted into a new galaxy. The shaded region in the left panel of Fig. 5.13 shows the ex-situ stellar mass fraction at $z = 0$ from the Illustris simulation [80]. Although it is difficult to measure the ex-situ component in observations, Huang et al. [51] has studied individual stellar halos out to 100 kpc in more massive galaxies ($10^{11} M_\odot \lesssim M_* \lesssim 10^{12} M_\odot$) at higher redshifts ($z \sim 0.4$) in HSC images, finding that the fraction of stellar mass between 10 and 100 kpc is in good agreement with the ex-situ fraction constraints from Illustris [80]. In addition, Wang et al. [97] probe the stellar halo around local ($0 \lesssim z \lesssim 0.25$) low-mass galaxies ($9.2\, M_\odot < \log M_* < 11.4\, M_\odot$) with a stacking analysis on HSC images in r-band. They stacked galaxies out to ~ 120 kpc within several stellar mass bins. For each bin, they split the sources into low and high concentration populations, defined by $C < 2.6$ and $C > 2.6$, where $C = R_{90}/R_{50}$ is the ratio of the radii that contain 90% and 50% of the r-band Petrosian flux.

CIBER extends the HSC measurements to higher redshifts and longer wavelength bands. Armed with light profile fits, we can quantify the luminosity fraction in the extended stellar halo around the stacked sources. The left panel of Fig. 5.13 shows the fraction of stellar flux between radii of 10 and 100 kpc, using the fitted galaxy profile from CIBER and HSC [97]. We observe that ∼ 50% of the flux originates at galactocentric distances between 10 and 100 kpc. [97] re-scaled their images to physical units before stacking, whereas in our analysis we stack sources in observed angular units. Therefore, the variations in our measurements are mostly due to the variation of the conversion factor from angular to physical units for each galaxy in our stack. Our constraints are consistent with the HSC results in the highest mass bin.

Both CIBER and HSC are consistent with the ex-situ fraction from Illustris at $z = 0$, but are systematically higher than the median value from Illustris (the grey line in Fig. 5.13). One possible explanation is that the flux between 10 and 100 kpc is not a perfect proxy of the ex-situ population for lower mass galaxies. For example, D'Souza et al. [34] has shown that the transition scale between in-situ and ex-situ components varies across a wide range from ∼ 10 to ∼ 50 kpc, depending on the stellar mass and concentration of the galaxies. Nevertheless, given the limited information in stacking, we use this definition to associate the luminosity from beyond 10 kpc with IHL.

We also show the fraction of flux with 20 kpc radius cut in the right panel of Fig. 5.13. A radius cut at 20 kpc is a more suitable choice to describe the stripped stellar populations. For example, Milky Way-sized simulations suggest that the infalling stellar debris is recaptured by the galaxy and results in disk thickening at $r \lesssim 10$ kpc [72]. We note that each galaxy has a different stellar halo profile, interpreting our stacking results requires knowing the stellar halo profile on average over a large sample of galaxies. Given the uncertainty in choosing an average IHL radius, we report our results in both 10 kpc and 20 kpc scales, while show the full radial range in Fig. 5.14. We find that ∼ 25% of galaxy fluxes are from outside 20 kpc. The CIBER constraints shown in Fig. 5.13 are summarized in Table .6.

EBL from Extended Stellar Halos

With the galaxy profile from CIBER and HSC, we can estimate the EBL contribution from the extended regions at the redshift of our stacked sources. We model this quantity in the following steps:

1. For any given radius cut r_{cut}, we model the fraction of light beyond r_{cut} as a function of stellar mass by fitting a line to all CIBER and HSC data points in logarithmic space;

2. We estimate total stellar mass density by integrating the stellar mass function from Muzzin et al. [74] (we take their single Schechter function fit with all samples in $0.2 \leqslant z < 0.5$ bin, approximately the redshift of our sources);

3. For each r_{cut}, we apply the fraction derived in step 1 to the stellar mass function, and integrate to get the stellar mass density from sources outside r_{cut};

4. Assuming the mass-to-light ratio is the same for all galaxies, the ratio between step 3 and step 2 is our estimate of the EBL fraction from extended sources as a function of r_{cut}.

The results are shown in Fig. 5.14. We get approximately 30/15 % of extended emission in the EBL with $r_{\text{cut}} = 10/20$ kpc, respectively. Note that these values are close to the fraction in the five individual stellar mass bins from our stacking results. This is expected as our samples are at $\sim L*$ scale, which are the representative population that contains the majority of the total stellar emission of their redshift.

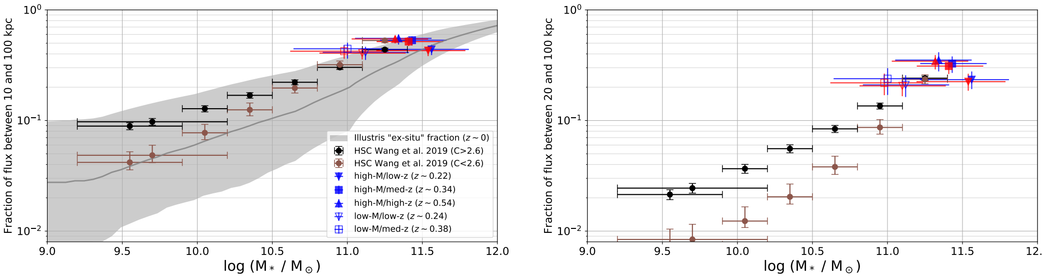

Figure 5.13: Fraction of flux between 10 (left)/20 (right) and 100 kpc from the galaxy profile derived from CIBER stacking (this work) in the 1.1 (blue) and 1.8 (red) μm bands and from HSC stacking [97]. The HSC stacking is performed on low and high concentration populations ($C < 2.6$ and $C > 2.6$) at optical wavelengths (r band). The horizontal error bars define the lower and upper bounds of the stellar mass of each stacking sample. The grey line and the shaded regions in the left panel are the median, 16th, and 84th percentile of ex-situ stellar mass fraction at $z = 0$ from Illustris simulations [80]. The shaded region shows the variance between individual galaxies in Illustris, whereas for CIBER and HSC, the error bars represent the standard error on the mean value.

Intra-halo Light Fraction

The fraction of the total emission from a dark matter halo associated with IHL, f_{IHL}, has been investigated with both observation and theoretical modeling [e.g., 18, 34, 35, 43, 63, 77]. With our stacking results, we can estimate the total halo

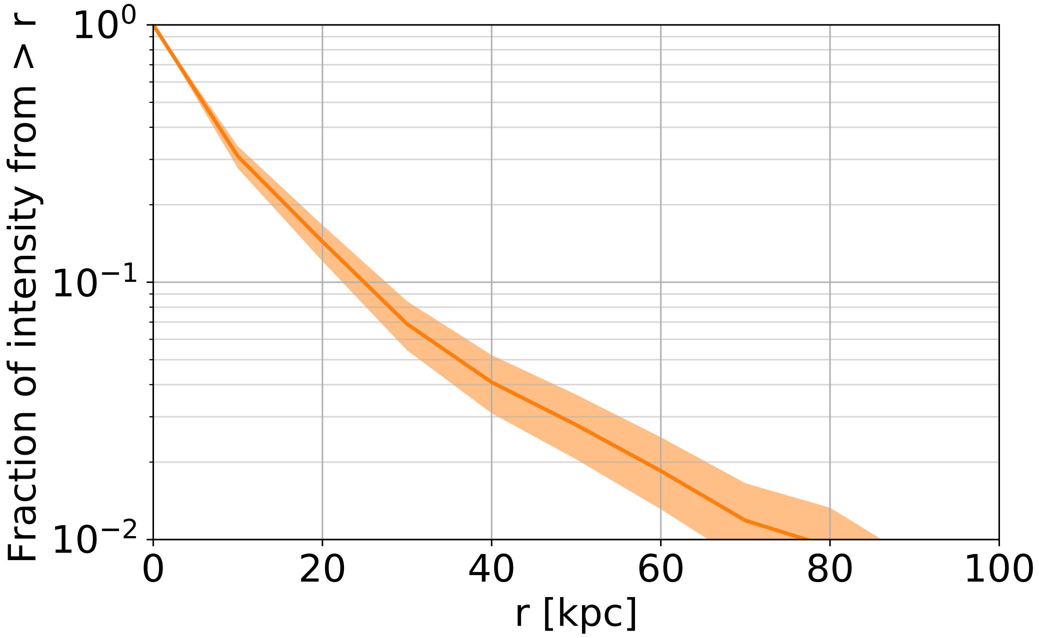

Figure 5.14: The fraction EBL intensity from galaxy extension as a function of $r_{\rm cut}$. This is estimated with the light profile fits from CIBER (this work) and HSC Wang et al. [97], and the stellar mass function from Muzzin et al. [74].

emission from the sum of the galaxy light and one-halo terms. For the IHL, we consider the extended galaxy emission beyond $r_{\rm cut} = 10/20$ kpc of all the bright ($m_{1.1} < 20$) galaxies in the halo, noting that $m_{1.1} = 20$ is also our choice of flux threshold for masking. Therefore, the IHL fraction $f_{\rm IHL}$ can be expressed as

$$f_{\rm IHL} = \frac{\sum_{m_{1.1}<20} L(> r_{\rm cut})}{\sum_{m_{1.1}<20} L + \sum_{\rm faint} L}, \tag{5.24}$$

where $\sum_{m_{1.1}<20} L$ is the total light associate with bright galaxies, and $\sum_{m_{1.1}<20} L(> r_{\rm cut})$ is the part of bright galaxy emission beyond $r_{\rm cut}$. $\sum_{\rm faint} L$ represents the light from faint galaxies as well as the unbound stars in the halo, captured in the one-halo luminosity. Note that we conservatively assume the one-halo luminosity arises entirely from faint, gravitationally bound galaxies. However it is certainly true that some one-halo light arises from unbound stars, as is readily observed in images of massive clusters at low redshift.

From our stacking profile, the faint source emission $\sum_{\rm faint} L$ can be described by the total emission in the one-halo term, L_{1h}[12]. For the bright sources, we define

$$\sum_{m_{1.1}<20} L = L_{\rm gal} \cdot N_{\rm eff}, \tag{5.25}$$

where $L_{\rm gal}$ is the total light in the galaxy profile term from our stacking results, which describes the averaged light of the galaxies within each stacking sample. $N_{\rm eff}$ accounts for the fact that there are multiple bright galaxies in the halo, and we infer the average $N_{\rm eff}$ value from MICECAT. For our five stacking sub-samples, we get $N_{\rm eff} \sim 2$ to 5. From our fitted galaxy profile, we can also calculate $L_{\rm gal}(> r_{\rm cut})$, and we apply the same $N_{\rm eff}$ to model the extension from other bright galaxies:

$$\sum_{m_{1.1}<20} L(> r_{\rm cut}) = L_{\rm gal}(> r_{\rm cut}) \cdot N_{\rm eff}. \tag{5.26}$$

This results in

$$f_{\rm IHL} = \frac{L_{\rm gal}(> r_{\rm cut})/L_{\rm gal}}{1 + L_{1h}/\left(N_{\rm eff} \cdot L_{\rm gal}\right)}. \tag{5.27}$$

We show our constraints on $f_{\rm IHL}$, as a function of halo mass and redshift in Fig. 5.15 and 5.16, respectively. The halo masses associated with our galaxies are inferred from the MICECAT simulation, and using the SDSS photometric redshifts. The CIBER data points shown in Fig. 5.15 and 5.16 are summarized in Table .7.

Note that the fraction of light beyond $r_{\rm cut}$ (the numerator in Eq. 5.27) is shown in Fig. 5.13, where the higher redshift bins have slightly higher values. However, in Fig. 5.16, they have lower $f_{\rm IHL}$. This is due to the increase of the one-halo term with redshift. We show the ratio of one-halo term and the stacked galaxy light in Fig. 5.17. Note that this observable quantity tracks the evolution of the one-halo luminosity, but lacks the $N_{\rm eff}$ term in Eq. 5.27 derived from simulations. We compare with the same quantity from the MICECAT simulation, where the one-halo term includes all the unmasked faint galaxies and residual bright source emission outside the mask due to the PSF. We detect a strong redshift evolution of one-halo contribution compared with the MICECAT simulation, which could be attributed to the unbound stars that are not included in MICECAT.

We compare our results with $f_{\rm IHL}$ from previous work, including the Milky Way [19], the Andromeda Galaxy [M31; 30], the ICL fraction in individual galaxy

---

[12] Our one-halo model also includes the outskirts of bright sources beyond the mask, but we checked that this component is negligible compared to the faint sources using the MICECAT simulation.

groups and clusters [18, 43, 44], and an analytical model [77, 78]. Our results follow a more gradual redshift evolution trend than reported in massive clusters [18] (see Fig. 5.16).

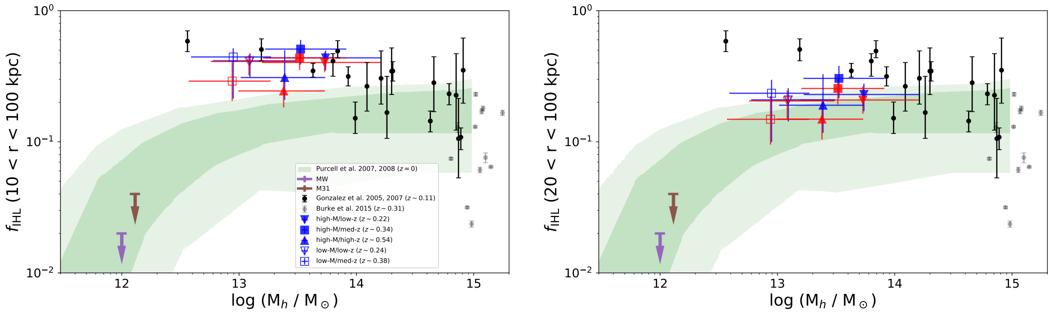

Figure 5.15: The IHL fraction $f_{\rm IHL}$ as a function of halo mass. The IHL is defined by the light beyond a radius $r_{\rm cut}$ around the galaxy. Here we consider three different $r_{\rm cut}$ values: 10 kpc (left) and 20 kpc (right). Blue and red data points show the constraints from this work in the 1.1 μm and 1.8 μm bands, respectively. Dark and light green shaded regions denote the 68% and 95% variations among galaxies from an analytical model at $z = 0$ [77, 78]. The ICL fraction in individual galaxy groups and clusters from Gonzalez et al. [43, 44] and Burke et al. [18] are shown in black and grey data points. The two downward arrows give upper limits for the Milky Way [19] and Andromeda (M31) [30].

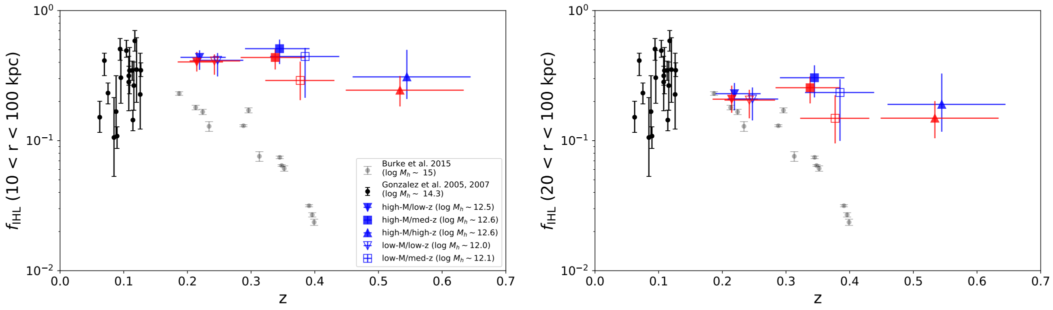

Figure 5.16: $f_{\rm IHL}$ constraints as in Fig. 5.15, but plotted as a function of redshift. The masses of the Burke et al. [18] clusters are 100-1000× the halo masses associated with our galaxies.

Color of the Galaxy Inner and Outer Regions

We calculate the $m_{1.1} - m_{1.8}$ color of the inner and outer region of the galaxy, defined by the total light inside and outside 20 kpc physical scale in the fitted galaxy profile. The results are summarized in Table 5.5. Note the definition of inner and outer component here is based on the intrinsic profile, which is different from the core/extension separation using the stacked PSF defined in Eq. 5.19. We have no

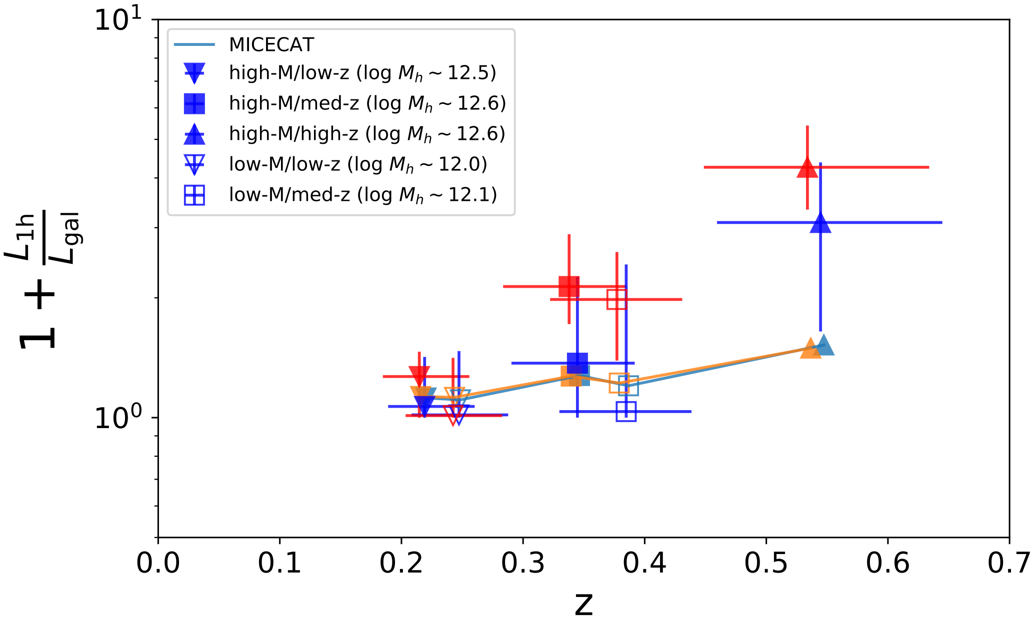

Figure 5.17: The ratio of the total one-halo term and stacked galaxy profile term from our stacking results (blue: 1.1 μm, red: 1.8 μm) compared with the MICECAT simulation (light blue: 1.1 μm, orange: 1.8 μm). We observe a somewhat stronger evolution, causing the fall-off of f_{IHL} with redshift seen in Fig. 5.16.

detection of a color difference between the inner and outer regions in the two CIBER bands. We also find similar inner and outer region color with 10 kpc radius cut. Previous measurements in optical bands found that the galaxy outskirts are bluer than their core [e.g., 34, 51]. For comparison, we calculate the $m_{1.1} - m_{1.8}$ color of galaxy cores in MICECAT sources selected from the same criteria, as well as from the empirical galaxy model of Helgason et al. [48] at $z = 0.3$, approximately the redshift of our samples. Our inner region color is consistent with these models. To model the extension, we use a collection of elliptical galaxy spectra from the population synthesis package GISSEL [16] redshifted to $z = 0.3$. We also estimate the extension color using an imaging study on the local spiral galaxy NGC 5907 [84]. We use their ratio of I band and J band flux in >1 arcmin regions to approximate the $m_{1.1} - m_{1.8}$ extension color. The rest-frame I and J band redshifted to $z \sim 0.3$ (approximately the redshift of our samples) are close to the two CIBER bands. NGC 5907 shows a redder spectrum than our galaxy extension, whereas the elliptical galaxy spectrum template is slightly bluer than our samples. In addition, the IHL constraints from Zemcov et al. [101] are also given in Table 5.5, but we note that Zemcov et al. [101] reflects the integrated IHL from all redshifts.

Table 5.5: Constraints on the color ($m_{1.1} - m_{1.8}$) of the galaxy inner and outer components. The +/− values indicate 68% interval ranges. The total row shows the weighted average of five sub-samples. For comparison, we also show models of core color from MICECAT and an analytical prescription from Helgason et al. [48] at $z = 0.3$. For the extension, we compare our results with spectra from a population synthesis code, GISSEL [16], and the outskirts of NGC 5907 redshifted to $z = 0.3$ [84]. The color of EBL fluctuations attributed to redshift-integrated IHL from Zemcov et al. [101] is also shown.

Name	Inner	Outer
high-M/low-z	$0.42^{+0.20}_{-0.17}$	$0.36^{+0.34}_{-0.31}$
high-M/med-z	$0.54^{+0.25}_{-0.27}$	$0.46^{+0.24}_{-0.25}$
high-M/high-z	$0.65^{+0.31}_{-0.28}$	$0.61^{+0.31}_{-0.25}$
low-M/low-z	$0.39^{+0.20}_{-0.18}$	$0.37^{+0.41}_{-0.37}$
low-M/med-z	$0.56^{+0.23}_{-0.24}$	$0.44^{+0.50}_{-0.44}$
total	$0.49^{+0.10}_{-0.10}$	$0.47^{+0.14}_{-0.15}$
MICECAT	0.44 ± 0.07	
Helgason et al. (2012)	0.41	
GISSEL		0.32 ± 0.08
NGC 5907		1.41 ± 0.61
Zemcov et al. (2014)		$0.89^{+1.17}_{-1.08}$

One-halo and Two-halo Clustering

The one-halo amplitude is detected in the 1.8 μm band at the $\sim 4\sigma$ level in the "total" and "high-M/high-z" cases, and at the $\sim 3\sigma$ level in "mag bin #4" and "high-M/med-z" cases. One-halo clustering is not clearly detected at the 1.1 μm band since the photocurrent from sources is lower in this band. The one-halo amplitude $A_{\rm 1h}$ is consistent with unity to within $\sim 2\sigma$, which implies that our one-halo templates built from MICECAT are sufficient to describe the clustering within halos of our stacked samples. However, from our stacking results, it is unclear if this emission actually consists of discrete galaxies as given in the MICECAT simulation. Two-halo clustering is not detected in all cases since the large-scale clustering signal is comparable to the current uncertainties in the measurement.

5.12 Conclusions

By stacking galaxies from CIBER imaging data in two near-infrared bands (1.1 and 1.8 μm), we detect extended emission in galaxies. The galaxies being stacked ($\sim 30,000$ galaxies in total) are split into five sub-samples from SDSS spanning redshifts $0.2 \lesssim z \lesssim 0.5$ and stellar masses $10^{10.5} M_\odot \lesssim M_* \lesssim 10^{12} M_\odot$, comparable to L_* galaxies at this redshift. We jointly fit a model for the inherent galaxy light profile and large-scale one- and two-halo clustering.

With the galaxy profile, we estimate that ∼ 20% of total light is missing in galaxy photometry due to the use of limited apertures, in agreement with previous estimates from the literature. We do not detect a 1.1-1.8 μm color difference in the inner and outer region of our galaxy samples.

While we do not detect two-halo clustering, we detect one-halo clustering in the 1.8 μm band at 4-σ significance over the full sample of galaxies. These results suggest non-linear clustering could have a significant impact on modeling the IHL, but is not accounted for in previous fluctuation analysis by Zemcov et al. [101]. An IHL fluctuation model with one-halo clustering [e.g., 36] is needed to fully account for the non-linear clustering in IHL modeling.

The intrinsic galaxy profile fitted from our stacking analysis suggests ∼ 50%/25% of the total galaxy light resides at $r > 10/20$ kpc, respectively. This result is in agreement with previous HSC measurements at lower redshifts ($0 \lesssim z \lesssim 0.25$) and lower stellar masses ($10^{9.2} M_\odot < M_* < 10^{11.4} M_\odot$). The galaxy extension accounts for significant fraction of luminosity in L_* galaxies, but falls off below $M_* \sim 10^{11} M_\odot$. We extrapolate the fraction of extended galaxy light we measure to all galaxy masses scales and assuming a Schechter luminosity function, we find ∼ 30%/15% of the total galaxy light are from $r > 10/20$ kpc, respectively. We measure a moderate increase in f_{IHL} with cosmic time, which we attribute to the decrease in one-halo contribution within the dark matter halo of our stacked samples. The previous fluctuation study using CIBER data [101] found that the IHL has comparable intensity to the IGL in the near-infrared EBL. While our study cannot constrain the whole IHL contribution to the EBL since we only study galaxies from a certain range of redshift and masses, our results suggest that $\sim L_*$ galaxy at $0.2 \lesssim z \lesssim 0.5$ have an extended light profile which contributes appreciable IHL to their host halos. As $\sim L_*$ galaxies are the representative population, which contain most of the IGL emission, the flux from the extension, and the one-halo term present in our galaxy samples, both need to be properly accounted for in future EBL photometry and fluctuation measurements.

.1 Extension and IHL Fraction

Table .6 summarize the fraction of light beyond 10 and 20 kpc, assuming our fitted light profile. These are the data presented in Fig. 5.13.

Table .7 summarize the f_{IHL} values with $r_{\mathrm{cut=}}$ 10 and 20 kpc, assuming our fitted light profile and the one-halo contribution from the MICECAT. These are the data

Table .6: Fraction of galaxy flux between 10/20 kpc and 100 kpc, assuming the galaxy light profile follows the stacking results in this work. These are the values shown in Fig. 5.13. The total row shows the weighted average of the five listed sub-samples.

Name	1.1 μm 10 kpc	1.1 μm 20 kpc	1.8 μm 10 kpc	1.8 μm 20 kpc
high-M/low-z	$0.44^{+0.05}_{-0.05}$	$0.23^{+0.04}_{-0.04}$	$0.43^{+0.05}_{-0.04}$	$0.22^{+0.04}_{-0.04}$
high-M/med-z	$0.53^{+0.05}_{-0.04}$	$0.31^{+0.05}_{-0.04}$	$0.52^{+0.05}_{-0.04}$	$0.30^{+0.04}_{-0.04}$
high-M/high-z	$0.55^{+0.07}_{-0.05}$	$0.33^{+0.06}_{-0.05}$	$0.55^{+0.05}_{-0.04}$	$0.33^{+0.05}_{-0.04}$
low-M/low-z	$0.41^{+0.06}_{-0.05}$	$0.21^{+0.04}_{-0.04}$	$0.41^{+0.05}_{-0.05}$	$0.20^{+0.04}_{-0.04}$
low-M/med-z	$0.45^{+0.08}_{-0.06}$	$0.23^{+0.06}_{-0.05}$	$0.42^{+0.06}_{-0.05}$	$0.21^{+0.05}_{-0.04}$
total	$0.48^{+0.02}_{-0.03}$	$0.25^{+0.02}_{-0.02}$	$0.47^{+0.02}_{-0.02}$	$0.25^{+0.02}_{-0.02}$

Table .7: IHL fraction (Eq. 5.27) with $r_{\mathrm{cut}=}$ 10/20 kpc, assuming the galaxy light profile and the one-halo terms follow our stacking results and the MICECAT simulation, respectively. These are the values shown in Fig. 5.15 and 5.16. The total row shows the weighted average of the five listed sub-samples.

Name	1.1 μm 10 kpc	1.1 μm 20 kpc	1.8 μm 10 kpc	1.8 μm 20 kpc
high-M/low-z	$0.44^{+0.09}_{-0.06}$	$0.23^{+0.06}_{-0.05}$	$0.40^{+0.06}_{-0.08}$	$0.21^{+0.04}_{-0.06}$
high-M/med-z	$0.51^{+0.12}_{-0.09}$	$0.30^{+0.09}_{-0.08}$	$0.44^{+0.08}_{-0.07}$	$0.26^{+0.06}_{-0.06}$
high-M/high-z	$0.31^{+0.10}_{-0.19}$	$0.19^{+0.07}_{-0.14}$	$0.24^{+0.06}_{-0.07}$	$0.15^{+0.04}_{-0.05}$
low-M/low-z	$0.41^{+0.10}_{-0.06}$	$0.21^{+0.07}_{-0.05}$	$0.41^{+0.09}_{-0.05}$	$0.21^{+0.06}_{-0.04}$
low-M/med-z	$0.44^{+0.23}_{-0.07}$	$0.23^{+0.14}_{-0.06}$	$0.29^{+0.09}_{-0.12}$	$0.15^{+0.05}_{-0.07}$
total	$0.43^{+0.03}_{-0.05}$	$0.23^{+0.03}_{-0.03}$	$0.36^{+0.03}_{-0.05}$	$0.19^{+0.02}_{-0.02}$

presented in Fig. 5.15 and 5.16.

References

[1] Abdalla, H., Adam, R., Aharonian, F., et al. 2020, MNRAS, 494, 5590, doi: `10.1093/mnras/staa999`

[2] Abdo, A. A., Ackermann, M., Ajello, M., et al. 2010, ApJ, 723, 1082, doi: `10.1088/0004-637X/723/2/1082`

[3] Abell, G. O. 1958, ApJS, 3, 211, doi: `10.1086/190036`

[4] Abeysekara, A. U., Archer, A., Benbow, W., et al. 2019, ApJ, 885, 150, doi: `10.3847/1538-4357/ab4817`

[5] Abraham, R. G., & van Dokkum, P. G. 2014, PASP, 126, 55, doi: `10.1086/674875`

[6] Acciari, V. A., Ansoldi, S., Antonelli, L. A., et al. 2019, MNRAS, 486, 4233, doi: `10.1093/mnras/stz943`

[7] Ackermann, M., Ajello, M., Allafort, A., et al. 2012, Science, 338, 1190, doi: `10.1126/science.1227160`

[8] Ackermann, M., Ajello, M., Baldini, L., et al. 2018, ApJS, 237, 32, doi: `10.3847/1538-4365/aacdf7`

[9] Aharonian, F., Akhperjanian, A. G., Bazer-Bachi, A. R., et al. 2006, Nature, 440, 1018, doi: `10.1038/nature04680`

[10] Aharonian, F., Akhperjanian, A. G., Barres de Almeida, U., et al. 2007, A&A, 475, L9, doi: `10.1051/0004-6361:20078462`

[11] Aihara, H., Allende Prieto, C., An, D., et al. 2011, ApJS, 193, 29, doi: `10.1088/0067-0049/193/2/29`

[12] Arnouts, S., Cristiani, S., Moscardini, L., et al. 1999, MNRAS, 310, 540, doi: `10.1046/j.1365-8711.1999.02978.x`

[13] Bernstein, R. A. 2007, ApJ, 666, 663, doi: `10.1086/519824`

[14] Blanton, M. R., Bershady, M. A., Abolfathi, B., et al. 2017, AJ, 154, 28, doi: `10.3847/1538-3881/aa7567`

[15] Bock, J., Sullivan, I., Arai, T., et al. 2013, ApJS, 207, 32, doi: `10.1088/0067-0049/207/2/32`

[16] Bruzual A., G., & Charlot, S. 1993, ApJ, 405, 538, doi: `10.1086/172385`

[17] Bullock, J. S., & Johnston, K. V. 2005, ApJ, 635, 931, doi: `10.1086/497422`

[18] Burke, C., Hilton, M., & Collins, C. 2015, MNRAS, 449, 2353, doi: `10.1093/mnras/stv450`

[19] Carollo, D., Beers, T. C., Chiba, M., et al. 2010, ApJ, 712, 692, doi: `10.1088/0004-637X/712/1/692`

[20] Carretero, J., Castander, F. J., Gaztañaga, E., Crocce, M., & Fosalba, P. 2015, MNRAS, 447, 646, doi: `10.1093/mnras/stu2402`

[21] Cavaliere, A., & Fusco-Femiano, R. 1978, A&A, 70, 677

[22] Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, arXiv e-prints, arXiv:1612.05560. `https://arxiv.org/abs/1612.05560`

[23] Conroy, C., Wechsler, R. H., & Kravtsov, A. V. 2007, ApJ, 668, 826, doi: `10.1086/521425`

[24] Cooper, A. P., D'Souza, R., Kauffmann, G., et al. 2013, MNRAS, 434, 3348, doi: 10.1093/mnras/stt1245

[25] Cooper, A. P., Parry, O. H., Lowing, B., Cole, S., & Frenk, C. 2015, MNRAS, 454, 3185, doi: 10.1093/mnras/stv2057

[26] Cooper, A. P., Cole, S., Frenk, C. S., et al. 2010, MNRAS, 406, 744, doi: 10.1111/j.1365-2966.2010.16740.x

[27] Cooray, A. 2016, Royal Society Open Science, 3, 150555, doi: 10.1098/rsos.150555

[28] Cooray, A., & Sheth, R. 2002, PhysRep, 372, 1, doi: 10.1016/S0370-1573(02)00276-4

[29] Cooray, A., Smidt, J., de Bernardis, F., et al. 2012, Nature, 490, 514, doi: 10.1038/nature11474

[30] Courteau, S., Widrow, L. M., McDonald, M., et al. 2011, ApJ, 739, 20, doi: 10.1088/0004-637X/739/1/20

[31] Crocce, M., Castander, F. J., Gaztañaga, E., Fosalba, P., & Carretero, J. 2015, MNRAS, 453, 1513, doi: 10.1093/mnras/stv1708

[32] Domínguez, A., Primack, J. R., Rosario, D. J., et al. 2011, MNRAS, 410, 2556, doi: 10.1111/j.1365-2966.2010.17631.x

[33] Driver, S. P., Andrews, S. K., Davies, L. J., et al. 2016, ApJ, 827, 108, doi: 10.3847/0004-637X/827/2/108

[34] D'Souza, R., Kauffman, G., Wang, J., & Vegetti, S. 2014, MNRAS, 443, 1433, doi: 10.1093/mnras/stu1194

[35] Elias, L. M., Sales, L. V., Creasey, P., et al. 2018, MNRAS, 479, 4004, doi: 10.1093/mnras/sty1718

[36] Fernandez, E. R., Komatsu, E., Iliev, I. T., & Shapiro, P. R. 2010, ApJ, 710, 1089, doi: 10.1088/0004-637X/710/2/1089

[37] Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306, doi: 10.1086/670067

[38] Fosalba, P., Crocce, M., Gaztañaga, E., & Castand er, F. J. 2015, MNRAS, 448, 2987, doi: 10.1093/mnras/stv138

[39] Fosalba, P., Gaztañaga, E., Castander, F. J., & Crocce, M. 2015, MNRAS, 447, 1319, doi: 10.1093/mnras/stu2464

[40] Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, A&A, 595, A1, doi: 10.1051/0004-6361/201629272

[41] Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, A&A, 616, A1, doi: 10.1051/0004-6361/201833051

[42] Gelman, A., & Rubin, D. B. 1992, Statistical Science, 7, 457, doi: 10.1214/ss/1177011136

[43] Gonzalez, A. H., Zabludoff, A. I., & Zaritsky, D. 2005, ApJ, 618, 195, doi: 10.1086/425896

[44] Gonzalez, A. H., Zaritsky, D., & Zabludoff, A. I. 2007, ApJ, 666, 147, doi: 10.1086/519729

[45] Graham, A. W., Driver, S. P., Petrosian, V., et al. 2005, AJ, 130, 1535, doi: 10.1086/444475

[46] H. E. S. S. Collaboration, Abdalla, H., Abramowski, A., et al. 2017, A&A, 606, A59, doi: 10.1051/0004-6361/201731200

[47] Hartlap, J., Simon, P., & Schneider, P. 2007, A&A, 464, 399, doi: 10.1051/0004-6361:20066170

[48] Helgason, K., Ricotti, M., & Kashlinsky, A. 2012, ApJ, 752, 113, doi: 10.1088/0004-637X/752/2/113

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

[50] Hoffmann, K., Bel, J., Gaztañaga, E., et al. 2015, MNRAS, 447, 1724, doi: 10.1093/mnras/stu2492

[51] Huang, S., Leauthaud, A., Greene, J. E., et al. 2018, MNRAS, 475, 3348, doi: 10.1093/mnras/stx3200

[52] Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, A&A, 457, 841, doi: 10.1051/0004-6361:20065138

[53] Kashlinsky, A., Arendt, R. G., Ashby, M. L. N., et al. 2012, ApJ, 753, 63, doi: 10.1088/0004-637X/753/1/63

[54] Kashlinsky, A., Arendt, R. G., Mather, J., & Moseley, S. H. 2005, Nature, 438, 45, doi: 10.1038/nature04143

[55] Kawara, K., Matsuoka, Y., Sano, K., et al. 2017, PASJ, 69, 31, doi: 10.1093/pasj/psx003

[56] Keenan, R. C., Barger, A. J., Cowie, L. L., & Wang, W. H. 2010, ApJ, 723, 40, doi: 10.1088/0004-637X/723/1/40

[57] Kim, M. G., Matsumoto, T., Lee, H. M., et al. 2019, PASJ, 71, 82, doi: 10.1093/pasj/psz063

[58] Korngut, P. M., Renbarger, T., Arai, T., et al. 2013, ApJS, 207, 34, doi: 10.1088/0067-0049/207/2/34

[59] Koushan, S., Driver, S. P., Bellstedt, S., et al. 2021, arXiv e-prints, arXiv:2102.12323. https://arxiv.org/abs/2102.12323

[60] Lang, D., Hogg, D. W., Mierle, K., Blanton, M., & Roweis, S. 2010, AJ, 139, 1782, doi: 10.1088/0004-6256/139/5/1782

[61] Lauer, T. R., Postman, M., Weaver, H. A., et al. 2020, arXiv e-prints, arXiv:2011.03052. https://arxiv.org/abs/2011.03052

[62] Levenson, L. R., Wright, E. L., & Johnson, B. D. 2007, ApJ, 666, 34, doi: 10.1086/520112

[63] Lin, Y.-T., & Mohr, J. J. 2004, ApJ, 617, 879, doi: 10.1086/425412

[64] Lucy, L. B. 1974, AJ, 79, 745, doi: 10.1086/111605

[65] MAGIC Collaboration, Albert, J., Aliu, E., et al. 2008, Science, 320, 1752, doi: 10.1126/science.1157087

[66] Martínez-Delgado, D., Gabany, R. J., Crawford, K., et al. 2010, AJ, 140, 962, doi: 10.1088/0004-6256/140/4/962

[67] Matsumoto, T., Kim, M. G., Pyo, J., & Tsumura, K. 2015, ApJ, 807, 57, doi: 10.1088/0004-637X/807/1/57

[68] Matsumoto, T., & Tsumura, K. 2019, PASJ, 71, 88, doi: 10.1093/pasj/psz070

[69] Matsumoto, T., Tsumura, K., Matsuoka, Y., & Pyo, J. 2018, AJ, 156, 86, doi: 10.3847/1538-3881/aad0f0

[70] Matsumoto, T., Seo, H. J., Jeong, W. S., et al. 2011, ApJ, 742, 124, doi: 10.1088/0004-637X/742/2/124

[71] Matsuura, S., Arai, T., Bock, J. J., et al. 2017, ApJ, 839, 7, doi: 10.3847/1538-4357/aa6843

[72] Mazzarini, M., Just, A., Macciò, A. V., & Moetazedian, R. 2020, A&A, 636, A106, doi: 10.1051/0004-6361/202037558

[73] Mitchell-Wynne, K., Cooray, A., Gong, Y., et al. 2015, Nature Communications, 6, 7945, doi: 10.1038/ncomms8945

[74] Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013, ApJ, 777, 18, doi: 10.1088/0004-637X/777/1/18

[75] Petrosian, V. 1976, ApJL, 210, L53, doi: 10.1086/182301

[76] Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 594, A13, doi: 10.1051/0004-6361/201525830

[77] Purcell, C. W., Bullock, J. S., & Zentner, A. R. 2007, ApJ, 666, 20, doi: 10.1086/519787

[78] —. 2008, MNRAS, 391, 550, doi: 10.1111/j.1365-2966.2008.13938.x

[79] Richardson, W. H. 1972, Journal of the Optical Society of America (1917-1983), 62, 55

[80] Rodriguez-Gomez, V., Pillepich, A., Sales, L. V., et al. 2016, MNRAS, 458, 2371, doi: 10.1093/mnras/stw456

[81] Rowan-Robinson, M., Gonzalez-Solares, E., Vaccari, M., & Marchetti, L. 2013, MNRAS, 428, 1958, doi: 10.1093/mnras/sts163

[82] Rowan-Robinson, M., Babbedge, T., Oliver, S., et al. 2008, MNRAS, 386, 697, doi: 10.1111/j.1365-2966.2008.13109.x

[83] Rudick, C. S., Mihos, J. C., Frey, L. H., & McBride, C. K. 2009, ApJ, 699, 1518, doi: 10.1088/0004-637X/699/2/1518

[84] Rudy, R. J., Woodward, C. E., Hodge, T., Fairfield, S. W., & Harker, D. E. 1997, Nature, 387, 159, doi: 10.1038/387159a0

[85] Saldana-Lopez, A., Domínguez, A., Pérez-González, P. G., et al. 2020, arXiv e-prints, arXiv:2012.03035. https://arxiv.org/abs/2012.03035

[86] Sano, K., Kawara, K., Matsuura, S., et al. 2015, ApJ, 811, 77, doi: 10.1088/0004-637X/811/2/77

[87] Sano, K., Matsuura, S., Yomo, K., & Takahashi, A. 2020, ApJ, 901, 112, doi: 10.3847/1538-4357/abad3d

[88] Seo, H. J., Lee, H. M., Matsumoto, T., et al. 2015, ApJ, 807, 140, doi: 10.1088/0004-637X/807/2/140

[89] Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163, doi: 10.1086/498708

[90] Symons, T., Zemcov, M., Bock, J., et al. 2021, ApJS, 252, 24, doi: 10.3847/1538-4365/abcaa5

[91] Tal, T., & van Dokkum, P. G. 2011, ApJ, 731, 89, doi: 10.1088/0004-637X/731/2/89

[92] Tal, T., van Dokkum, P. G., Nelan, J., & Bezanson, R. 2009, AJ, 138, 1417, doi: 10.1088/0004-6256/138/5/1417

[93] Thompson, R. I., Eisenstein, D., Fan, X., Rieke, M., & Kennicutt, R. C. 2007, ApJ, 666, 658, doi: 10.1086/520634

[94] Tsumura, K., Matsumoto, T., Matsuura, S., Sakon, I., & Wada, T. 2013, PASJ, 65, 121, doi: 10.1093/pasj/65.6.121

[95] Tsumura, K., Arai, T., Battle, J., et al. 2013, ApJS, 207, 33, doi: 10.1088/0067-0049/207/2/33

[96] van Dokkum, P. G., Abraham, R., & Merritt, A. 2014, ApJL, 782, L24, doi: 10.1088/2041-8205/782/2/L24

[97] Wang, W., Han, J., Sonnenfeld, A., et al. 2019, MNRAS, 487, 1580, doi: 10.1093/mnras/stz1339

[98] Wen, Z. L., Han, J. L., & Liu, F. S. 2012, ApJS, 199, 34, doi: 10.1088/0067-0049/199/2/34

[99] Zemcov, M., Immel, P., Nguyen, C., et al. 2017, Nature Communications, 8, 15003, doi: 10.1038/ncomms15003

[100] Zemcov, M., Arai, T., Battle, J., et al. 2013, ApJS, 207, 31, doi: 10.1088/0067-0049/207/2/31

[101] Zemcov, M., Smidt, J., Arai, T., et al. 2014, Science, 346, 732, doi: 10.1126/science.1258168

[102] Zhang, Y., Yanny, B., Palmese, A., et al. 2019, ApJ, 874, 165, doi: 10.3847/1538-4357/ab0dfd

[103] Zibetti, S., White, S. D. M., Schneider, D. P., & Brinkmann, J. 2005, MNRAS, 358, 949, doi: 10.1111/j.1365-2966.2005.08817.x

[104] Zu, Y., & Mandelbaum, R. 2015, MNRAS, 454, 1161, doi: 10.1093/mnras/stv2062

Chapter 6

COSMIC NEAR-INFRARED BACKGROUND TOMOGRAPHY WITH SPHEREX USING GALAXY CROSS-CORRELATIONS

Cheng, Y.-T., & Chang, T.-C. 2021, in prep.

The extragalactic background light (EBL) consists of integrated light from all sources of emission throughout the history of the Universe. At near-infrared wavelengths, the EBL is dominated by stellar emission across cosmic time; however, spectral and redshift information of emitting sources cannot be probed in absolute photometry or fluctuation measurements. Cross-correlating near-infrared maps with tracers of known redshift enables EBL redshift tomography, as EBL emission will only correlate with external tracers from the same redshift. Here we forecast the sensitivity of probing the EBL spectral energy distribution as a function of redshift by cross-correlating the upcoming near-infrared spectro-imaging survey, SPHEREx, with several current and future galaxy redshift surveys. Using a model galaxy luminosity function, and the cross-power spectrum clustering amplitude on large scales, we forecast that the redshift-dependent near-infrared EBL spectrum can be detected out to $z \sim 6$. We also predict a high significance measurement ($\sim 10^2 - 10^4\sigma$) of the small-scale cross-power spectrum out to $z \sim 10$. The large-scale cross power spectrum amplitudes is proportional to the redshift-dependent EBL, which can constrain the cosmic evolution of the stellar synthesis process through both continuum and the line emission. On the non-linear and the Poisson-noise scales, the high sensitivity measurements can used to probe the non-linear clustering and the mean spectra associate with the tracer population across redshift.

6.1 Introduction

The extragalactic background light (EBL) is the aggregate light from all sources of emission across cosmic time. EBL measurements have been made from gamma-rays to radio [see e.g., 22, 70, for recent reviews]. At near-infrared wavelengths, the EBL is mostly produced by redshifted ultraviolet and optical stellar emission, and thus carries essential information on the history of stellar synthesis processes in our universe. However, observations also suggest other sources of the near-infrared EBL, including diffuse light in the dark matter halos from stripped stars, sometimes referred to as intra-halo light [IHL, 23, 109], and primordial stars and galaxies from the epoch of reionization [47, 48, 68, 75]. Theories also proposed possible near-infrared EBL emission from direct collapse black holes from the dark ages [105, 106] and the decay of axion-like particles as dark matter candidate [e.g., 14, 40].

Extensive studies have attempted to probe the near-infrared EBL through different methods. Galaxy counts measure the emission from resolved galaxies and extrapolate their observations to estimate the contributions from faint galaxies below the detection limit [31, 36, 44, 49, 53, 86]. This sets a lower bound on the EBL by the integrated galaxy light (IGL) component, while EBL from the diffuse light or sources not associated with galaxies are not included. Direct measurements using absolute photometry captures all emission in the EBL [11, 56, 58, 66, 69, 87, 88, 100, 108]. However, absolute photometry is challenging since the systematic errors and foreground emission have to be tightly controlled. Fluctuation analysis is an alternative approach that is less susceptible to these systematics, since different sources of emission have distinct spatial and spectral correlations [23, 47, 48, 52, 67, 68, 75, 93, 98, 109], but deriving the absolute intensity of the EBL from fluctuation measurements depends on model assumptions. Another limitation of absolute photometry and fluctuations is the redshift resolution, as the measured intensity is a projection of all emission along the line of sight. One way to infer the redshift dependency of the EBL is to measure the opacity of gamma-ray photons from individual blazars. Near-infrared photons along the line of sight will interact with gamma-ray photons by pair production, and thus the EBL can be constrained by absorption features in blazar spectra [e.g., 1–8, 43, 63]. This method enables a redshift tomography of the EBL by observing blazars from different distances. However, these estimates have low spectral resolution and depend on the assumption of intrinsic blazar spectra.

As the EBL intensity is a mixture of emission from all redshifts, and different sources have distinct spectral features, it will be insightful to decompose the EBL signal by its redshift and spectral dependencies. Cross correlation provides a path to perform EBL tomography while retaining spectral information. By cross-correlating external sources (e.g., a galaxy catalog) that trace underlying large-scale structures from a certain redshift, one can extract the EBL from the IGL and other components associated with the large-scale structures, such as the IHL, from the same cosmic volume. The cross-correlation technique has previously been applied to derive the redshift distribution of photometric catalogs [71, 72, 77], perform redshift tomography of broadband imaging surveys [18, 19, 91], probe the cosmic infrared background [94] and the thermal or kinematic Sunyaev-Zel'dovich signal from the cosmic microwave background [e.g., 17, 45]. Cross correlation has also been proposed to probe the direct collapse blackholes [13, 74] and the axion decay [25] in the optical to near-infrared EBL. In line intensity mapping [see 54, for a review], where the emission from a certain spectral line is used to trace the three-dimensional large-scale structure, cross correlation can also probe line emission from 21 cm [e.g., 15, 65], CO [e.g., 81], [C II] [e.g., 83, 104], and Lyα [e.g., 26, 27].

SPHEREx [33, 34] is an approved NASA MIDEX mission that will carry out an all-sky spectro-imaging survey at near-infrared wavelengths (see Sec. 6.5 for details). SPHEREx will achieve unparalleled spectral resolution, sensitivity, and sky coverage for spectrally mapping the near-infrared sky, which provides an exceptional dataset for studying EBL through fluctuation analysis.

Next generation spectroscopic and photometric galaxy surveys will come online in the next few years, including DESI [30], the Rubin Observatory LSST [59], Euclid [9], and the Roman Space Telescope [97]. These surveys will provide galaxy catalogs in unprecedented depth and synergy with the SPHEREx spectro-imaging dataset. In this work, we forecast constraints on EBL enabled by cross-correlations of SPHEREx and upcoming galaxy surveys. SPHEREx spectro-images encode the spatial and spectral information of the EBL, whereas galaxies trace the three-dimensional cosmic structures with redshift information. With cross-correlation, we perform redshift tomography of the EBL spectrum, which constrains the redshift composition and spectral response of the integrated background emission, as well as its scale dependence. These measurements primarily consist of the IGL, and will provide valuable insights on the cosmic star formation and stellar mass history. Furthermore, the information on the IHL or other possible EBL emitting sources that trace the large-scale structure (e.g. the decaying dark matter candidates) is also encoded in the cross correlation measurements.

A recent study by Scott et al. [92] investigated the ultraviolet to optical EBL constraints from cross correlation, using images from the future ultraviolet survey Cosmological Advanced Survey Telescope for Optical-ultraviolet Research [CASTOR; 24] for lower redshift, and SPHEREx for the epoch of reionization. Their study focuses on the UV broadband emission and the Lyα line. In this work, we conduct a comprehensive forecast of constraining the EBL spectra from SPHEREx with all accessible wavelength and redshift ranges.

This paper is organized as follows. We first introduce the redshift tomography formalism of cross-power spectrum estimation in Sec. 6.2. We model components of the EBL signal in Sec. 6.3. Sec. 6.4 details the evaluation metrics and the fiducial cases we choose to present the results. Sec. 6.5 introduces SPHEREx, and Sec. 6.6 describes the galaxy surveys considered in this work. The results are presented in Sec. 6.7. Discussion and further science interpretations are given in Sec. 6.8 and Sec. 6.9, respectively. Sec. 6.10 concludes the paper. Throughout this work, we assume a flat ΛCDM cosmology with $n_s = 0.97$, $\sigma_8 = 0.82$, $\Omega_m = 0.26$, $\Omega_b = 0.049$, $\Omega_\Lambda = 0.69$, and $h = 0.68$, consistent with the measurement from *Planck* [80]. All fluxes are quoted in the AB magnitude system.

6.2 Formalism

We detail the analytical expression for a cross power spectrum, and its error estimation, between a continuous density field (such as a diffuse intensity map) and a discrete density tracer sample (such as a galaxy catalog). An intensity field I and a galaxy tracer field g can each be expanded in terms of spherical harmonics $Y_{\ell m}$ as

$$\delta I\left(\hat{n}\right) = \sum_{\ell,m} a^I_{\ell m} Y_{\ell m}\left(\hat{n}\right), \tag{6.1}$$

and

$$\delta g\left(\hat{n}\right) = \sum_{\ell,m} a^g_{\ell m} Y_{\ell m}\left(\hat{n}\right). \tag{6.2}$$

The angular cross power spectrum of these two fields, $C^\times_\ell$, binned in a width of $\Delta\ell$, is given by

$$C^\times_\ell = \left\langle \frac{1}{\left(2\ell+1\right) f_{\rm sky}} \sum_{m=-\ell}^{\ell} \left(a^I_{\ell m}\right)^* a^g_{\ell m} \right\rangle_{\ell\in\Delta\ell}, \tag{6.3}$$

with a variance of

$$\left(\delta C^\times_\ell\right)^2 = \frac{1}{N_\ell}\left[(C^\times_\ell)^2 + C^I_\ell C^g_\ell\right], \tag{6.4}$$

where $N_\ell = \Delta\ell\left(2\ell+1\right) f_{\rm sky}$ is the number of ℓ modes in the band power spectrum at ℓ with a width of $\Delta\ell$, and $f_{\rm sky}$ is the fraction of sky area used in the cross correlation. C^g_ℓ, C^I_ℓ, and $C^\times_\ell$ are the galaxy auto power spectrum, intensity field auto power spectrum, and their cross power spectrum, respectively. Eq. 6.4 is the Gaussian variance that assumes all the multipole modes are independent. This is a good approximation for large-scales, but on the small scales, the non-Gaussian terms that correlate the modes become non-negligible. Therefore, we also consider the trispectrum term, and the cross power spectrum covariance is

$$\mathrm{Cov}\left(C_\ell^{\mathrm{x}}, C_{\ell'}^{\mathrm{x}}\right) = \delta^K_{\ell\ell'}\left(\delta C_\ell^{\mathrm{x}}\right)^2 + \frac{\mathcal{T}^{\mathrm{xx}}_{\ell,-\ell,\ell',-\ell'}}{\Omega_{\mathrm{sur}}}, \tag{6.5}$$

where δ^K is the Kronecker delta, $\Omega_{\mathrm{sur}} = 4\pi f_{\mathrm{sky}}$ is the the survey area, and $\mathcal{T}^{\mathrm{xx}}_{\ell,-\ell,\ell',-\ell'}$ is the trispectrum of the field [89].

Galaxy Sample Auto Power Spectrum C_ℓ^g

We describe a galaxy auto power spectrum as comprised of clustering and (shot) Poisson-noise terms:

$$C_\ell^g = C_{\ell,\mathrm{clus}}^g + C_{\ell,\mathrm{shot}}^g. \tag{6.6}$$

Using the Limber approximation, the clustering term can be expressed as

$$C_{\ell,\mathrm{clus}}^g = \int dz \frac{H(z)}{c}\frac{f^g(z)f^g(z)}{\chi^2(z)} b_g^2(z) P_{\mathrm{m}}\left(k = \frac{\ell+\frac{1}{2}}{\chi(z)}, z\right), \tag{6.7}$$

where H, χ, c, $f^g(z)$ are the Hubble parameter, co-moving distance, speed of light, and the galaxy selection function, respectively. P_m is the matter power spectrum. In this work, we focus on the linear clustering regime and restrict our analysis to large scales at $k < 0.2$ h/Mpc. We assume b_g is a scale-independent linear bias. In practice, with an assumed cosmological model, b_g can be measured from the amplitude of the galaxy auto spectrum. We consider the limit where galaxies are from a narrow redshift bin ($\Delta z^g \ll 1$) centered at z^g, and use the following approximation:

$$f^g(z) = \begin{cases} \frac{1}{\Delta z^g} & \text{if} \quad z_g - \frac{\Delta z^g}{2} < z < z_g + \frac{\Delta z^g}{2} \\ 0 & \text{Otherwise,} \end{cases} \tag{6.8}$$

and

$$C_{\ell,\mathrm{clus}}^g = \frac{1}{\Delta z^g}\frac{H(z^g)}{c\chi^2(z)} b_g^2(z) P_{\mathrm{m}}(k = \frac{\ell+\frac{1}{2}}{\chi(z^g)}, z^g). \tag{6.9}$$

The galaxy (shot) Poisson noise power is the reciprocal of the surface number density $dN_g/d\Omega$,

$$C_{\ell,\mathrm{shot}}^g = \left(\frac{dN_g}{d\Omega}\right)^{-1} = \left(\Delta z^g \frac{dN_g}{dz d\Omega}\right)^{-1}, \tag{6.10}$$

where the second equality uses the same assumption of a narrow redshift bin.

Intensity Field Auto Power Spectrum C_ℓ^I

The angular auto power spectrum of an intensity field includes contributions from clustering, Poisson noise, and instrument noise,

$$C_\ell^I = C_{\ell,\text{clus}}^I + C_{\ell,\text{shot}}^I + C_{\ell,n}^I. \tag{6.11}$$

Using the Limber approximation, the clustering power spectrum of the intensity map at frequency ν is

$$\begin{aligned} C_{\ell,\text{clus}}^I = &\int dz \frac{H(z)}{c} \frac{f^I(z) f^I(z)}{\chi^2(z)} b_I^2(z,\nu) \\ &\cdot \left[\frac{d(\nu I_\nu)(z, R(\nu))}{dz}\right]^2 P_\text{m}\left(k = \frac{\ell+\frac{1}{2}}{\chi(z)}, z\right), \end{aligned} \tag{6.12}$$

where b_I is the large-scale bias, and f^I the redshift selection function. We take $f^I(z) = 1$ for simplicity as the intensity field is comprised from emission over a large redshift range. $\nu I_\nu(z, R(\nu))$ is the intensity of emitting sources at redshift z observed by a filter with frequency response function $R(\nu)$, i.e.,

$$\frac{d(\nu I_\nu)}{dz}(z, R(\nu)) = \frac{1}{\int d\nu R(\nu)} \int d\nu R(\nu) \frac{d(\nu I_\nu)}{dz}(z,\nu), \tag{6.13}$$

where $\nu I_\nu(z,\nu)$ is the intrinsic spectral energy distribution of the EBL from sources at redshift z observed at frequency ν. In this work, we consider filters with a narrow spectral width ($\Delta\nu/\nu \ll 1$), and use the approximation:

$$\frac{d(\nu I_\nu)}{dz}(z, R(\nu)) = \frac{d(\nu I_\nu)}{dz}(z,\nu). \tag{6.14}$$

The EBL is the aggregate intensity from all sources across cosmic time, defined as

$$\nu I_\nu(\nu) = \int dz \frac{d(\nu I_\nu)}{dz}(z,\nu). \tag{6.15}$$

In the optical to near-infrared wavelengths, the emitting sources contributing to the measured intensity include galaxies, quasars, stars, from all luminosity ranges and from all redshifts [e.g., 22, 70].

We model the integrated galaxy light (IGL) as the main emitting source of the EBL, which can be considered as a lower bound to $\nu I_\nu(z,\nu)$. We model the continuum and spectral line emission of the IGL spectrum separately,

$$\begin{aligned} \left.\frac{d(\nu I_\nu)}{dz}\right|_\text{IGL}(z,\nu) = &\left.\frac{d(\nu I_\nu)}{dz}\right|_\text{cont}(z,\nu) \\ &+ \sum_\text{line} \left.\frac{d(\nu I_\nu)}{dz}\right|_\text{line}(z,\nu). \end{aligned} \tag{6.16}$$

The continuum can be expressed in terms of a volume averaged galaxy emissivity, consisting of the galaxy luminosity function $\Phi(m, z, \nu_{\rm rf}) = dN/dV/dm$, where N is the galaxy number count, V is the co-moving volume, m is the AB magnitude, and $\nu_{\rm rf}$ is the rest-frame frequency. The spectrum of the continuum intensity can be written as

$$\begin{aligned} \left.\frac{d(\nu I_\nu)}{dz}\right|_{\rm cont}(z,\nu) = \int_{m_{\rm th}(z,\nu)}^{\infty} & dm \Phi(m, z, \nu(1+z)) \\ & \cdot \nu F_\nu(m) \frac{d\chi}{dz}(z) D_A^2(z), \end{aligned} \tag{6.17}$$

where $F_\nu(m) = 3631 \cdot 10^{m/2.5}$ Jy is the specific flux density at magnitude m, $d\chi/dz = c/H(z)$, and D_A is the co-moving angular diameter distance, which equals the co-moving distance in a flat universe. The lower bound of the magnitude integration, $m_{\rm th}(z, \nu)$, is the masking magnitude threshold, below which sources are masked in the diffuse intensity map. Masking reduces the foreground contributions to the cross power spectrum Poisson noise, as discussed in Sec. 6.8. The models for Φ, b_I, and the choice of $m_{\rm th}$ used in this work are detailed in Sec. 6.3.

For spectral lines, we build the emission model based on the halo model formalism and the mass M and line luminosity $L_{\rm line}$ relation:

$$\begin{aligned} \left.\frac{d(\nu I_\nu)}{dz}\right|_{\rm line}(z,\nu) = \int_{M_{\rm min}}^{M_{\rm th}(z)} & dM \frac{dn}{dM}(M,z) \frac{\nu L_{\rm line}(M,z)}{4\pi D_L^2(z)} \\ & \cdot \frac{d\chi}{dz}(z) D_A^2(z) \delta^D(\nu - \nu_{\rm rf}^{\rm line}/(1+z)), \end{aligned} \tag{6.18}$$

where dn/dM is the halo mass function [95], D_L is the luminosity distance, δ^D is the Dirac delta function, and $\nu_{\rm rf}^{\rm line}$ is the rest frame frequency of the spectral line. Here we treat the intrinsic spectral line width as a delta function, as the intrinsic line width in a galaxy is unresolved in SPHEREx's low-resolution spectral bands. The lower limit of the integration, $M_{\rm min}$, is the minimum halo mass hosting sources with line emission, and we set $M_{\rm min} = 10^8 M_\odot$/h. The upper limit, $M_{\rm th}(z)$, is determined by the line masking threshold and detailed in Sec. 6.8.

The line intensity measured at frequency band ν with width $\Delta\nu$ is

$$\left.\frac{d(\nu I_\nu)}{dz}\right|_{\rm line}(z,\nu,\Delta\nu) = \frac{1}{\Delta z^{\rm line}} \left.\nu I_\nu\right|_{\rm line}(z^{\rm line}), \tag{6.19}$$

where

$$\begin{aligned} \nu I_\nu|_{\rm line}(z) = &\int_{M_{\rm min}}^{M_{\rm th}(z)} dM \frac{dn}{dM}(M,z) \\ &\cdot \frac{\nu L_{\rm line}(M,z)}{4\pi D_L^2(z)} \frac{d\chi}{d\nu}(z) D_A^2(z), \end{aligned} \tag{6.20}$$

$z^{\rm line} = \nu_{\rm rf}^{\rm line}/\nu - 1$, $\Delta z^{\rm line} = \Delta\nu \left(\nu_{\rm rf}^{\rm line}/\nu^2\right)$, and $d\chi/d\nu = c(1+z)^2/(\nu_{\rm rf}^{\rm line} H(z))$. Thus,

$$\begin{aligned} \left.\frac{d(\nu I_\nu)}{dz}\right|_{\rm line}(z,\nu,\Delta\nu) = &\frac{c}{4\pi(1+z)^2 H(z)} \frac{\nu}{\Delta\nu} \\ &\cdot \int_{M_{\rm min}}^{M_{\rm th}(z)} dM \frac{dn}{dM}(M,z) L_{\rm line}(M,z). \end{aligned} \tag{6.21}$$

Note the line intensity is inversely proportional to the observed spectral width $\Delta\nu$.

The intensity Poisson noise comprises galaxy continuum and line emission, and foreground emission. Here we consider Galactic stars as foreground sources that contribute to the fluctuation power spectrum, and thus

$$C_{\ell,{\rm shot}}^I = \left.C_{\ell,{\rm shot}}^I\right|_{\rm cont} + \sum_{\rm line} \left.C_{\ell,{\rm shot}}^I\right|_{\rm line} + \left.C_{\ell,{\rm shot}}^I\right|_{\rm star}. \tag{6.22}$$

In the limit of narrow spectral channel width, the Poisson-noise components are given by

$$\begin{aligned} \left.C_{\ell,{\rm shot}}^I\right|_{\rm cont} = &\int dz \int_{m_{\rm th}(z,\nu)}^{\infty} dm \Phi(m,z,\nu(1+z)) \\ &\cdot [\nu F_\nu(m)]^2 \frac{d\chi}{dz}(z) D_A^2(z), \end{aligned} \tag{6.23}$$

$$\begin{aligned} \left.C_{\ell,{\rm shot}}^I\right|_{\rm line} = &\sum_{\rm line} \frac{1}{\Delta z^{\rm line}} \frac{H(z^{\rm line})}{c\chi^2(z^{\rm line})} \int_{M_{\rm min}}^{M_{\rm th}(z)} dM \\ &\cdot \frac{dn}{dM}(M,z^{\rm line}) \left[\frac{\nu L_{\rm line}(M,z^{\rm line})}{4\pi D_L^2(z^{\rm line})} \frac{d\chi}{d\nu}(z^{\rm line}) D_A^2(z^{\rm line})\right]^2, \end{aligned} \tag{6.24}$$

and

$$\left.C_{\ell,{\rm shot}}^I\right|_{\rm star} = \int_{m_{\rm th}(z,\nu)}^{\infty} dm \frac{dN_{\rm star}}{dm d\Omega}(m,\nu) \left[\nu F_\nu(m)\right]^2, \tag{6.25}$$

where $dN_{\rm star}/dm/d\Omega$ is the star count per magnitude per solid angle.

Here we ignore the correlation between continuum and and line emissions, and model them as independent source populations that contribute to the Poisson noise.

This correlation is negligible in narrow band measurements ($\Delta\nu/\nu \ll 1$) considered in this work, as the line emission in a given frequency channel only comes from a narrow redshift range ($\Delta z^{\rm line} \ll 1$).

We assume instrument noise to be Gaussian with a variance $\sigma_n^2(\nu)$, so the noise power spectrum is given by

$$C^{\rm I}_{\ell,n} = \sigma_n^2(\nu)\Omega_{\rm pix}, \tag{6.26}$$

where $\Omega_{\rm pix}$ is the pixel size. Note that $\sigma_n^2(\nu)$ has the same units as νI_ν.

Galaxy-Intensity Cross Power Spectrum C_ℓ^x

The cross power spectrum also contains clustering and Poisson noise components:

$$C^{\rm x}_\ell = C^{\rm x}_{\ell,\rm clus} + C^{\rm x}_{\ell,\rm shot}. \tag{6.27}$$

The clustering term in the Limber approximation has the expression:

$$\begin{aligned} C^{\rm x}_{\ell,\rm clus} = \int dz \frac{H(z)}{c}\frac{f^g(z)f^I(z)}{\chi^2(z)} r_{\rm x}(z) b_g(z) b_I(z) \\ \cdot \frac{d(\nu I_\nu)}{dz}(z,\nu) P_{\rm m}(k = \frac{\ell+\frac{1}{2}}{\chi(z)}, z), \end{aligned} \tag{6.28}$$

where $r_{\rm x}$ is the cross correlation coefficient between galaxy sample and intensity field. We set $r_{\rm x} = 1$ throughout this work as we consider linear scales only.

We separate the clustering term into continuum and line emission,

$$C^{\rm x}_{\ell,\rm clus} = C^{\rm x}_{\ell,\rm clus}\Big|_{\rm cont} + \sum_{\rm line} C^{\rm x}_{\ell,\rm clus}\Big|_{\rm line}. \tag{6.29}$$

In the limit of a narrow observed spectral channel ($\Delta\nu/\nu \ll 1$) and redshift width ($\Delta z^g \ll 1$), we get

$$\begin{aligned} C^{\rm x}_{\ell,\rm clus}\Big|_{\rm cont} = \frac{H(z^g)}{c\chi^2(z^g)} b_g(z^g) b_{I,\rm cont}(z^g) \\ \cdot \frac{d(\nu I_\nu)}{dz}\Big|_{\rm cont}(z^g,\nu) P_{\rm m}\left(k = \frac{\ell+\frac{1}{2}}{\chi(z^g)}, z^g\right), \end{aligned} \tag{6.30}$$

$$\begin{aligned} C^{\rm x}_{\ell,\rm clus}\Big|_{\rm line} = \frac{H(z^g)}{c\chi^2(z^g)} \frac{\Delta z^{\rm g,line}}{\Delta z^g \Delta z^{\rm line}} b_g(z^g) b_{I,\rm line}(z^g) \\ \cdot \nu I_\nu|_{\rm line}(z^g) P_{\rm m}\left(k = \frac{\ell+\frac{1}{2}}{\chi(z^g)}, z^g\right), \end{aligned} \tag{6.31}$$

where $\Delta z^{\mathrm{g,line}}$ is the redshift range where the galaxy tracer redshift z^g overlaps with the line redshift z^{line}.

The Poisson noise cross power spectrum is proportional to the total intensity of the tracer galaxies, which is equivalent to stacking on tracer galaxies in the real space. The Poisson noise cross power spectrum can similarly be written as the sum of continuum and line contributions:

$$\begin{aligned} C^{\mathrm{x}}_{\ell,\mathrm{shot}} &= \left(\frac{dN_g}{d\Omega}\right)^{-1} \Delta z^g \left.\frac{d(\nu I_\nu)}{dz}\right|_g (z^g,\nu) \\ &= \left.C^{\mathrm{x}}_{\ell,\mathrm{shot}}\right|_{\mathrm{cont}} + \left.C^{\mathrm{x}}_{\ell,\mathrm{shot}}\right|_{\mathrm{line}}. \end{aligned} \tag{6.32}$$

We use "g" to represent the intensity of the tracer (galaxy) sample. The continuum and line Poisson noise components are given by

$$\left.C^{\mathrm{x}}_{\ell,\mathrm{shot}}\right|_{\mathrm{cont}} = \left(\frac{dN_g}{d\Omega}\right)^{-1} \Delta z^g \left.\frac{d(\nu I_\nu)}{dz}\right|_{\mathrm{cont},g} (z^g,\nu), \tag{6.33}$$

and

$$\left.C^{\mathrm{x}}_{\ell,\mathrm{shot}}\right|_{\mathrm{line}} = \left(\frac{dN_g}{d\Omega}\right)^{-1} \frac{\Delta z^{g,\mathrm{line}}}{\Delta z^{\mathrm{line}}} \left.\nu I_\nu\right|_{\mathrm{line},g} (z^g,\nu), \tag{6.34}$$

where the last term in both equations is the integration of Eq. 6.17 and 6.20 over the galaxy sample, respectively. Here we assume the galaxy tracers are the brightest continuum and line emitting sources, and thus there exist a limiting magnitude $m_g(z,\nu)$ (halo mass $M_g(z)$) for continuum (lines) where the tracers are all the galaxies below (above) that limit. For a given masking threshold $m_{\mathrm{th}}(z,\nu)$ ($M_{\mathrm{th}}(z)$), if the tracer catalog is deeper than the mask, i.e., $m_g > m_{\mathrm{th}}$ ($M_g < M_{\mathrm{th}}$), we get

$$\begin{aligned} \left.\frac{d(\nu I_\nu)}{dz}\right|_{\mathrm{cont},g} (z,\nu) = \int_{m_{\mathrm{th}}(z,\nu)}^{m_g(z,\nu)} & dm\Phi(m,z,\nu(1+z)) \\ & \cdot \nu F_\nu(m)\frac{d\chi}{dz}(z)D_A^2(z), \end{aligned} \tag{6.35}$$

and

$$\begin{aligned} \left.\nu I_\nu\right|_{\mathrm{line},g} (z,\nu) = \int_{M_g(z)}^{M_{\mathrm{th}}(z)} & dM\frac{dn}{dM}(M,z) \\ & \cdot \frac{\nu L_{\mathrm{line}}(M,z)}{4\pi D_L^2(z)}\frac{d\chi}{d\nu}(z)D_A^2(z). \end{aligned} \tag{6.36}$$

On the other hand, if $m_g \geqslant m_{\mathrm{th}}$ ($M_g \leqslant M_{\mathrm{th}}$), all the tracer galaxies are masked, and thus

$$\left.\frac{d(\nu I_\nu)}{dz}\right|_{\mathrm{cont},g} (z,\nu) = \left.\nu I_\nu\right|_{\mathrm{line},g} (z,\nu) = 0, \tag{6.37}$$

and the cross Poisson noise vanishes.

Galaxy-Intensity Cross Trispectrum $\mathcal{T}^{\rm xx}_{\ell,-\ell,\ell',-\ell'}$

Since the non-Gaussian effects are negligible on the large scales, we only model the Poisson term of the trispectrum, which is given by [89]

$$\mathcal{T}^{\rm xx}_{\ell,{\rm shot}} = \left[C^I_{\ell,{\rm shot}}\Big|_{{\rm cont},g} + \sum_{\rm line} C^I_{\ell,{\rm shot}}\Big|_{{\rm line},g} \right] \cdot \left(C^g_{\ell,{\rm shot}} \right)^2, \tag{6.38}$$

where $C^g_{\ell,{\rm shot}}$ is given by Eq. 6.10, and $C^I_{\ell,{\rm shot}}\Big|_{\rm cont,g}$ and $\sum_{\rm line} C^I_{\ell,{\rm shot}}\Big|_{\rm line,g}$ is the integration of Eq. 6.22 and Eq. 6.23 over galaxy sample, respectively. If $m_g > m_{\rm th}$ ($M_g < M_{\rm th}$),

$$\begin{aligned} C^I_{\ell,{\rm shot}}\Big|_{{\rm cont},g} = \int dz \int_{m_{\rm th}(z,\nu)}^{m_g(z,\nu)} & dm \Phi(m, z, \nu(1+z)) \\ & \cdot \left[\nu F_\nu(m)\right]^2 \frac{d\chi}{dz}(z) D_A^2(z) \end{aligned} \tag{6.39}$$

and

$$\begin{aligned} C^I_{\ell,{\rm shot}}\Big|_{{\rm line},g} = & \sum_{\rm line} \frac{1}{\Delta z^{\rm line}} \frac{H(z^{\rm line})}{c\chi^2(z^{\rm line})} \\ & \cdot \int_{M_g(z)}^{M_{\rm th}(z)} dM \cdot \frac{dn}{dM}(M, z^{\rm line}) \\ & \cdot \left[\frac{\nu L_{\rm line}(M, z^{\rm line})}{4\pi D_L^2(z^{\rm line})} \frac{d\chi}{d\nu}(z^{\rm line}) D_A^2(z^{\rm line}) \right]^2. \end{aligned} \tag{6.40}$$

Otherwise, if $m_g \geqslant m_{\rm th}$ ($M_g \leqslant M_{\rm th}$), we get

$$C^I_{\ell,{\rm shot}}\Big|_{{\rm cont},g} = C^I_{\ell,{\rm shot}}\Big|_{{\rm line},g} = \mathcal{T}^{\rm xx}_{\ell,{\rm shot}} = 0. \tag{6.41}$$

6.3 Emission Model

We assume the integrated galaxy light (IGL) constitutes the majority of the near-IR emission, and decompose it into continuum and spectral line components. In addition, we also consider another model for quasars, which is used for studying the constraints on quasar spectrum with cross correlation. Note that we do not add the quasar model to the IGL model, since IGL already includes the quasar populations. We also consider Galactic stars as foreground emissions.

Galaxy Spectral Energy Distribution

Continuum Emission

Our model for the IGL continuum emission is based on a luminosity function prescription from [44], where the luminosity function, $\Phi(m, z, \nu_{\rm rf})$, has a Schechter functional form [90] that depends on magnitude m, redshift z, and rest-frame frequency $\nu_{\rm rf}$. The parameters are calibrated to observations in several ultraviolet to mid-infrared wavelength bands, and we interpolate the Schechter parameters to all wavelengths from their model values.

Our model for the continuum intensity bias, $b_{I,\rm cont}(z, \nu)$, is also from [44], who uses a halo occupation distribution (HOD) framework [111] and assumes no frequency dependence:

$$b_{I,\rm cont}(z) = \frac{\int dM \frac{dn}{dM}(M,z) b(M,z) \left\langle N_{\rm gal} \right\rangle}{\int dM \frac{dn}{dM}(M,z) \left\langle N_{\rm gal} \right\rangle}, \tag{6.42}$$

where $\left\langle N_{\rm gal} \right\rangle$ is the total halo occupation number. Following [44], we use the HOD parameters given by the SDSS measurement in [107].

Line Emission

In this work, we consider five prominent optical spectral lines: Hα (656.3 nm), [O III] (500.7 nm), Hβ (486.1 nm), [O II] (372.7 nm), and Lyα (121.6 nm). These lines are of particular interest as they are visible in SPHEREx bands across a range of redshifts.

Our line emission model is built on an empirical halo mass and line luminosity relation $L_{\rm line}(M, z)$. We adopt and summarize the prescription from [41] below. First, we assume a simple linear relation between star formation rate (SFR) and halo mass, and calibrate the scaling factor using the star formation rate density constraints from [46]. We then further assume a linear relation between the line luminosity and SFR, using SFR–$L_{\rm line}$ relations from Kennicutt [50] and Ly et al. [61] for Hα, [O II] and [O III] emission:

$$\frac{\rm SFR}{M_\odot/{\rm yr}} = 7.9 \times 10^{-42} \frac{L_{H\alpha}}{\rm erg/s}, \tag{6.43}$$

$$\frac{\rm SFR}{M_\odot/{\rm yr}} = 1.4 \times 10^{-41} \frac{L_{[O\,II]}}{\rm erg/s}, \tag{6.44}$$

$$\frac{\rm SFR}{M_\odot/{\rm yr}} = 7.6 \times 10^{-42} \frac{L_{[O\,III]}}{\rm erg/s}. \tag{6.45}$$

For Hβ, we assume a fixed line ratio of $H\beta/H\alpha = 0.35$ [79]. As Lyα emission is not discussed in [41], we use a model with a constant halo mass to luminosity M–$L_{\rm Ly\alpha}$ ratio at halo masses of $M = 10^8$–10^{15} Mpc/h, and calibrate the scaling factor to match the mean Lyα intensity of the analytical model from Pullen et al. [82, Eq. 51].

The line emission bias factor, $b_{I,\rm line}$, is given by the the halo bias, $b(M,z)$, weighted by line luminosity:

$$b_{I,\rm line}(z) = \frac{\int_{M_{\rm min}}^{M_{\rm th}(z)} dM \frac{dn}{dM}(M,z) b(M,z) L_{\rm line}(M,z)}{\int_{M_{\rm min}}^{M_{\rm th}(z)} dM \frac{dn}{dM}(M,z) L_{\rm line}(M,z)}. \tag{6.46}$$

We use the same lower and upper integration limits ($M_{\rm min}$ and $M_{\rm th}(z)$) as Eq. 6.18.

Quasar Spectrum

Our quasar emission model assumes a quasar luminosity function and an averaged quasar spectrum. We adopt the fiducial model of the quasar luminosity function in [21], and use the quasar spectrum template compiled in [64] that splices together a composite spectra from [102], [85], and [60].

Foreground Stars

Galactic stars contribute to the Poisson noise (Eq. 6.25) in the auto spectrum of the intensity field. We model star counts, $dN_{\rm star}/dm/d\Omega$ in Eq. 6.25, using the stellar population synthesis code, *Trilegal*[1] [38, 39]. We use their model in the *UBVRIJHKLMN* photometric system at the north galactic pole ($\ell = 0°$, $b = 90°$).

6.4 Cross-Correlation Forecast

Evaluation Metrics

We focus on constraining the redshift-dependent IGL spectrum, $d(\nu I_\nu)/dz(\nu, z)$, from cross-correlating galaxies and the intensity field. The clustering and Poisson-noise amplitudes in the cross power spectrum are proportional to $b_I(z) d(\nu I_\nu)/dz(z,\nu)$ and $d(\nu I_\nu)/dz|_{\rm g}\,(z,\nu)$, respectively (Eq. 6.28, 6.32). Therefore we forecast the signal-to-noise ratio (SNR) on these two quantities obtained from the cross power spectrum. Note $d(\nu I_\nu)/dz|_{\rm g}\,(z,\nu)$ represents the IGL spectrum from the specific galaxies (tracers) used in the cross-correlation measurements, and does not equal the cosmic mean IGL spectrum, $d(\nu I_\nu)/dz(z,\nu)$.

As the matter power spectrum is linear on large scales, here defined as $k \lesssim 0.2$ h/Mpc, we evaluate constraints on $b_I(z) d(\nu I_\nu)/dz(z,\nu)$ from the amplitude of all $\ell_{\rm clus}^{\rm min} < \ell < \ell_{\rm clus}^{\rm max}$ modes, where $\ell_{\rm clus}^{\rm max} = 0.2[{\rm h/Mpc}]\chi(z) - 1/2$, and the minimum accessible mode $\ell_{\rm clus}^{\rm min} = 51$, which corresponds to SPHEREx's field of view, as the modes larger than the field size will be susceptible to the foreground filtering

---

[1] `http://stev.oapd.inaf.it/cgi-bin/trilegal`

process. Therefore, the SNR on the clustering amplitude is

$$\mathrm{SNR}_{\mathrm{clus}}(b_I(z)\frac{d(\nu I_\nu)}{dz}(z,\nu)) = \sqrt{\sum_{\ell\in[\ell^{\mathrm{min}}_{\mathrm{clus}},\ell^{\mathrm{max}}_{\mathrm{clus}}]}\left[\frac{C^{\mathrm{x}}_{\ell,\mathrm{clus}}}{\delta C^{\mathrm{x}}_{\ell}}\right]^2}. \tag{6.47}$$

In the Poisson-noise regime, defined as $k \gtrsim 1$ h/Mpc, we use the modes in $\ell^{\mathrm{min}}_{\mathrm{shot}} < \ell < \ell^{\mathrm{max}}_{\mathrm{shot}}$ to constrain $\frac{d(\nu I_\nu)}{dz}\Big|_{\mathrm{g}}(z,\nu)$, where $\ell^{\mathrm{shot}}_{\mathrm{min}} = 1[\mathrm{h/Mpc}]\chi(z) - 1/2$, and $\ell^{\mathrm{max}}_{\mathrm{shot}} = 10^5$, which is the highest multipole mode available given SPHEREx pixel size (6.2″). Incorporating the effects from both Gaussian and non-Guassian terms, the SNR on the Poisson noise amplitude is given by

$$\begin{aligned}&\mathrm{SNR}_{\mathrm{shot}}(\left.\frac{d(\nu I_\nu)}{dz}\right|_g(z,\nu))\\ &= \frac{C^{\mathrm{x}}_{\ell,\mathrm{shot}}}{\sqrt{\frac{1}{N_{\ell>\ell_{\mathrm{min}}}}\sum_{\ell\in[\ell^{\mathrm{min}}_{\mathrm{shot}},\ell^{\mathrm{max}}_{\mathrm{shot}}]}\left(\delta C^{\mathrm{x}}_{\ell}\right)^2 + \frac{\mathcal{T}^{\mathrm{xx}}_{\ell,-\ell,\ell',-\ell'}}{\Omega_{\mathrm{sur}}}}},\end{aligned} \tag{6.48}$$

where Ω_{sur} is the survey size used for cross correlation.

Since $b_I(z)d(\nu I_\nu)/dz(z,\nu)$ and $d(\nu I_\nu)/dz|_g\,(z,\nu)$ are the main quantities of interest, we forecast below their constraints as a function of redshift and wavelength from upcoming surveys.

Masking

In practice, pixels that contain bright stars and bright galaxies will have to be masked in data processing to reduce sample variance noise in the cross power spectrum, and the resulting power spectrum amplitude will depend on the depth of the masking limit. To model the effect of masking on the power spectrum, we assume a photometric source catalog with sufficient depth provides the location and fluxes of bright stars and galaxies to be masked. The masking magnitude threshold, $m^{\mathrm{ph}}_{\mathrm{th}}$, is set by the limiting magnitude of a photometric band at frequency ν^{ph} (or the corresponding wavelength λ^{ph}).

For simplicity, we mask all stars and galaxies brighter than our masking threshold (see Sec. 6.8 for discussion). Since only the flux density information is needed, we can use a photometric source catalog for this purpose, which is usually deeper than a spectroscopic catalog. Given $m^{\mathrm{ph}}_{\mathrm{th}}$ and ν^{ph}, we adopt the abundance matching concept [10, 20, 55, 76, 101] to infer the masking magnitude threshold $m_{\mathrm{th}}(\nu, z)$ for continuum emission at a given frequency ν; this assumes that the brightest N sources in a given frequency bin are also the brightest ones in another frequency bin. We therefore first calculate the number density of galaxies being masked at $\nu^{\rm ph}$:

$$N = \int_{-\infty}^{m_{\rm th}^{\rm ph}} dm \Phi(m, z, \nu^{\rm ph}(1+z)), \tag{6.49}$$

and then at each frequency solve for the $m_{\rm th}(\nu, z)$ value which gives

$$N = \int_{-\infty}^{m_{\rm th}(\nu,z)} dm \Phi(m, z, \nu(1+z)). \tag{6.50}$$

Analogously, for line emission modeled with a line luminosity and halo mass relation, given the halo mass function we can derive its maximum halo mass $M_{\rm th}(z)$ by solving for

$$N = \int_{M_{\rm th}(z)}^{\infty} dM \frac{dn}{dM}(M, z). \tag{6.51}$$

Throughout this work, we set the fiducial masking threshold to $m_{\rm th}^{\rm ph} = 20$ at frequency $\nu^{\rm ph} = 1\mu$m. Our model suggests that $\sim 3\%$ of the SPHEREx pixel (6.2″) contain stars or galaxies above this limit, and the intensity Poisson noise from unmasked stars is approximately an order of magnitude below the galaxy Poisson noise. Here we ignore the sensitivity loss from mode mixing due to masking (see Sec. 6.8 for more discussions).

Fiducial Case

We consider cross power spectra between SPHEREx intensity map and several current and future galaxy redshift surveys. For simplicity, we present results with a few chosen parameters as our fiducial case.

Fig. 6.1 shows the redshift-dependent IGL spectra, $d(\nu I_\nu)/dz$, from our model (Sec. 6.3) at fiducial redshifts. The spectra include continuum and spectral line emission, and are presented with SPHEREx spectral binning (Sec. 6.5) of 96 channels over $0.75-5\mu$m[2]. The IGL spectral amplitude builds up with cosmic time. Since SPHEREx has higher spectral resolution at $\lambda > 3.8\ \mu$m, the line intensity at longer wavelength is stronger compared to the continuum as the line signal is inversely proportional to the spectral width (Eq. 6.19). The model we used for IGL continuum [44] is only calibrated to observations at $z \lesssim 5$ and rest frame

---

[2] SPHEREx recently modified its design to have 102 frequency channels covering the same wavelength range. As this has negligible impact on our results, we use their 96 channel configuration with published sensitivity forecasts for this work.

wavelengths $\lambda > 0.15$ μm. The higher redshifts and shorter wavelengths model are derived from extrapolation.

At low redshifts ($z < 3$), cross correlations between SPHEREx and spectroscopic redshift surveys with high redshift accuracy ($\sigma_z/(1+z) \sim 10^{-4} - 10^{-3}$) allow selecting galaxies in thin redshift slices. We choose a redshift bin width of $\Delta z^g/(1+z^g) = 0.03$, and calculate the cross-correlation at five fiducial redshifts, $z^g = [0.25, 0.5, 1, 2, 3]$. Between a redshift of 0 and 3 there are $\sim$ 46 such measurements. We choose a fiducial masking threshold $m^{\rm ph}_{\rm th} = 20$ at $\lambda^{\rm ph} = 1$ μm for all these cases, except for at $z = 0.25$ where we use $m^{\rm ph}_{\rm th} = 18$. This threshold is chosen to optimize the trade off between losing signal and reducing foreground emission (see Sec. 6.8 for more discussion). SPHEREx will achieve point source sensitivity $m_{\rm AB} \sim 20$ at $\lambda/\Delta\lambda \sim 4$, and therefore our masking depths is feasible with the SPHEREx source catalog.

At higher redshifts ($z > 3$), spectroscopic catalogs are either unavailable or lacking sufficient number density and/or sky coverage, and therefore we consider cross-correlating SPHEREx with photometric redshift catalogs, which have a redshift accuracy of $\sigma_z/(1+z) \sim 0.02 - 0.05$. We choose eight fiducial redshifts: $z^g = [3, 4, 5, 6, 7, 8, 9, 10]$. For $3 \leqslant z \leqslant 6$, we use a redshift bin width of $\Delta z^g/(1+z^g) = 0.1$, which gives $\sim$ 6 such measurements. At $z > 6$, the redshift bins are given by the resolution from the Lyman break technique (see Sec. 6.6 for details). We also set the masking threshold to $m^{\rm ph}_{\rm th} = 20$ at $\lambda^{\rm ph} = 1$ μm.

6.5 SPHEREx Intensity Mapping

SPHEREx is a NASA MIDEX mission scheduled to launch in 2024 [33, 34][3]. SPHEREx will carry out the first all-sky near-infrared spectro-imaging survey from 0.75 to 5 μm. It has a pixel size of $6.2''$, and a wavelength-dependent spectral resolution: a spectral resolving power of $R = 41$ at wavelengths between 0.75 and 2.42 μm, (48 spectral channels), $R = 35$ between 2.42 and 3.82 μm (16 spectral channels), $R = 110$ between 3.82 and 4.42 μm (16 spectral channels), and $R = 130$ between 4.42 and 5.00 μm (16 spectral channels), The top panel of Fig. 6.2 shows the SPHEREx spectral resolution at each channel, and the bottom panel is the expected noise rms per spectral channel per pixel of SPHEREx of the all-sky and the two $\sim$ 100 deg^2 deep surveys[4], which we use to calculate the noise contribution to

---

[3] `http://spherex.caltech.edu`

[4] Data downloaded from: `https://github.com/SPHEREx/Public-products/blob/master/Surface_Brightness_v28_base_cbe.txt`

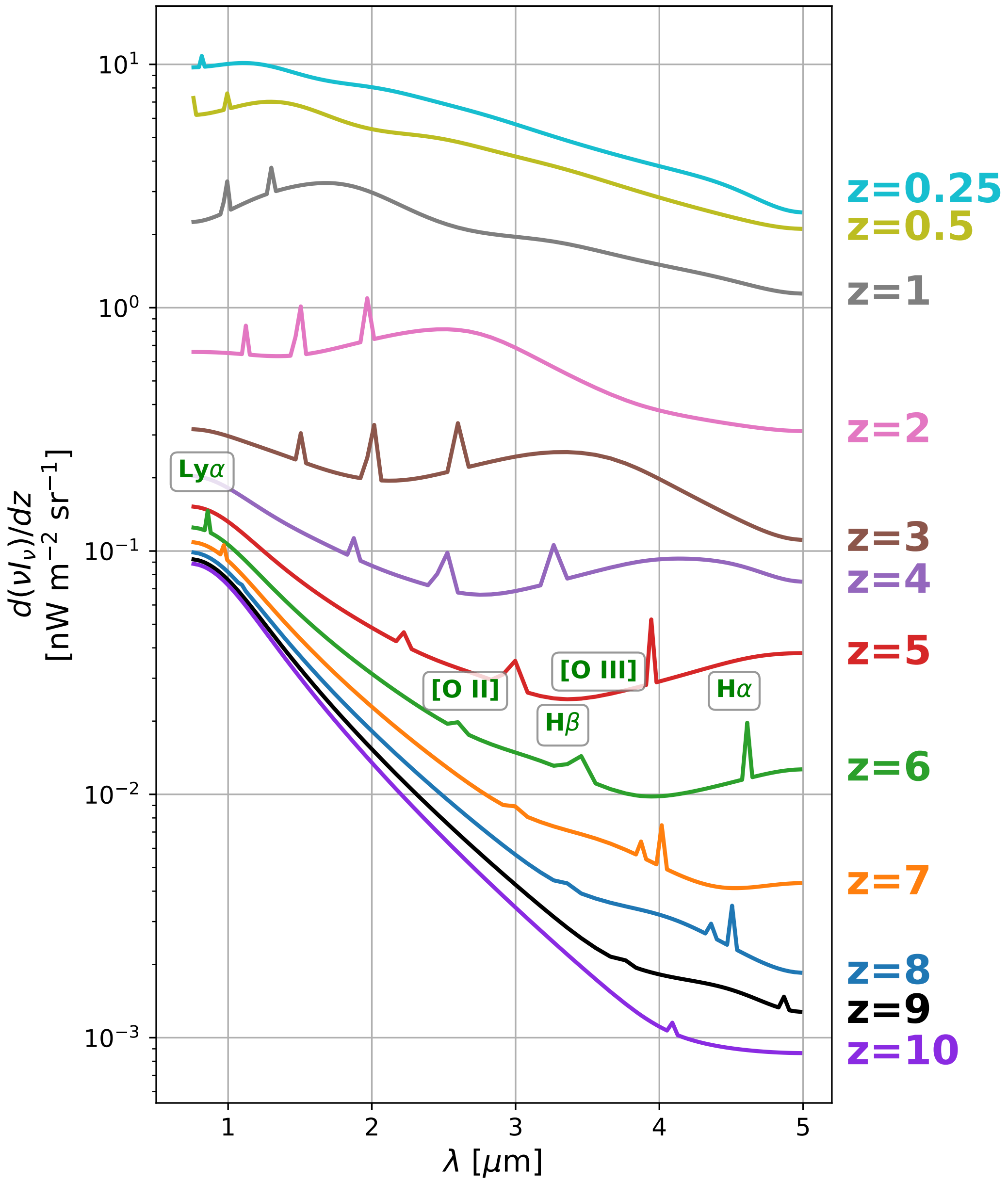

Figure 6.1: Redshift-dependent IGL spectra from our model (Sec. 6.3) at fiducial redshifts, using SPHEREx spectral binning. The spectra contain continuum and line emission. The spectral lines from left to right are Lyα, [O II], Hβ, [O III], and Hα, respectively, as labeled in $z = 6$ case. For some cases, Hβ and [O III] lines are mixed in the same spectral channel since their frequencies are close to each other.

cross-power spectrum (Eq. 6.26). Here we adopt SPHEREx's optimistic sensitivity model, noting that the pessimistic forecast is $\sim$ 0.5 mag lower. For SPHEREx all-sky survey, we only use 75% of the the full sky coverage to avoid the Galactic

Table 6.1: Summary of the sky area and redshift coverage of the surveys used in this work.

survey name	A_{sur} [deg^2]	f_{sky}	redshift range
Intensity Maps			
SPHEREx	30,940	0.75	N/A
SPHEREx deep	2×100	4.85×10^{-3}	N/A
Spectroscopic Surveys			
BOSS CMASS	10,000	0.242	$0 < z < 1$
eBOSS LRG	7,500	0.182	$0 < z < 1.2$
eBOSS ELG	620	0.0150	$0 < z < 1.5$
eBOSS QSO	7,500	0.182	$0 < z < 3.5$
DESI BGS	14,000	0.339	$0 < z < 0.5$
DESI LRG	14,000	0.339	$0.6 < z < 1.2$
DESI ELG	14,000	0.339	$0.6 < z < 1.7$
DESI QSO	14,000	0.339	$0.6 < z < 4.2$
Euclid ELG	15,000	0.364	$0.65 < z < 2.05$
Euclid deep ELG	10	2.42×10^{-4}	$0.65 < z < 2.05$
Roman ELG	2,200	0.053	$0.5 < z < 3$
SPHEREx spec.	30,940	0.75	$0 < z < 4.6$
Photometric Surveys			
Rubin phot.	18,000	0.436	$0 < z < 4$
Rubin phot. (gold sample)	18,000	0.436	$0 < z < 3$
Rubin Lyman break	18,000	0.436	$2.4 < z < 3.7$
	18,000	0.436	$3.7 < z < 4.8$
	18,000	0.436	$4.8 < z < 5.9$
	18,000	0.436	$5.9 < z < 6.7$
Roman phot.	2,200	0.053	$0 < z < 6$
Roman Lyman Break	2,200	0.053	$5.5 < z < 6.5$
	2,200	0.053	$6.5 < z < 7.5$
	2,200	0.053	$7.5 < z < 8.5$
	2,200	0.053	$8.5 < z < 9.5$
	2,200	0.053	$9.5 < z < 10.5$

contamination.

6.6 Redshift Surveys

We consider cross-correlating intensity maps from SPHEREx and several current and future galaxy surveys. Table 6.1 summarizes the survey parameters of all galaxy surveys. The model for source number density as a function of redshift is shown in Fig. 6.3.

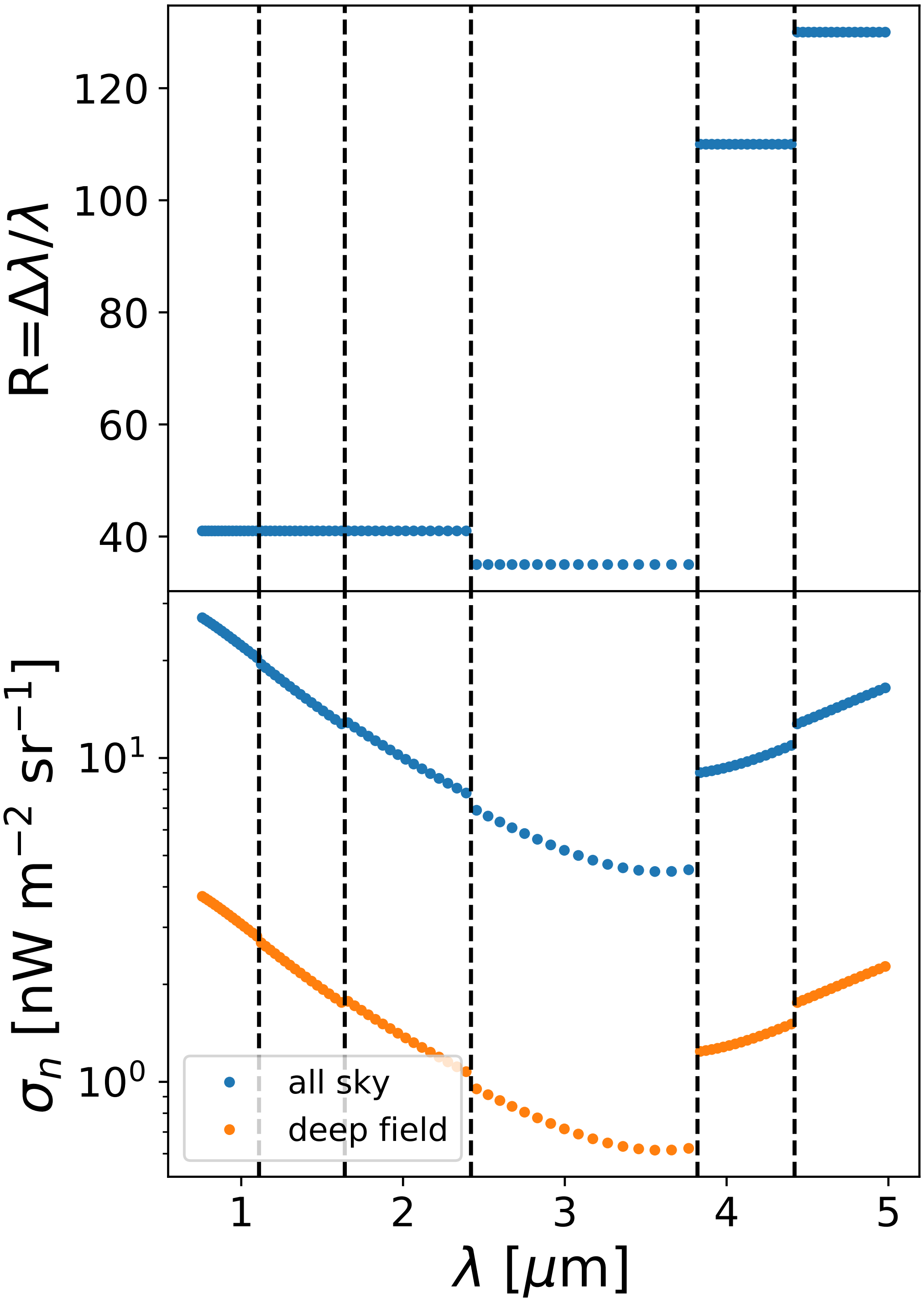

Figure 6.2: Top: SPHEREx spectral resolution in each channel. The vertical dashed lines mark their six frequency bands, each contains 16 channels. Bottom: SPHEREx noise rms per spectral channel on all sky (blue) and the deep fields (orange) in a 6.2″ sky pixel.

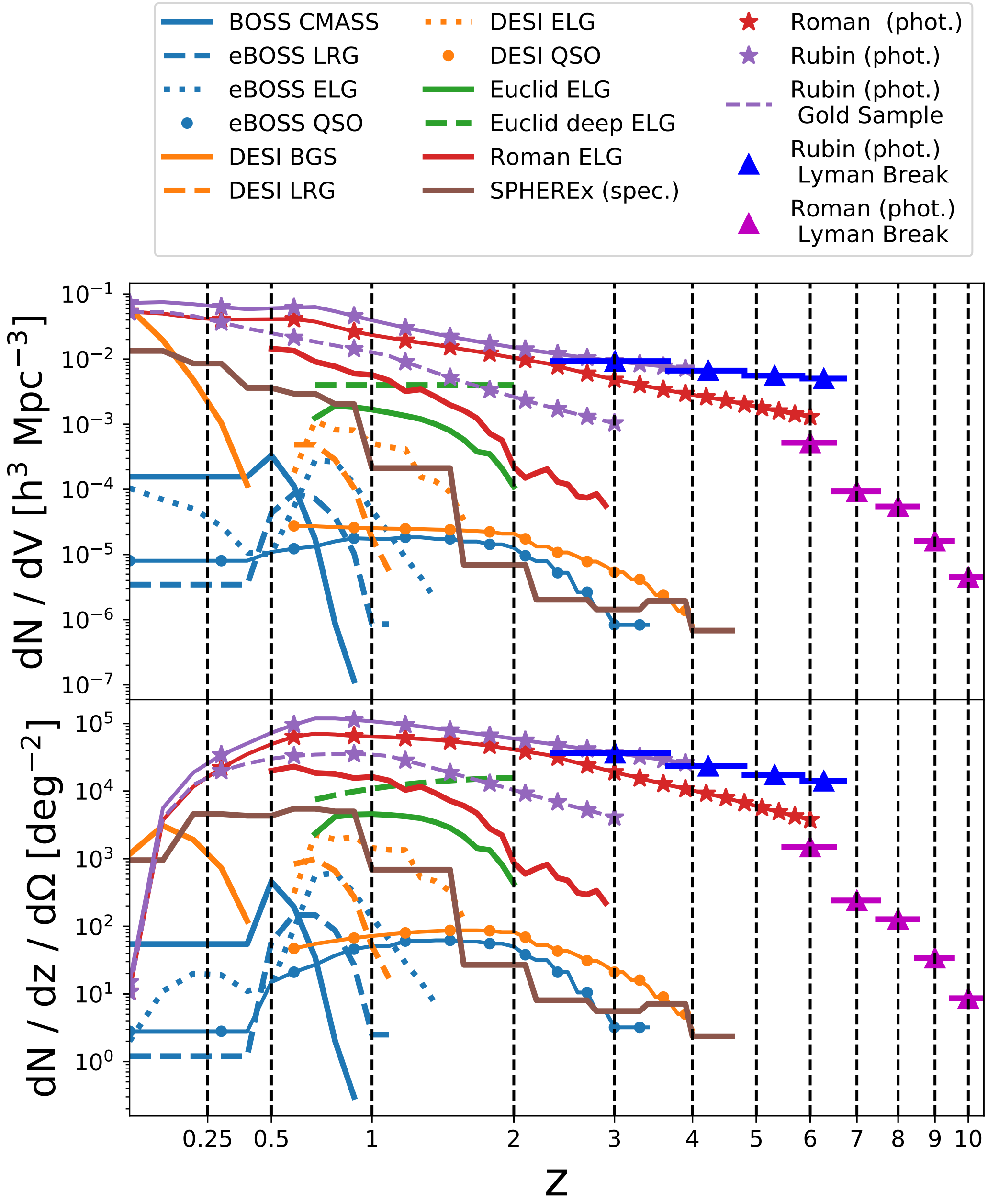

Figure 6.3: Model of number density of the spectroscopic and photometric galaxy catalogs from each survey considered in this work. The values are compiled from the literature or predicted with our model, detailed in Sec. 6.6. We show the number density per co-moving volume on the top panel, and the number density per redshift and solid angle on the bottom panel. The vertical dashed lines mark the fiducial redshifts used in this work.

Spectroscopic Galaxy Surveys

We consider the following current and planned future spectroscopic redshift surveys: SDSS BOSS/eBOSS, DESI, Euclid, the High Latitude Survey with the Roman

Space Telescope. Each survey has one or multiple types of density tracers. Below we describe the parameters we use for each tracer catalog, including the number density, sky coverage, and the galaxy bias factors.

SDSS BOSS/eBOSS

SDSS BOSS [28] and eBOSS [29] are spectroscopic surveys targeting several types of galaxy tracers. In this work, we use four spectroscopic samples from the SDSS: the BOSS CMASS survey, eBOSS LRG (luminous red galaxies), eBOSS ELG (emission line galaxies), and eBOSS QSO (quasars) surveys.

For BOSS CMASS, eBOSS LRG, and eBOSS QSO, we use the predicted number density from Dawson et al. [29] Table 1 ($z_{\rm conf} > 1$ for LRG, and `QSO_CORE` New + Known for QSO); for eBOSS ELG, the number density is given in Table 4 of Raichoor et al. [84]. We assume a tophat functional form of $dN/dV(z)$, the number density per co-moving volume, between the redshift bins quoted in their tables. Our model for galaxy bias, b_g, also follows the prescription of Dawson et al. [29] and Raichoor et al. [84]: $b_g(\text{LRG}) = 1.7/\sigma_8(0)/\sigma_8(z)$, $b_g(\text{ELG}) = 1.0/\sigma_8(0)/\sigma_8(z)$, $b_g(\text{QSO}) = 0.53 + 0.29(1+z)^2$, and we set $b_g(\text{CMASS}) = 2$ according to Nuza et al. [78].

DESI

The Dark Energy Spectroscopic Instrument (DESI) [30] is an ongoing spectroscopic survey that observes four types of tracers across 14,000 deg^2: bright galaxy samples (BGS), luminous red galaxies (LRG), star-forming emission line galaxies (ELG), and quasars (QSO). The predicted number density for each tracer catalog is given in Table 2.3, 2.5, and 2.7 of [30], and we assume a tophat functional form of $dN/dV(z)$ between the redshift bins quoted in their tables. For LRG, ELG, and QSO, we use the same bias model as the eBOSS: $b_g(\text{LRG}) = 1.7/\sigma_8(0)/\sigma_8(z)$, $b_g(\text{ELG}) = 1.0/\sigma_8(0)/\sigma_8(z)$, $b_g(\text{QSO}) = 0.53 + 0.29(1+z)^2$, and we set $b_g(\text{BGS}) = 2$, same as the BOSS CMASS case.

Euclid

The Euclid spectroscopic survey [57] will map ELGs (Hα emitters) over 15,000 deg^2 between redshifts $0.9 \lesssim z \lesssim 2$. The expected number density as a function of redshift is given by Table 3 of Amendola et al. [9] (we use the n_2 reference case), and we assume a tophat $dN/dV(z)$ between the redshift bins quoted in their table. For galaxy bias, we use the model from Merson et al. [73]: $b_g(z) = 0.7z + 0.7$.

In addition, Euclid will also have deep fields that overlap with the SPHEREx deep field at the north ecliptic pole. We therefore consider a case of cross-correlating the Euclid deep field catalog with SPHEREx deep field images. For Euclid deep field sources, we adopt the same bias and redshift coverage as the wide-field ELGs, and assume a constant number density of $n_g = 4 \times 10^{-3}$ (h/Mpc)^3 over this redshift range.

Roman Space Telescope

The Roman Space Telescope [97] plans a spectroscopic survey of ELGs using the Hα and [O III] lines as part of the High Latitude Survey [HLS; 35]. Compared to Euclid, the Roman Space Telescope-HLS covers a smaller sky area (2200 deg^2) but with better point source sensitivity. We adopt the ELG galaxy number density predictions from Zhai et al. [110]. We use the Hα number density (Table 1) for $z < 2$, the [O III] number density (Table 2) for $2 < z < 3$, and assume a tophat $dN/dV(z)$ between the redshift bins quoted in their tables. For both spectral lines we use their “dust model fit at high redshifts” and 1×10^{-16} erg s^{-1} cm^{-2} flux limit case.

SPHEREx Spectroscopic Catalog

In addition to diffuse imaging, SPHEREx will also produce a spectroscopic galaxy catalog [33, 34]. We therefore also consider the case of cross-correlating the SPHEREx galaxy catalog with its intensity maps. For the SPHEREx all-sky survey, we use the predicted number density and galaxy bias of the SPHEREx source catalog at $z < 4.6$ from their public products [5]. Specifically, we use the forecast with galaxy photometric redshift error of $0.01 < \sigma_z/(1+z) < 0.03$, and consider a tophat functional form of $dN/dV(z)$ between the quoted redshift bins. The SPHEREx deep field with higher sensitivity will produce a deeper catalog, which we leave it to future works.

---

[5] `https://github.com/SPHEREx/Public-products/blob/master/galaxy_density_v28_base_cbe.txt`

Photometric Galaxy Surveys

We consider cross-correlating SPHEREx and two upcoming photometric surveys: the Rubin Observatory Legacy Survey of Space and Time (LSST) and the Roman Space Telescope High Latitude Survey (HLS). Despite the larger redshift uncertainties in photometric samples, they have the advantage of reaching higher redshifts and fainter sources.

Rubin Observatory LSST

The Rubin Observatory LSST will carry out an 18,000 deg^2 photometric galaxy survey at $0 < z < 4$ in optical bands [59]. The expected 5-σ point source depth after a ten-year observation is a magnitude of 27.5 at r band[6] . A subset of galaxies from $0 < z < 3$ with higher redshift accuracy ($\sigma_z/(1+z) < 0.05$) constitute the "gold sample," which has a depth of 25.3 in i band [59]. We predict the number density of these two samples by applying these magnitude thresholds to our galaxy luminosity function model from Helgason et al. [44].

Photometric redshift estimates are not available for $z > 4$ sources from the Rubin Observatory LSST; nevertheless, redshift information of high-z sources can be determined by the Lyman-break drop-out technique. We thus estimate Lyman-break galaxy samples in a few broad redshift bands (listed in Table 6.1) defined by the LSST filter boundaries and the redshifted Lyα wavelength (rest-frame 121.6 nm). For these Lyman-break selected galaxies, we estimate their number density by applying the same magnitude threshold (27.5 in r band) to our luminosity function model. We use Eq. 6.42 to estimate galaxy bias in both the photometric and Lyman-break samples.

The Roman Space Telescope

The planned photometric galaxy survey as part of the Roman Space Telescope HLS [35] covers 2200 deg^2 with a 5-σ depth of $\sim$ 26.5 in Y, J, and H band [97]. We apply these magnitude thresholds to our IGL luminosity function model to predict the number density of the HLS photometric galaxies from $z = 0$ to $z = 6$. For $z > 6$, we use the expected Lyman-break selected galaxy number density from the Roman Space Telescope collaboration (S. Finkelstein, private communication)[7].

---

[6] https://www.lsst.org/scientists/keynumbers

[7] See slides from the presentation given by S. Finkelstein in the conference *Astronomy in the 2020s: Synergies with WFIRST*, available at the time of writing at https://www.stsci.edu/~dlaw/WFIRST2020s/slides/finkelstein.pdf

We also use Eq. 6.42 to estimate galaxy biases in both photometric and high-redshift samples.

6.7 Results

An Example Cross Power Spectrum

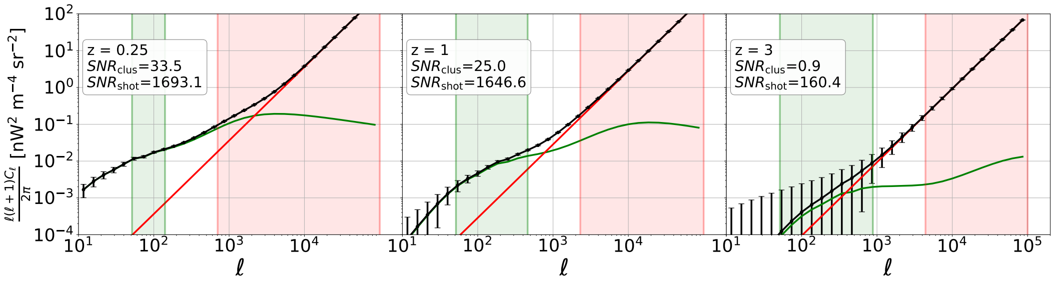

Figure 6.4: Examples of cross power spectra (black) of SPHEREx 3 μm image and SPHEREx spectroscopic galaxy catalog at $z = 0.25$, 1, and 3. Green and red lines are the clustering and Poisson-noise components, respectively. We constrain $b_I(z)d(\nu I_\nu)/dz(z,\nu)$ and $d\nu I_\nu/dz|_g\,(z,\nu)$ from the amplitudes of clustering and Poisson-noise modes in green and red bands (see Sec. 6.4), respectively, and their SNR values are given in the legend.

In Fig. 6.4, we show a set of example cross power spectra of a all-sky SPHEREx intensity map at 3 μm and SPHEREx spectroscopic catalog at $z = 0.25$, 1, and 3, respectively. The cross power spectra are given by Eq. 6.27 with error bars estimated by Eq. 6.4. In the Poisson-noise regime, we expect high significance measurements of the cross power spectra at all three redshifts due to the large number of modes on small scales, while in the clustering regime, we expect significant detection at the lower two redshifts only. For the $z = 3$ case, the low tracer number density gives a high galaxy Poisson noise level that reduces sensitivity on the cross power spectrum (see Sec. 6.8 for discussions).

We aim to extract two quantities related to the redshift-dependent IGL spectra, $b_I(z)d(\nu I_\nu)/dz(z,\nu)$ and $d(\nu I_\nu)/dz|_g\,(z,\nu)$ from the large (clustering) and small (Poisson-noise) scale amplitudes of the cross power spectrum. Their respective SNRs are denoted as $\mathrm{SNR_{clus}}$ and $\mathrm{SNR_{sh}}$, respectively. Note that for $\mathrm{SNR_{sh}}$, we also include the non-Gaussian effect from the trispectrum term (Eq. 6.48).

Sensitivity to the Clustering Term

The top row of Fig. 6.5 shows our model for $b_I(z)d(\nu I_\nu)/dz(z,\nu)$ as a function of the 96 SPHEREx spectral bins and at the five fiducial $z \leqslant 3$ redshifts and masking

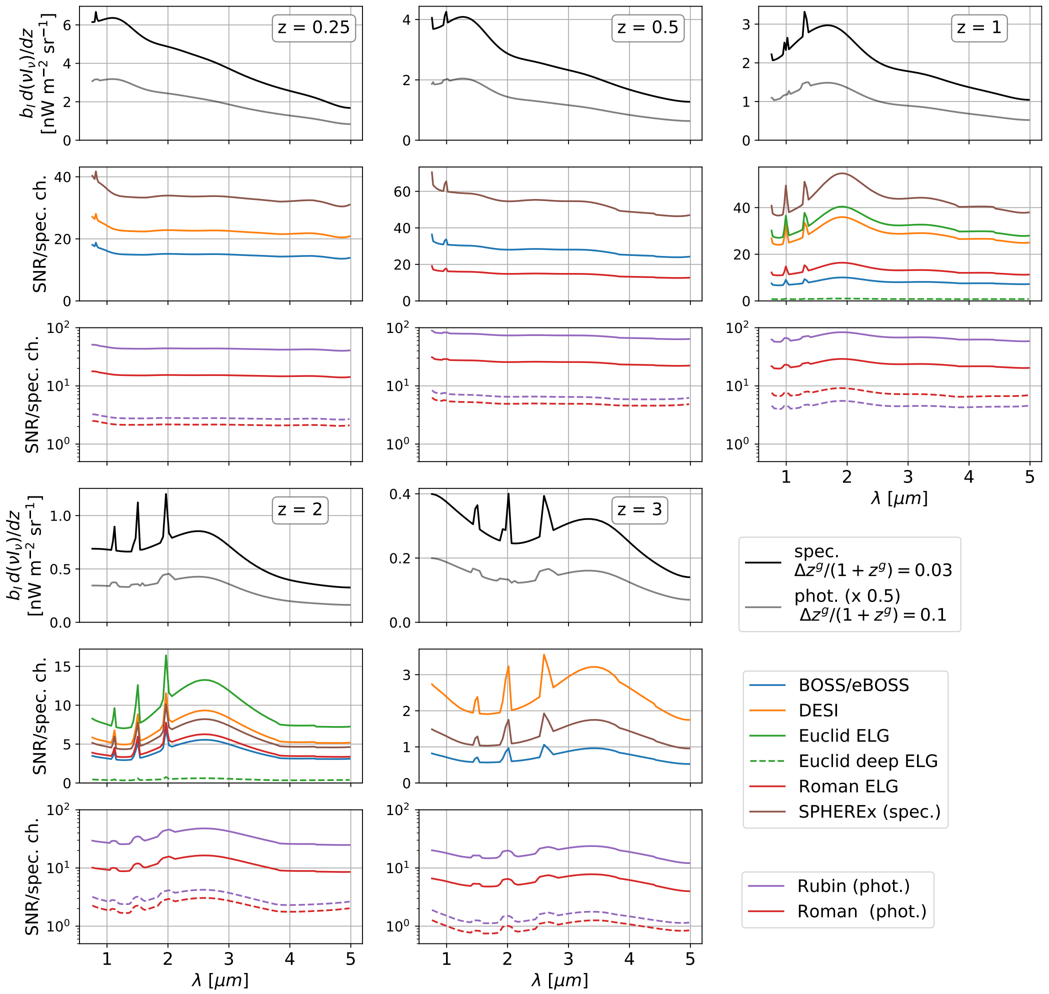

Figure 6.5: Constraints on the amplitude of cross power spectra of SPHEREx and galaxy surveys on large scales at $z \leqslant 3$. Top: $b_I(z)d(\nu I_\nu)/dz(z,\nu)$ from our model using spectroscopic (black) and photometric (grey, shifted down by ×0.5 for presentation purposes) galaxy samples. The photometric surveys have wider redshift bins and thus lower spectral resolution. Middle: SNR forecast on $b_I(z)d(\nu I_\nu)/dz(z,\nu)$ ($\mathrm{SNR_{clus}}$ in Eq. 6.47) per SPHEREx spectral channel from cross-correlating SPHEREx images with different spectroscopic surveys. For the surveys that have multiple tracer catalogs, we plot the one that gives the highest SNR values (BOSS CMASS and DESI BGS for $z = 0.25$; BOSS CMASS for $z = 0.5$; eBOSS QSO and DESI ELG for $z = 1$; eBOSS QSO and DESI QSO for $z = 2$ and 3). Solid lines are the cases using SPHEREx all-sky survey, and dashed lines are the cases with two SPHEREx deep fields, which have lower instrument noise but smaller sky coverage. Bottom: same as the middle rows but with photometric surveys.

thresholds. This quantity can be derived from the amplitude of the cross power spectra on large scales. We consider both spectroscopic and photometric galaxy surveys, where we use different redshift bin widths ($\Delta z^g/(1+z^g) = 0.03$ and 0.1 for spectroscopic and photometric, respectively) to account for the respective redshift accuracies. For cross-correlation with photometric galaxies, the SPHEREx spectral line features are smoothed out due to the wide redshift bins.

The middle and bottom rows of Fig. 6.5 show the SNR forecasts on $b_I(z)d(\nu I_\nu)/dz(z,\nu)$, i.e., $\mathrm{SNR}_{\mathrm{clus}}$ in Eq. 6.47, per SPHEREx spectral channel by cross-correlating SPHEREx maps with different spectroscopic (middle) and photometric (bottom) galaxy catalogs. For surveys that have multiple tracer catalogs, we plot the ones that give the highest SNR values. With spectroscopic galaxies, we expect to achieve a $\mathrm{SNR}_{\mathrm{clus}} > 5$ out to $z \sim 2$. However, at $z = 3$, due to the low number density of tracers, $\mathrm{SNR}_{\mathrm{clus}} \sim 1$ to 3. Photometric galaxies, on the other hand, have a much higher source density, and can reach a $\mathrm{SNR}_{\mathrm{clus}}$ of $\sim 10(5)$ with Rubin observatory LSST (Roman Space Telescope HLS) at $z = 3$. Dashed lines indicate results from SPHEREx deep fields, which have lower instrument noise but smaller sky coverage. According to our model, the SNR in the deep fields are much lower than the all-sky cases, due to their small f_{sky} that results in a large sample variance noise contribution. Note that with our chosen redshift binning there are ~ 37 measurements at $z < 2$ and ~ 9 measurements at $2 < z < 3$.

Fig. 6.6 shows the same forecast as Fig. 6.5 at high redshifts ($4 < z < 10$) with photometric surveys. Note with our chosen redshift binning there are ~ 6 measurements available for $3 < z < 6$. For $z > 6$ where the universe has not yet fully ionized, we cut off the spectrum below the Lyman-α wavelength (121.6 nm rest frame) since those high-energy photons will be absorbed by the neutral gas in the intergalactic medium.

The Lyman break selected galaxy bins from Rubin Observatory LSST and Roman Space Telescope are also shown in the panels close to their central redshift. Note that the Lyman break galaxy samples have a much wider redshift bin, and therefore they have higher sensitivity but lower spectral resolution on the inferred IGL spectrum.

Our estimate suggests that the IGL spectrum can be robustly measured over the full $0.75 - 5\mu$m wavelength range across redshift, to a significance level of a few σ per SPHEREx spectral channel out to $z = 6$ by cross-correlating SPHEREx all-sky and upcoming photometric galaxy surveys.

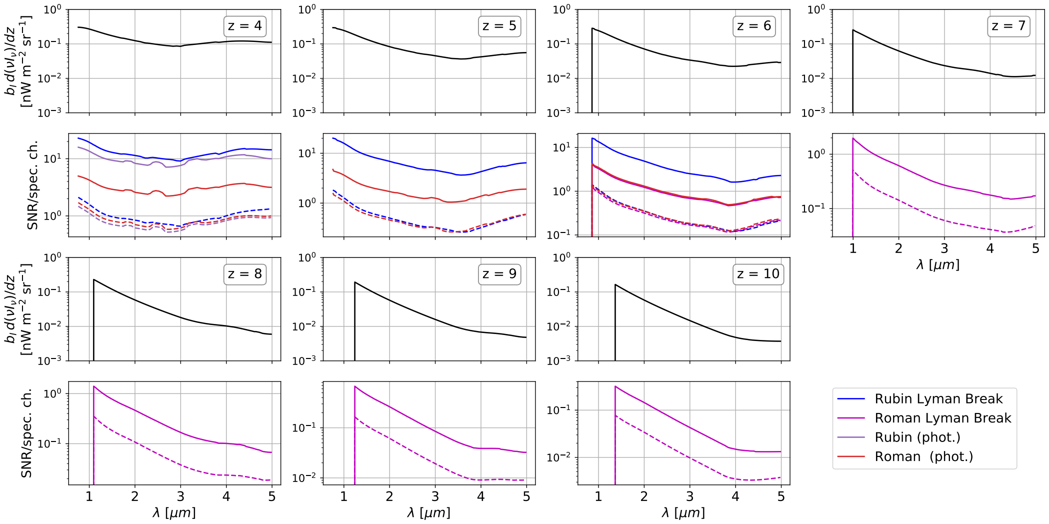

Figure 6.6: Constraints on the amplitude of cross power spectra of SPHEREx and galaxy surveys on large scales at $4 < z < 10$. Top: $b_I(z)d(\nu I_\nu)/dz(z,\nu)$ from our model using photometric surveys. We cut off the spectrum below the Lyman-α (121.6 nm) for $z > 6$ since those high-energy photons will be absorbed by neutral gas in the intergalactic medium. Bottom: SNR forecast on $b_I(z)d(\nu I_\nu)/dz(z,\nu)$ ($\mathrm{SNR_{clus}}$ in Eq. 6.47) per SPHEREx spectral channel from cross-correlating SPHEREx images with different photometric surveys. The Lyman break selected galaxy bins from Rubin Observatory LSST and the Roman Space Telescope are also shown in the panels close to their central redshifts. Solid lines are the cases using SPHEREx all-sky survey, and dashed lines are the cases with two SPHEREx deep fields.

Sensitivity to the Poisson-noise Term

Fig. 6.7 shows our forecast on the cross power spectrum Poisson noise amplitude of SPHEREx and galaxy surveys at $z \leqslant 3$. The top row compares the IGL spectrum from all sources, $d(\nu I_\nu)/dz$, and from the subset of sources in the galaxy samples used for cross-correlation, $d(\nu I_\nu)/dz|_g$, which is proportional to the Poisson noise amplitude. Naturally, deeper catalogs capture a larger fraction of the total IGL in the cross Poisson noise amplitude. The bottom row shows our SNR forecast on $d(\nu I_\nu)/dz|_g$, i.e. $\mathrm{SNR_{sh}}$, in Eq. 6.48. In most cases, the cross-correlation results in a highly significant $\mathrm{SNR_{sh}}$ ($\sim 10^2 - 10^4$) because of the large number of modes available on small scales. Dashed lines indicate the cross-correlation of the two SPHEREx deep field images with a Euclid deep field. Compared to all-sky, the deep fields capture a higher fraction of the IGL, as shown in the top row, but the SNR is lower due to the small sky coverage.

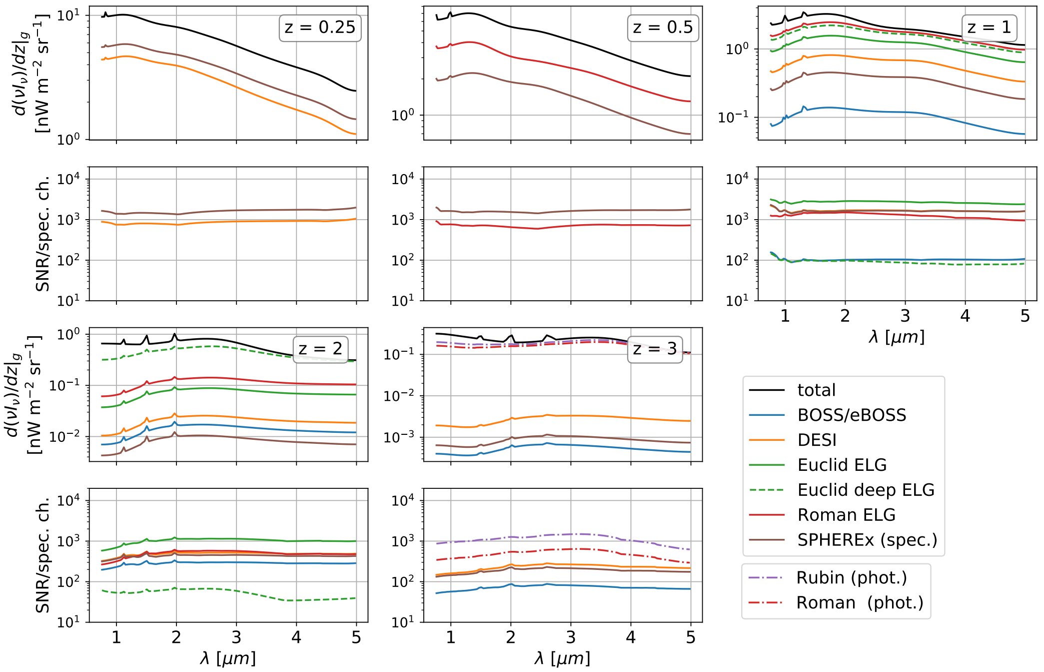

Figure 6.7: Constraints on the cross power spectrum Poisson noise amplitude at $z \leqslant 3$. Top: the IGL spectrum from all galaxies, $d(\nu I_\nu)/dz$ (black), and from the subset of the galaxies used for cross-correlations, $d\nu I_\nu/dz|_g$ (colored), which is proportional to the Poisson noise amplitude. Solid lines give the cases using the SPHEREx all-sky survey, and dashed lines are the cases with the two SPHEREx deep fields, which have lower instrument noise but smaller sky coverage. Bottom: SNR forecast on $d(\nu I_\nu)/dz|_g$ (SNR$_{\rm sh}$ in Eq. 6.48) per SPHEREx spectral channel from cross-correlating SPHEREx images with different spectroscopic surveys. For the surveys that have multiple tracer catalogs, we use the same tracer as in Fig. 6.5, which gives the highest SNR values. We use the fiducial masking depths as described in Sec. 6.4, and in the CMASS case at $z = 0.25$ and $z = 0.5$, the mask is deeper than the galaxy catalog, and therefore we only have clustering but not Poisson-noise measurement for these two cases.

Fig. 6.8 shows the same forecast as Fig. 6.7 at high redshifts ($4 < z < 10$) with photometric surveys and the Lyman-break selected samples. We observe a slight kink in SNR at $\lambda \sim 3.8$ μm due to the change in SPHEREx spectral resolution and thus instrument noise level (see Fig. 6.2). In practice, one can bin the measurements at $\lambda > 3.8$ μm to increase the sensitivity. The same effect is not evident at the low redshift cases as shown in Fig. 6.5, because instrument noise only becomes the dominant source of uncertainties in δC_ℓ^x (Eq. 6.4) at higher redshifts. Our forecast suggests that high significance measurements with SNR ~ 100, can be achieved out to $z = 6$ with upcoming photometric galaxy surveys.

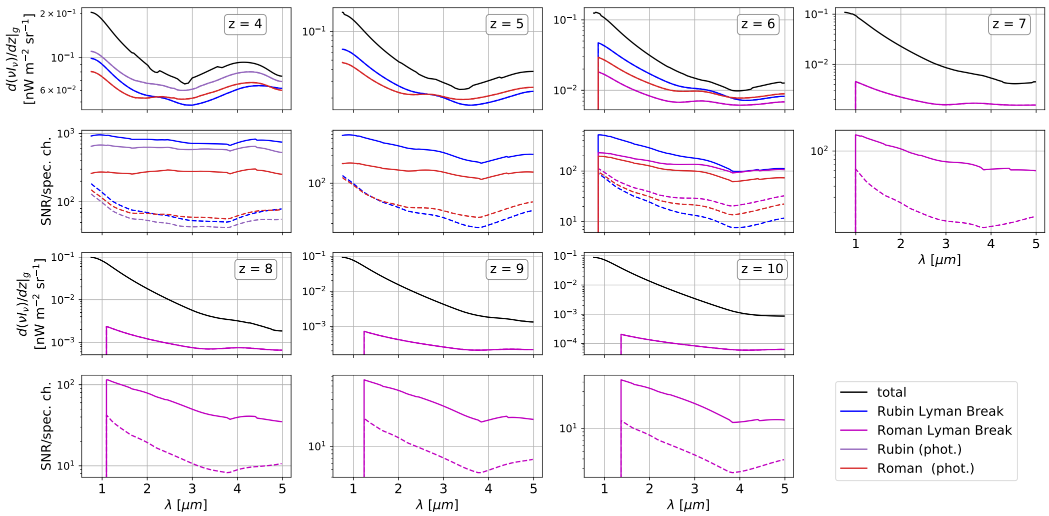

Figure 6.8: Constraints on the cross power spectrum Poisson noise amplitude for $4 < z < 10$. Top: the IGL spectrum from all galaxies, $d(\nu I_\nu)/dz$ (black), and from the subset of the galaxies used for cross-correlations, $d\nu I_\nu/dz|_g$ (colored), which is proportional to the Poisson noise amplitude. Solid lines give the cases using the SPHEREx all-sky survey, and dashed lines are the cases with the two SPHEREx deep fields, which have lower instrument noise but smaller sky coverage. The Lyman break selected galaxy bins from Rubin Observatory LSST and Roman Space Telescope are also shown in the panels close to their central redshifts. We cut off the spectrum below the Lyman-α wavelength (121.6 nm) for $z > 6$ since those high energy photons will be absorbed by neutral gas in the intergalactic medium. Bottom: SNR forecast on $d(\nu I_\nu)/dz|_g$ (SNR$_{\rm sh}$ in Eq. 6.48) per SPHEREx spectral channel from cross-correlations. The kink in SNR above $\lambda \sim 3.8$ μm is due to the change in SPHEREx resolving power and instrument noise (see Fig. 6.2).

Note that $d(\nu I_\nu)/dz|_g$ is the mean spectrum of sources in the galaxy catalog, which is not the same as the mean IGL spectrum from the averaged emission from all sources. Nevertheless, for simplicity, we assume any subset of galaxies has the same mean spectrum as the IGL model used in this calculation. We also highlight potentially different measurements from the cross Poisson noise term with an example of quasars in Sec. 6.9.

6.8 Discussion

Error Estimation

The variance of a cross power spectrum is the sum of several contributing factors in Eq. 6.4, including instrument noise, galaxy Poisson (shot) noise, and sample variance. Fig. 6.9 compares three cross power spectra and their different error

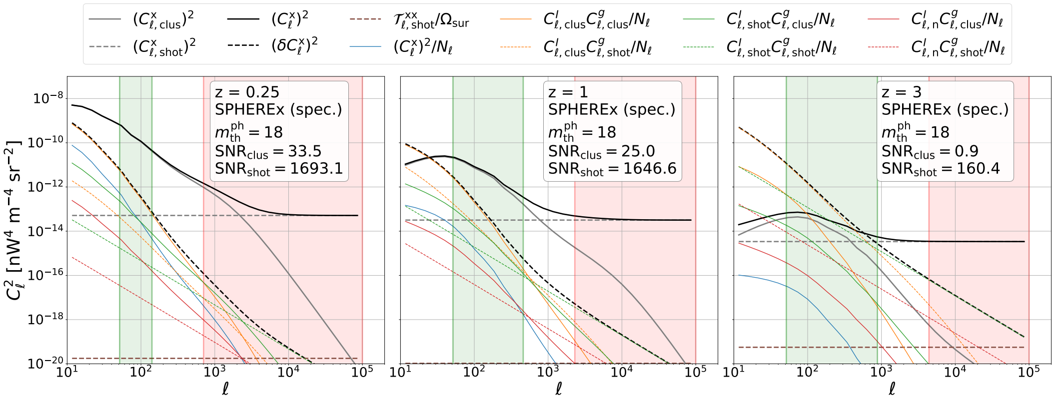

Figure 6.9: Cross power spectra (black solid lines) and their error components (colored lines) from cross-correlating SPHEREx all-sky maps at 3-μm channel with SPHEREx spectroscopic catalogs at $z = 0.25$ (left), 1 (middle), and 3 (right), using a masking threshold of $m^{\rm ph}_{\rm th} = 18$. The thick black dashed lines show the total Gaussian noise power spectrum variance, i.e., the sum of all colored lines (except for the trispectrum term $\mathcal{T}^{\rm xx}_{\ell,\rm shot}/\Omega_{\rm sur}$). Grey solid and dashed lines are the clustering and Poisson noise terms in the cross power spectrum, respectively. The green and red bands mark the multipole modes used for constraining the clustering and Poisson noise amplitudes, respectively. The two lower redshift cases are background-limited where the cross spectrum errors are dominated by $C^I_{\ell,\rm clus}C^g_{\ell,\rm clus}/N_\ell$, whereas at $z = 3$, the sensitivity is also limited by the galaxy Poisson noise due to its low tracer number density.

components from cross correlating SPHEREx all-sky maps at 3-μm channel with SPHEREx spectroscopic catalogs at $z = 0.25$, 1, and 3 using a masking threshold of $m^{\rm ph}_{\rm th} = 18$. The Gaussian error on the cross power spectrum (Eq. 6.4) comprises a quadratic sum of individual power spectrum terms:

$$\begin{aligned}\left(\delta C^{\rm x}_\ell\right)^2 &= \frac{1}{N_\ell}\left[(C^{\rm x}_\ell)^2 + C^I_\ell C^g_\ell\right]\\ &= \frac{1}{N_\ell}\Big[\ (C^{\rm x}_\ell)^2 + \left(C^I_{\ell,\rm clus} + C^I_{\ell,\rm shot} + C^I_{\ell,n}\right)\\ &\qquad \cdot\left(C^g_{\ell,\rm clus} + C^g_{\ell,\rm shot}\right)\ \Big]\\ &= \frac{1}{N_\ell}\Big[\ (C^{\rm x}_\ell)^2 + C^I_{\ell,\rm clus}C^g_{\ell,\rm clus} + C^I_{\ell,\rm clus}C^g_{\ell,\rm shot}\\ &\qquad + C^I_{\ell,\rm shot}C^g_{\ell,\rm clus} + C^I_{\ell,\rm shot}C^g_{\ell,\rm shot}\\ &\qquad + C^I_{\ell,n}C^g_{\ell,\rm clus} + C^I_{\ell,n}C^g_{\ell,\rm shot}\ \Big]\,.\end{aligned}\tag{6.52}$$

First, we see that the cross- spectrum amplitude decreases with redshift following the redshift dependence of the IGL spectrum (Fig. 6.1) and the underlying matter density fluctuations. For the two lower redshift cases, the dominant noise term is $C^{I}_{\ell,\mathrm{clus}}C^{g}_{\ell,\mathrm{clus}}/N_\ell$. While $C^{g}_{\ell,\mathrm{clus}}$ is determined by the underlying cosmology, the only way to further bring down the cross power spectrum error is to reduce $C^{I}_{\ell,\mathrm{clus}}$, which is dominated by aggregate emission along line of sight. Therefore, these measurements are in the background-limited regime, where the background refers to all emission not from the redshift of interest. We can apply a deeper mask to suppress the background. However,this will also reduce the signal, and therefore there exists a masking threshold that optimizes the trade-off between foreground and signal removal. We detail this comparison in Sec. 6.8.

At $z = 3$, the dominant noise term becomes $C^{I}_{\ell,\mathrm{clus}}C^{g}_{\ell,\mathrm{shot}}/N_\ell$ due to the increased galaxy Poisson noise power spectrum amplitude, $C^{g}_{\ell,\mathrm{shot}}$, in available surveys. This means that a deeper catalog with a higher tracer number density can effectively improve the sensitivity.

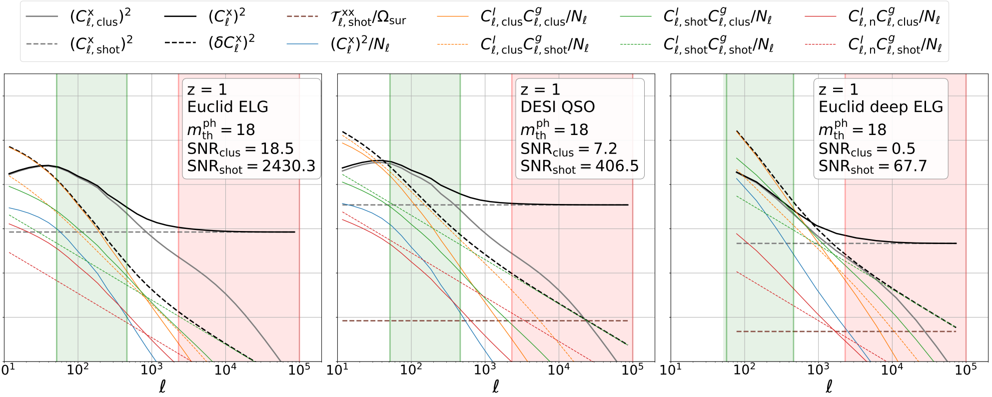

Figure 6.10: Cross power spectra (black solid lines) and their error components (colored lines) from cross-correlating SPHEREx all-sky maps at 3-μm channel with Euclid ELG (left) and DESI QSO (middle), and SPHEREx deep-field maps with the Euclid deep field ELG catalog (right). All three cases are at $z = 1$ with a masking threshold of $m^{\mathrm{ph}}_{\mathrm{th}} = 18$. The curves are the same as in Fig. 6.9. The Euclid ELG (left) and the DESI QSO (middle) have similar survey area, but the former has lower tracer number density and thus the higher Poisson noise suppresses the sensitivity. The Euclid deep field case (right) has even higher ELG number density than the wide field case (left), however it has lower SNR due to the larger sample variance from its smaller sky coverage.

Fig. 6.10 compares three different galaxy surveys at $z = 1$ with the same masking threshold of $m^{\mathrm{ph}}_{\mathrm{th}} = 18$ for the SPHEREx 3-μm channel. The first two, Euclid ELG and DESI QSO surveys, have a similar survey area, f_{sky}, and thus similar N_ℓ's, but Euclid ELG has a higher source number density. As a result, the two cases have similar cross spectrum amplitudes in the clustering regime, but DESI QSO has a higher Poisson noise amplitude, and the noise terms that contain $C^{g}_{\ell,\mathrm{shot}}$ (all dashed lines) are therefore higher in the DESI QSO case.

We can also compare the Euclid wide and deep field cross-correlations with the SPHEREx all-sky and deep fields, respectively. As shown on the left and right panels of Fig. 6.10, we see that the terms associated with galaxy Poisson noise (all dashed lines) are lower in the deep fields because of the higher galaxy number density. However, the deep fields cannot access the lower ℓ modes, and the overall noise is higher in the deep field due to the larger sample variance (small N_ℓ) from its smaller sky coverage.

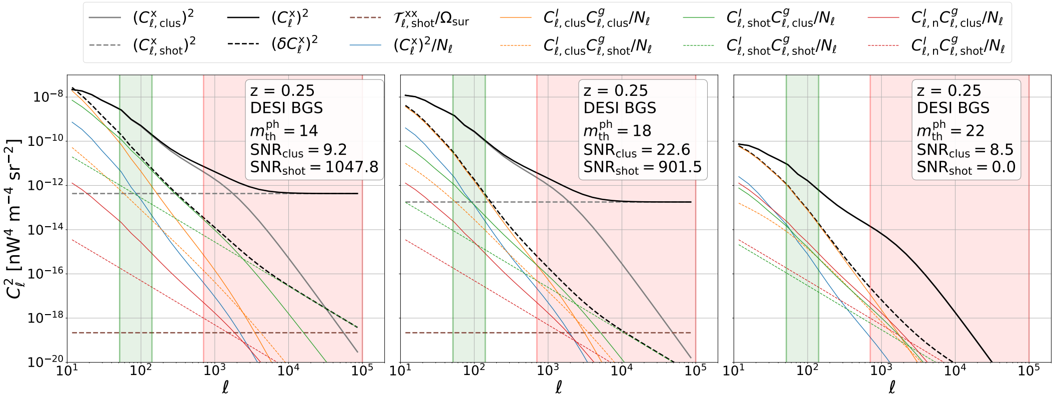

Figure 6.11: Cross power spectra (black solid lines) and their error components (colored lines) from cross-correlating SPHEREx all-sky maps at 3-μm channel with DESI BGS at $z = 0.25$ using a masking threshold of $m^{\mathrm{ph}}_{\mathrm{th}} = 14$ (left), 18 (middle), and 22 (right), respectively. The curves are the same as in Fig. 6.9. The SNR is improved by raising the masking threshold from $m^{\mathrm{ph}}_{\mathrm{th}} = 14$ to $m^{\mathrm{ph}}_{\mathrm{th}} = 18$, since the line-of-sight contamination decreases ($C^{\mathrm{I}}_{\ell,\mathrm{clus}}$ and $C^{\mathrm{I}}_{\ell,\mathrm{shot}}$ terms), while signal reduction is negligible (black solid lines). However, if we further increase the masking threshold to $m^{\mathrm{ph}}_{\mathrm{th}} = 22$, the SNR becomes lower due to the significant loss in the signal.

Fig. 6.11 demonstrates the impact of masking depth, using the cross-correlation of SPHEREx all-sky maps at 3 μm and DESI BGS galaxies at $z = 0.25$ with three different masking depths, $m^{\mathrm{ph}}_{\mathrm{th}} = 14, 18, 22$. Comparing $m^{\mathrm{ph}}_{\mathrm{th}} = 14$ and $m^{\mathrm{ph}}_{\mathrm{th}} = 18$, the signals are almost the same (black solid line) but the total Gaussian error (black dashed line) is lower with a deeper mask; this suggests that by masking sources to $14 < m^{\mathrm{ph}}_{\mathrm{th}} < 18$, a significant portion of the foreground can be removed, while the signal reduction is negligible which leads to an improvement of SNR. However, if we further increase the masking depth to $m_{\rm th}^{\rm ph} = 22$, the signal is drastically reduced, which means a significant fraction of IGL at this redshift is from galaxies with $18 < m_{\rm th}^{\rm ph} < 22$, and we cannot get a higher SNR by this deeper masking depth.

Masking Depth

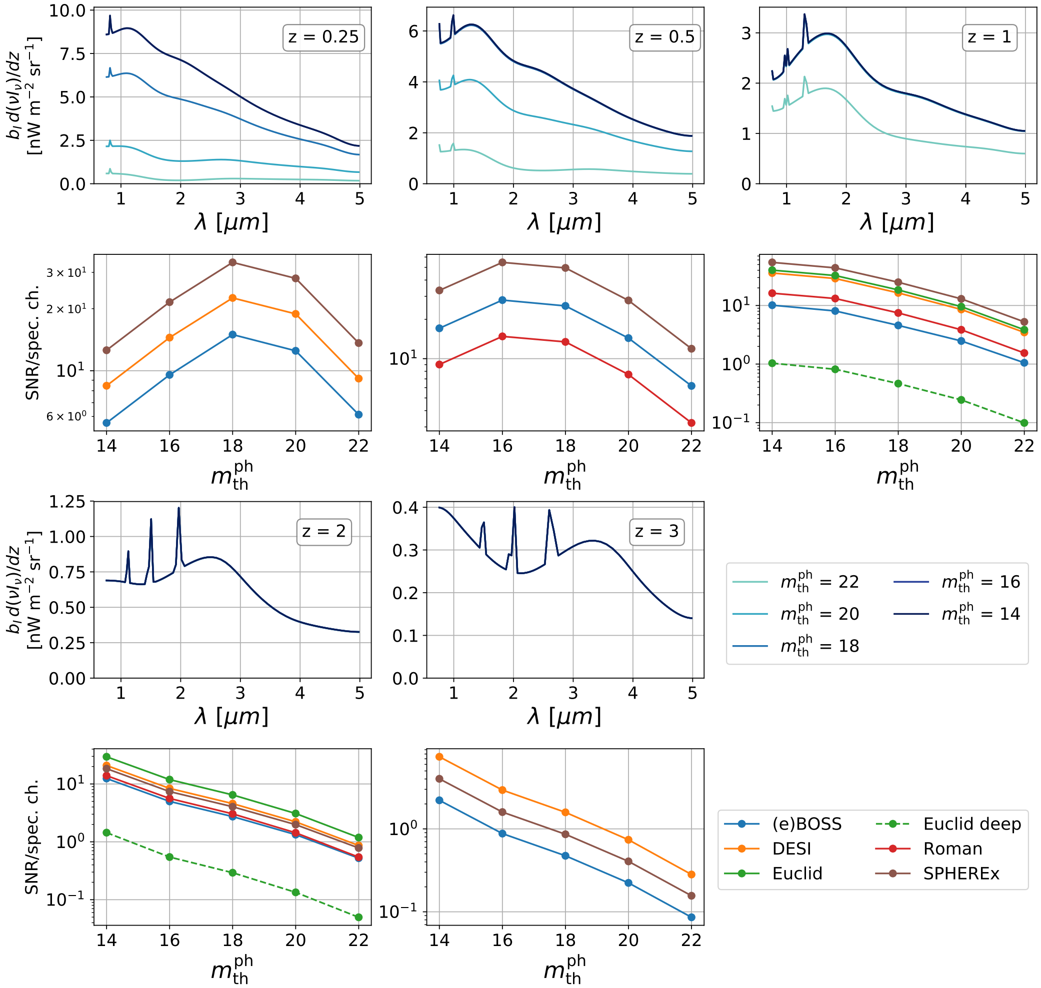

Figure 6.12: Top: the IGL spectrum $d(\nu I_\nu)/dz$ as a function of masking depth in the SPHEREx 3 μm channel. Bottom: SNR$_{\rm clus}$ as a function of $m_{\rm th}^{\rm ph}$ with different galaxy surveys shown in Fig. 6.5.

We consider masking all sources from an external catalog that are brighter than a flux threshold, and therefore the choice of masking depth is a trade-off between signal loss and foreground reduction. A deeper masking depth tends to remove more foreground sources contributing to the intensity maps, but at the same time, bright galaxies in the target redshift range of interest can also be masked. Fig. 6.12 demonstrates how the IGL amplitude and the error depends on masking depth. At lower redshift, the IGL amplitude strongly depends on the masking depths $m_{\text{th}}^{\text{ph}}$. At higher redshift, the signals are almost unchanged with varying masking depth, as the sources that dominate the IGL at these redshifts are below the masking thresholds. In other words, at low redshift, there is a masking depth that optimizes the tradeoff between signal loss and foreground reduction. At higher redshift, deeper masks perform better since they reduce more foreground emission while having negligible impact on the signal.

Here we only consider masking all stars and galaxies brighter than a magnitude threshold. In practice, if the external catalog provides redshift information of individual sources, a redshift-dependent magnitude threshold might also help to minimize foreground contamination from lower-redshift clustering signals. We leave this investigation to future work.

We ignore the effect of pixel loss due to masking in our forecast. In reality, masking results in highly non-uniform images, which not only reduces the number of multipole modes in the data, but also introduces mode mixing, the coupling of signal on different scales that needs to be corrected for in analysis. The reduction in power spectral sensitivity due to masking as a function of multipole depends on the source luminosity function and clustering amplitude, the masking size around each source, and the mode-coupling correction method. A more detailed study requires realistic simulated images, which is beyond the scope of this work.

Caveats of the Abundance Matching Model

In our model, we use abundance matching to relate the underlying matter distribution to galaxy tracers, continuum emitters, and spectral lines. In other words, if we have N galaxy tracers, we assume they are in the most massive N dark matter halos, and are also the strongest emitters of continuum in all frequencies and spectral lines. In reality, these relations have large scatters. It has been shown that red and blue galaxies have different spatial clustering and thus they trace the matter density field differently [107]. Each galaxy has distinct color and spectral line strength depending on its stellar composition, metallicity, dust attenuation, and other ISM and IGM properties. Moreover, the tracer galaxies are selected with certain spectral features, which makes them unrepresentative for the average EBL emitters. For example, the ELG samples will bias toward the star-forming galaxies which have strong emission lines. The decorrelations between matter, continuum, and lines will not only introduce uncertainties to our model, but also reduce the cross correlation signal. As discussed in Chiang et al. [19], the stochasticity in the relation between mass, tracers, and EBL light could be a possible explanation to their low EBL intensity bias values. Quantifying these decorrelations requires a coherent model for matters, continuum and lines, which we leave to future work.

Intensity-bias Degeneracy

We derive constraints on $b_I(z)d(\nu I_\nu)/dz(z,\nu)$ from the cross power spectrum amplitude in the clustering regime, but we do not further consider decoupling these two terms. In practice, we can apply the methods used in [19] to jointly fit the redshift an frequency dependence of bias and intensity, using all the cross spectrum measurements. In [19], they use the intensity from resolved sources to calibrate intensity at local universe, as the source catalog has sufficiently depth that contains most of the intensity emitters. For our case, we can cross-correlate SPHEREx maps with a deep catalog at low redshifts, and assume $d(\nu I_\nu)/dz|_g \approx d(\nu I_\nu)/dz$ to break the degeneracy. According to our model, SPHEREx spectroscopic catalog at $z = 0.1$ contains $\sim 95\%$ of the total EBL intensity, and thus this is feasible for our measurements.

6.9 Science Interpretation

Poisson Noise Constraints on the Tracer Spectrum

In the Poisson-noise sensitivity forecast presented in Sec. 6.7, we assume the tracer population has the same averaged spectrum as our IGL model spectrum. This is not necessarily true if we select a certain type of source for cross correlation (e.g., ELG, QSO), where the Poisson noise signal will be the averaged spectrum of the tracer sources. Therefore we can extract potentially distinct information from the Poisson-noise spectrum. As a demonstration, we consider a case of cross-correlating SPHEREx with quasars from DESI at $z = 2$, using a quasar spectrum and luminosity function model described in Sec. 6.3. The constraints on the averaged quasar spectrum, $d(\nu I_\nu)/dz|_Q$, from the cross Poisson spectrum is shown in Fig. 6.13. Using the small-scale cross-correlation information of SPHEREx with the quasar sample, we can obtain a high- SNR measurement on the average quasar spectrum at near-infrared wavelengths across redshift.

Given the high SNR value from the Poisson-noise measurements in all cases considered in this work, it is also feasible to gain more information from Poisson-noise by splitting the tracers into sub-samples by different properties. For example, we can select galaxies by their stellar mass, star formation rate, morphology, and constrain their average spectrum individually.

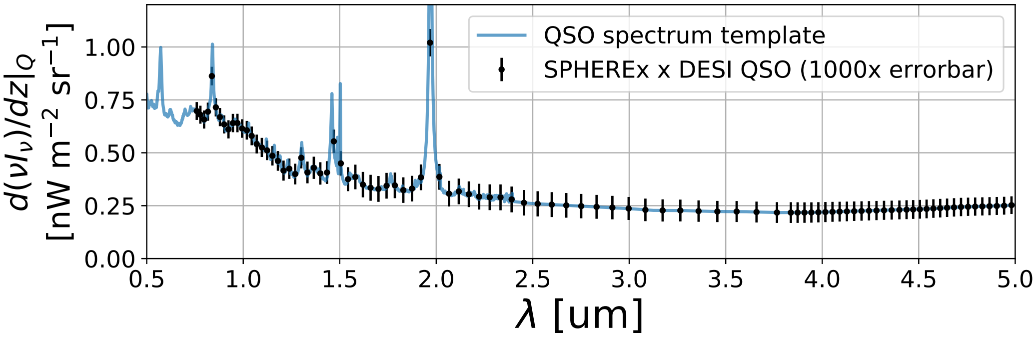

Figure 6.13: Average quasar's spectrum constraints (per SPHEREx spectral bin) from cross-correlating SPHEREx with DESI quasar samples at $z = 2$. The blue line is the modeled quasar spectrum template from Mas-Ribas & Hennawi [64].

Science from Non-Linear Clustering

In this work, we only consider information from the linear and Poisson noise scales of the cross power spectrum. Non-linear clustering scales also contain crucial information on galaxy evolution as well as the underlying cosmology. For example, emission from satellite galaxies and the contribution of IHL can be constrained by the amplitude and shape of the non-linear power spectrum [e.g., 16, 23, 109]. As we have demonstrated that a high-SNR measurement can be achieved by cross-correlating SPHEREx with current or upcoming galaxy surveys, it is promising to extract more constraints from non-linear scales.

Science from Near-Infrared EBL Tomography

The rest-frame near-infrared light from galaxies are dominated by the low mass stars, and thus the near-infrared photometry has been used for measuring the stellar mass function across redshifts. Most of the current stellar mass function constraints are based on deep photometric samples [65, 99, 103]. While individual galaxies can be resolve down to $\sim 10^{-3} M*$ at lower redshift, for $z \gtrsim 3$, the photometric surveys only complete to $\sim M*$ scales, and the faint end estimation are extrapolated from the bright end with a functional fit [99]. SPHEREx EBL measurement, $d\nu I_\nu/dz(\nu, z)$, provides complimentary information to the current detection-based measurement. Although SPHEREx cannot extract the individual galaxy spectra, SPHEREx EBL measurements contain emission from all sources with high spectral and redshift resolution, and a large sky coverage, which can be used to calibrate the stellar mass function model build from the resolved galaxies, and also probe the emission from diffuse origins such as the IHL. At higher redshift, SPHEREx bands cover the rest-frame optical and ultraviolet spectra, which can also help constraining the star formation history measurements [62].

Science from Spectral Lines

Here we do not separate spectral lines and continuum in our results. Our forecast suggest that the lines can be detected with high significance at lower redshift ($z \lesssim 3$) in both clustering and Poisson-noise regime. The line emission contains valuable information on the galaxy properties. For example, the Hα, Lyα, and [O II] line flux can be used for probing the star formation rate [12, 42, 51, 96]; the Balmer decrement, the ratio of Hα and Hβ line flux, is an crucial indicator to the dust extinctions [32]. The SPHEREx cross-correlation measurements of the average line emission have the potential to address these galaxy properties across redshifts. Moreover, with the high spectral resolution and sensitivity, SPEHREx can also probe the three-dimensional large scale structures with line intensity mapping [54], which measures the emission field from purely the spectral lines.

Cosmological Constraints

So far we have focused on, EBL spectrum constraints derived from cross-correlation, which only depend on the large-scale and Poisson noise amplitude of the cross power spectrum. In general, abundant cosmological information can be inferred from the power spectrum; for example, baryon acoustic oscillation signals can be extracted from the cross power spectrum, which can then be used to constrain the growth and geometry of the universe [37].

6.10 Conclusion

We forecast the sensitivity of near-infrared EBL tomography using spectro-imaging from the upcoming SPHEREx mission through cross-correlation with several current and future galaxy surveys, spanning redshifts of $0.25 < z < 10$. We consider IGL as the only emitting source that constitutes the near-infrared EBL, and build a model for IGL continuum and spectral line emissions. We then model the cross power spectrum of SPHEREx images and galaxy surveys as a function of SPHEREx spectral channel and galaxy redshift.

From the amplitudes of the cross power spectrum on linear scales, we infer the redshift-dependent EBL spectrum, multiplied by a bias factor, $b_I(z)d(\nu I_\nu)/dz(z,\nu)$.

Our forecast suggests that this quantity can be constrained to a significance of at least a few σ out to $z = 6$, using the SPHEREx all-sky survey in cross-correlation with a combination of spectroscopic (at $z \lesssim 3$) and photometric galaxy surveys (at $0.25 < z < 10$). At $3 \lesssim z \lesssim 10$, photometric galaxy catalogs from the upcoming Rubin Observatory and the Roman Space Telescope can constrain $b_I(z)d(\nu I_\nu)/dz(z,\nu)$, albeit at a lower redshift and spectral resolution.

The Poisson noise level of the cross power spectrum represents the average intensity of sources used in cross correlation. Our forecast shows that the Poisson noise level can be extracted at high significance ($\gtrsim 10^2$) out to $z \sim 10$, suggesting the SPHEREx all-sky and deep-field data can provide high sensitivity measurements of the average near-infrared spectrum of any selected population of tracer sources.

In summary, a high-sensitivity tomographic measurement of the EBL spectrum can be achieved by cross-correlating SPHEREx with current and future galaxy spectroscopic and photometric surveys. Making use of the clustering and Poisson noise of the cross power spectra, we expect further cosmological and astrophysical information can be extracted from this rich dataset.

References

[1] Abdalla, H., Adam, R., Aharonian, F., et al. 2020, MNRAS, 494, 5590, doi: 10.1093/mnras/staa999

[2] Abdo, A. A., Ackermann, M., Ajello, M., et al. 2010, ApJ, 723, 1082, doi: 10.1088/0004-637X/723/2/1082

[3] Abeysekara, A. U., Archer, A., Benbow, W., et al. 2019, ApJ, 885, 150, doi: 10.3847/1538-4357/ab4817

[4] Acciari, V. A., Ansoldi, S., Antonelli, L. A., et al. 2019, MNRAS, 486, 4233, doi: 10.1093/mnras/stz943

[5] Ackermann, M., Ajello, M., Allafort, A., et al. 2012, Science, 338, 1190, doi: 10.1126/science.1227160

[6] Ackermann, M., Ajello, M., Baldini, L., et al. 2018, ApJS, 237, 32, doi: 10.3847/1538-4365/aacdf7

[7] Aharonian, F., Akhperjanian, A. G., Bazer-Bachi, A. R., et al. 2006, Nature, 440, 1018, doi: 10.1038/nature04680

[8] Aharonian, F., Akhperjanian, A. G., Barres de Almeida, U., et al. 2007, A&A, 475, L9, doi: 10.1051/0004-6361:20078462

[9] Amendola, L., Appleby, S., Avgoustidis, A., et al. 2018, Living Reviews in Relativity, 21, 2, doi: 10.1007/s41114-017-0010-3

[10] Behroozi, P. S., Conroy, C., & Wechsler, R. H. 2010, ApJ, 717, 379, doi: 10.1088/0004-637X/717/1/379

[11] Bernstein, R. A. 2007, ApJ, 666, 663, doi: 10.1086/519824

[12] Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, MNRAS, 351, 1151, doi: 10.1111/j.1365-2966.2004.07881.x

[13] Cappelluti, N., Kashlinsky, A., Arendt, R. G., et al. 2013, ApJ, 769, 68, doi: 10.1088/0004-637X/769/1/68

[14] Caputo, A., Vittino, A., Fornengo, N., Regis, M., & Taoso, M. 2020, arXiv e-prints, arXiv:2012.09179. https://arxiv.org/abs/2012.09179

[15] Chang, T.-C., Pen, U.-L., Bandura, K., & Peterson, J. B. 2010, Nature, 466, 463, doi: 10.1038/nature09187

[16] Cheng, Y.-T., Arai, T., Bangale, P., et al. 2021, arXiv e-prints, arXiv:2103.03882. https://arxiv.org/abs/2103.03882

[17] Chiang, Y.-K., Makiya, R., Ménard, B., & Komatsu, E. 2020, ApJ, 902, 56, doi: 10.3847/1538-4357/abb403

[18] Chiang, Y.-K., & Ménard, B. 2019, ApJ, 870, 120, doi: 10.3847/1538-4357/aaf4f6

[19] Chiang, Y.-K., Ménard, B., & Schiminovich, D. 2019, ApJ, 877, 150, doi: 10.3847/1538-4357/ab1b35

[20] Conroy, C., Wechsler, R. H., & Kravtsov, A. V. 2006, ApJ, 647, 201, doi: 10.1086/503602

[21] Conroy, C., & White, M. 2013, ApJ, 762, 70, doi: 10.1088/0004-637X/762/2/70

[22] Cooray, A. 2016, Royal Society Open Science, 3, 150555, doi: 10.1098/rsos.150555

[23] Cooray, A., Smidt, J., de Bernardis, F., et al. 2012, Nature, 490, 514, doi: 10.1038/nature11474

[24] Cote, P., Abraham, B., Balogh, M., et al. 2019, in Canadian Long Range Plan for Astronomy and Astrophysics White Papers, Vol. 2020, 18, doi: 10.5281/zenodo.3758463

[25] Creque-Sarbinowski, C., & Kamionkowski, M. 2018, PhRvD, 98, 063524, doi: 10.1103/PhysRevD.98.063524

[26] Croft, R. A. C., Miralda-Escudé, J., Zheng, Z., Blomqvist, M., & Pieri, M. 2018, MNRAS, 481, 1320, doi: 10.1093/mnras/sty2302

[27] Croft, R. A. C., Miralda-Escudé, J., Zheng, Z., et al. 2016, MNRAS, 457, 3541, doi: 10.1093/mnras/stw204

[28] Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, AJ, 145, 10, doi: 10.1088/0004-6256/145/1/10

[29] Dawson, K. S., Kneib, J.-P., Percival, W. J., et al. 2016, AJ, 151, 44, doi: 10.3847/0004-6256/151/2/44

[30] DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 2016, arXiv e-prints, arXiv:1611.00036. https://arxiv.org/abs/1611.00036

[31] Domínguez, A., Primack, J. R., Rosario, D. J., et al. 2011, MNRAS, 410, 2556, doi: 10.1111/j.1365-2966.2010.17631.x

[32] Domínguez, A., Siana, B., Henry, A. L., et al. 2013, ApJ, 763, 145, doi: 10.1088/0004-637X/763/2/145

[33] Doré, O., Bock, J., Ashby, M., et al. 2014, ArXiv e-prints. https://arxiv.org/abs/1412.4872

[34] Doré, O., Werner, M. W., Ashby, M. L. N., et al. 2018, arXiv e-prints, arXiv:1805.05489. https://arxiv.org/abs/1805.05489

[35] Doré, O., Hirata, C., Wang, Y., et al. 2018, arXiv e-prints, arXiv:1804.03628. https://arxiv.org/abs/1804.03628

[36] Driver, S. P., Andrews, S. K., Davies, L. J., et al. 2016, ApJ, 827, 108, doi: 10.3847/0004-637X/827/2/108

[37] Eisenstein, D. J., Zehavi, I., Hogg, D. W., et al. 2005, ApJ, 633, 560, doi: 10.1086/466512

[38] Girardi, L., Bertelli, G., Bressan, A., et al. 2002, A&A, 391, 195, doi: 10.1051/0004-6361:20020612

[39] Girardi, L., Groenewegen, M. A. T., Hatziminaoglou, E., & da Costa, L. 2005, A&A, 436, 895, doi: 10.1051/0004-6361:20042352

[40] Gong, Y., Cooray, A., Mitchell-Wynne, K., et al. 2016, ApJ, 825, 104, doi: 10.3847/0004-637X/825/2/104

[41] Gong, Y., Cooray, A., Silva, M. B., et al. 2017, ApJ, 835, 273, doi: 10.3847/1538-4357/835/2/273

[42] Gunawardhana, M. L. P., Hopkins, A. M., Bland-Hawthorn, J., et al. 2013, MNRAS, 433, 2764, doi: 10.1093/mnras/stt890

[43] H. E. S. S. Collaboration, Abdalla, H., Abramowski, A., et al. 2017, A&A, 606, A59, doi: `10.1051/0004-6361/201731200`

[44] Helgason, K., Ricotti, M., & Kashlinsky, A. 2012, ApJ, 752, 113, doi: `10.1088/0004-637X/752/2/113`

[45] Hill, J. C., Ferraro, S., Battaglia, N., Liu, J., & Spergel, D. N. 2016, Phys. Rev. Lett., 117, 051301, doi: `10.1103/PhysRevLett.117.051301`

[46] Hopkins, A. M., & Beacom, J. F. 2006, ApJ, 651, 142, doi: `10.1086/506610`

[47] Kashlinsky, A., Arendt, R. G., Ashby, M. L. N., et al. 2012, ApJ, 753, 63, doi: `10.1088/0004-637X/753/1/63`

[48] Kashlinsky, A., Arendt, R. G., Mather, J., & Moseley, S. H. 2005, Nature, 438, 45, doi: `10.1038/nature04143`

[49] Keenan, R. C., Barger, A. J., Cowie, L. L., & Wang, W. H. 2010, ApJ, 723, 40, doi: `10.1088/0004-637X/723/1/40`

[50] Kennicutt, Robert C., J. 1998, Annual Review of Astronomy and Astrophysics, 36, 189, doi: `10.1146/annurev.astro.36.1.189`

[51] Khostovan, A. A., Sobral, D., Mobasher, B., et al. 2015, MNRAS, 452, 3948, doi: `10.1093/mnras/stv1474`

[52] Kim, M. G., Matsumoto, T., Lee, H. M., et al. 2019, PASJ, 71, 82, doi: `10.1093/pasj/psz063`

[53] Koushan, S., Driver, S. P., Bellstedt, S., et al. 2021, arXiv e-prints, arXiv:2102.12323. `https://arxiv.org/abs/2102.12323`

[54] Kovetz, E. D., Viero, M. P., Lidz, A., et al. 2017, ArXiv e-prints. `https://arxiv.org/abs/1709.09066`

[55] Kravtsov, A. V., Berlind, A. A., Wechsler, R. H., et al. 2004, ApJ, 609, 35, doi: `10.1086/420959`

[56] Lauer, T. R., Postman, M., Weaver, H. A., et al. 2020, arXiv e-prints, arXiv:2011.03052. `https://arxiv.org/abs/2011.03052`

[57] Laureijs, R., Amiaux, J., Arduini, S., et al. 2011, arXiv e-prints, arXiv:1110.3193. `https://arxiv.org/abs/1110.3193`

[58] Levenson, L. R., Wright, E. L., & Johnson, B. D. 2007, ApJ, 666, 34, doi: `10.1086/520112`

[59] LSST Science Collaboration, Abell, P. A., Allison, J., et al. 2009, arXiv e-prints, arXiv:0912.0201. `https://arxiv.org/abs/0912.0201`

[60] Lusso, E., Worseck, G., Hennawi, J. F., et al. 2015, MNRAS, 449, 4204, doi: 10.1093/mnras/stv516

[61] Ly, C., Malkan, M. A., Kashikawa, N., et al. 2007, ApJ, 657, 738, doi: 10.1086/510828

[62] Madau, P., & Dickinson, M. 2014, ARA&A, 52, 415, doi: 10.1146/annurev-astro-081811-125615

[63] MAGIC Collaboration, Albert, J., Aliu, E., et al. 2008, Science, 320, 1752, doi: 10.1126/science.1157087

[64] Mas-Ribas, L., & Hennawi, J. F. 2018, AJ, 156, 66, doi: 10.3847/1538-3881/aace5f

[65] Masui, K. W., Switzer, E. R., Banavar, N., et al. 2013, ApJL, 763, L20, doi: 10.1088/2041-8205/763/1/L20

[66] Matsumoto, T., Kim, M. G., Pyo, J., & Tsumura, K. 2015, ApJ, 807, 57, doi: 10.1088/0004-637X/807/1/57

[67] Matsumoto, T., & Tsumura, K. 2019, PASJ, 71, 88, doi: 10.1093/pasj/psz070

[68] Matsumoto, T., Seo, H. J., Jeong, W. S., et al. 2011, ApJ, 742, 124, doi: 10.1088/0004-637X/742/2/124

[69] Matsuura, S., Arai, T., Bock, J. J., et al. 2017, ApJ, 839, 7, doi: 10.3847/1538-4357/aa6843

[70] Mattila, K., & Väisänen, P. 2019, Contemporary Physics, 60, 23, doi: 10.1080/00107514.2019.1586130

[71] McQuinn, M., & White, M. 2013, MNRAS, 433, 2857, doi: 10.1093/mnras/stt914

[72] Ménard, B., Scranton, R., Schmidt, S., et al. 2013, arXiv e-prints, arXiv:1303.4722. https://arxiv.org/abs/1303.4722

[73] Merson, A., Smith, A., Benson, A., Wang, Y., & Baugh, C. 2019, MNRAS, 486, 5737, doi: 10.1093/mnras/stz1204

[74] Mitchell-Wynne, K., Cooray, A., Xue, Y., et al. 2016, ApJ, 832, 104, doi: 10.3847/0004-637X/832/2/104

[75] Mitchell-Wynne, K., Cooray, A., Gong, Y., et al. 2015, Nature Communications, 6, 7945, doi: 10.1038/ncomms8945

[76] Moster, B. P., Somerville, R. S., Maulbetsch, C., et al. 2010, ApJ, 710, 903, doi: 10.1088/0004-637X/710/2/903

[77] Newman, J. A. 2008, ApJ, 684, 88, doi: 10.1086/589982

[78] Nuza, S. E., Sánchez, A. G., Prada, F., et al. 2013, MNRAS, 432, 743, doi: 10.1093/mnras/stt513

[79] Osterbrock, D. E., & Ferland, G. J. 2006, Astrophysics of gaseous nebulae and active galactic nuclei

[80] Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 594, A13, doi: 10.1051/0004-6361/201525830

[81] Pullen, A. R., Chang, T.-C., Doré, O., & Lidz, A. 2013, ApJ, 768, 15, doi: 10.1088/0004-637X/768/1/15

[82] Pullen, A. R., Doré, O., & Bock, J. 2014, ApJ, 786, 111, doi: 10.1088/0004-637X/786/2/111

[83] Pullen, A. R., Serra, P., Chang, T.-C., Doré, O., & Ho, S. 2018, MNRAS, 478, 1911, doi: 10.1093/mnras/sty1243

[84] Raichoor, A., Comparat, J., Delubac, T., et al. 2017, MNRAS, 471, 3955, doi: 10.1093/mnras/stx1790

[85] Richards, G. T., Lacy, M., Storrie-Lombardi, L. J., et al. 2006, ApJS, 166, 470, doi: 10.1086/506525

[86] Saldana-Lopez, A., Domínguez, A., Pérez-González, P. G., et al. 2020, arXiv e-prints, arXiv:2012.03035. https://arxiv.org/abs/2012.03035

[87] Sano, K., Kawara, K., Matsuura, S., et al. 2015, ApJ, 811, 77, doi: 10.1088/0004-637X/811/2/77

[88] Sano, K., Matsuura, S., Yomo, K., & Takahashi, A. 2020, ApJ, 901, 112, doi: 10.3847/1538-4357/abad3d

[89] Schaan, E., Ferraro, S., & Spergel, D. N. 2018, PhRvD, 97, 123539, doi: 10.1103/PhysRevD.97.123539

[90] Schechter, P. 1976, ApJ, 203, 297, doi: 10.1086/154079

[91] Schmidt, S. J., Ménard, B., Scranton, R., et al. 2015, MNRAS, 446, 2696, doi: 10.1093/mnras/stu2275

[92] Scott, B., Upton Sanderbeck, P., & Bird, S. 2021, arXiv e-prints, arXiv:2104.00017. https://arxiv.org/abs/2104.00017

[93] Seo, H. J., Lee, H. M., Matsumoto, T., et al. 2015, ApJ, 807, 140, doi: 10.1088/0004-637X/807/2/140

[94] Serra, P., Lagache, G., Doré, O., Pullen, A., & White, M. 2014, A&A, 570, A98, doi: 10.1051/0004-6361/201423958

[95] Sheth, R. K., Mo, H. J., & Tormen, G. 2001, MNRAS, 323, 1, doi: 10.1046/j.1365-8711.2001.04006.x

[96] Sobral, D., Smail, I., Best, P. N., et al. 2013, MNRAS, 428, 1128, doi: 10.1093/mnras/sts096

[97] Spergel, D., Gehrels, N., Baltay, C., et al. 2015, arXiv e-prints, arXiv:1503.03757. https://arxiv.org/abs/1503.03757

[98] Thompson, R. I., Eisenstein, D., Fan, X., Rieke, M., & Kennicutt, R. C. 2007, ApJ, 666, 658, doi: 10.1086/520634

[99] Thorne, J. E., Robotham, A. S. G., Davies, L. J. M., et al. 2021, MNRAS, doi: 10.1093/mnras/stab1294

[100] Tsumura, K., Matsumoto, T., Matsuura, S., Sakon, I., & Wada, T. 2013, PASJ, 65, 121, doi: 10.1093/pasj/65.6.121

[101] Vale, A., & Ostriker, J. P. 2004, MNRAS, 353, 189, doi: 10.1111/j.1365-2966.2004.08059.x

[102] Vanden Berk, D. E., Richards, G. T., Bauer, A., et al. 2001, AJ, 122, 549, doi: 10.1086/321167

[103] Wright, A. H., Driver, S. P., & Robotham, A. S. G. 2018, MNRAS, 480, 3491, doi: 10.1093/mnras/sty2136

[104] Yang, S., Pullen, A. R., & Switzer, E. R. 2019, MNRAS, 489, L53, doi: 10.1093/mnrasl/slz126

[105] Yue, B., Ferrara, A., Salvaterra, R., Xu, Y., & Chen, X. 2013, MNRAS, 433, 1556, doi: 10.1093/mnras/stt826

[106] —. 2014, MNRAS, 440, 1263, doi: 10.1093/mnras/stu351

[107] Zehavi, I., Zheng, Z., Weinberg, D. H., et al. 2011, ApJ, 736, 59, doi: 10.1088/0004-637X/736/1/59

[108] Zemcov, M., Immel, P., Nguyen, C., et al. 2017, Nature Communications, 8, 15003, doi: 10.1038/ncomms15003

[109] Zemcov, M., Smidt, J., Arai, T., et al. 2014, Science, 346, 732, doi: 10.1126/science.1258168

[110] Zhai, Z., Benson, A., Wang, Y., Yepes, G., & Chuang, C.-H. 2019, MNRAS, 490, 3667, doi: 10.1093/mnras/stz2844

[111] Zheng, Z., Berlind, A. A., Weinberg, D. H., et al. 2005, ApJ, 633, 791, doi: 10.1086/466510