Multilayered Bayesian Inference for EoR 21-cm Signal: Gibbs Sampling and GCR Reconstruction

M.Sc. Thesis

Mayukh Mandal



Department of Astronomy, Astrophysics and Space Engineering

INDIAN INSTITUTE OF TECHNOLOGY INDORE

May 17, 2026

Multilayered Bayesian Inference for EoR 21-cm Signal: Gibbs Sampling and GCR Reconstruction

M.Sc.Thesis

Submitted in partial fulfillment of the requirements for the awards of the degree
of
Master of Science

by

Mayukh Mandal



Department of Astronomy, Astrophysics and Space Engineering

INDIAN INSTITUTE OF TECHNOLOGY INDORE

May 17, 2026

INDIAN INSTITUTE OF TECHNOLOGY INDORE
CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the thesis entitled "Physics under extreme magnetic fields: Probing the curious case of peculiar magnetars" in the partial fulfillment of the requirements for the award of the degree of MASTER OF SCIENCE and submitted in the DISCIPLINE OF ASTRONOMY, Indian Institute of Technology Indore, is an authentic record of my own work carried out during the time period from June, 2025 to May, 2026 under the supervision of Dr. Suman Majumdar, Associate professor.

The matter presented in this thesis has not been submitted by me for the award of any other degree of this or any other institute.

INDIAN INSTITUTE OF TECHNOLOGY INDORE

CANDIDATE'S DECLARATION

I hereby certify that the work which is being presented in the thesis entitled **Multilayered Bayesian Inference for EoR 21-cm Signal: Gibbs Sampling and GCR Reconstruction** in the partial fulfillment of the requirements for the award of the degree of **MASTER OF SCIENCE** and submitted in the **DEPARTMENT OF ASTRONOMY , ASTROPHYSICS AND SPACE ENGINEERING , Indian Institute of Technology Indore**, is an authentic record of my own work carried out during the time period from **June 2024 to May 2026** under the supervision of **Dr. Suman Majumdar, Associate Professor, Department of Astronomy, Astrophysics and Space Engineering.**

The matter presented in this thesis has not been submitted by me for the award of any other degree of this or any other institute.

Mayukh
08/05/26

Signature of the student with date
(Mayukh Mandal)

This is to certify that the above statement made by the candidate is correct to the best of my/our knowledge.

Suman Majumdar
07/05/2026

Signature of the Supervisor of
M.Sc. thesis (with date)

(Suman Majumdar)

Mayukh Mandal has successfully given his/her M.Sc. Oral Examination held on **06/05/26**.

Suman Majumdar
Signature(s) of Supervisor(s) of MSc thesis
Date: 07/05/2026

Manoneeta Chakraborty
Programme Coordinator, M.Sc.
Date: 08/05/2026

Manoneeta Chakraborty
Convenor, DPGC
Date: 08/05/2026

HoD, DAASE
Date: 08/05/2026

Acknowledgements

The past two years have been an incredible journey of learning and personal growth. I am deeply grateful to all the people who have supported and guided me along the way.

I would like to express my sincere gratitude to my supervisor, Dr. Suman Majumdar, for his constant guidance and support throughout this work. I deeply appreciate the help and mentorship he has provided over the past year. I am especially grateful for the freedom he gave me to explore my own ideas and choose the direction of my work, while always guiding me with clarity and insight whenever needed. I would also like to thank him for his understanding and support during challenging times, particularly for giving me time off during my backlogs and examinations. His encouragement during those phases made a significant difference.

Beyond academics, I am grateful for the friendly and welcoming environment he created in the lab. His sense of humour, the ease with which he interacted with everyone, and the many treats he shared with us made the lab not just a place of work, but a place I genuinely enjoyed being in. His mentorship has been truly invaluable, both academically and personally.

I am thankful to my labmates — Shirin, Mohit, Yashraj, and Renuka — for their help, discussions, and for making the lab a great place to work and learn. The personal banter we shared, the jokes we made at each other's expense, and the occasional rage-baiting made the experience all the more enjoyable. I would also like to extend my gratitude to Leon and Vishrut, whose names I had not mentioned earlier but who I have probably disturbed the most. Their patience and willingness to explain things repeatedly have meant a lot to me.

I would also like to thank my classmates — Manpreet, Nityananda, Kartik, and Kaivan — along with many others, for their support and companionship during this journey. The many trips to the city and the jokes we shared made this experience all the more memorable.

A special thanks to my friends — Rahul, Vishal, Jadura, and Sachin. The time we spent playing football and just loofing around made these years truly memorable. I am especially grateful for Rahul's cooking — without it, I might have run away long ago.

I am equally grateful to my friends back home — Dirpo, Arka, Pritam, Zaheer, and Zunaid — for the late-night Discord calls, football and gaming sessions that kept me going.

And finally, peanut butter and mixed fruit jam — for getting me through many long nights.

Abstract

The Cosmic Dawn and Epoch of Reionization mark a critical phase in the history of the Universe, during which the first luminous sources transformed the predominantly neutral intergalactic medium into an ionized state. The redshifted 21-cm signal, arising from the hyperfine transition of neutral hydrogen, provides a powerful probe of this era, encoding information about the thermal evolution, ionization history, and early structure formation. However, extracting physical parameters from this signal remains challenging due to its faintness, strong foreground contamination, and the computational cost of forward modelling.

In this work, we investigate Bayesian inference techniques for recovering key reionization parameters ($M_{\rm min}, N_{\rm ion}, R_{\rm mfp}$) from the 21-cm power spectrum. A Bayesian Neural Network (BNN) emulator is employed to provide fast predictions of the power spectrum along with associated uncertainties, enabling a realistic likelihood framework. In an initial study under idealized conditions, we compare two sampling strategies: a standard Metropolis–Hastings Markov Chain Monte Carlo sampler and a Metropolis-within-Gibbs approach. While both methods recover the correct posterior distributions, the Gibbs sampler demonstrates improved computational efficiency through higher acceptance rates and better mixing in the presence of parameter correlations.

The analysis is then extended to a more realistic setting by incorporating incomplete data arising from foreground contamination and associated foreground avoidance strategy. We develop a reconstruction-based framework in which missing modes (due to the avoidance of the foreground wedge in k-space) of the two-dimensional power spectrum are inferred using a Gaussian Constrained Realization-inspired approach. The results show that the reconstructed power spectra are consistent with the true signal within expected uncertainties across a range of avoidance scenarios, validating the feasibility of reconstruction-based inference.

This work establishes a two-stage framework combining efficient sampling and statistical reconstruction of the signal, and provides a foundation for future extensions to full three-dimensional signal reconstruction and emulator-based inference in the presence of foregrounds and observational systematics.

Table of Contents

1 Cosmic Dawn and Epoch of Reionization (EoR) 1
- 1.1 Introduction . . . 1
- 1.2 Key Observables: Lyman-α, CMB, and 21 cm Line . . . 2
  - 1.2.1 The Lyman-α Forest . . . 2
  - 1.2.2 CMB Anisotropies . . . 3
  - 1.2.3 21 cm Signal . . . 3

2 The 21 cm Signal 5
- 2.1 Observable . . . 5
- 2.2 Challenges: Foregrounds and Systematics . . . 6
- 2.3 Statistical Quantification through Power Spectrum . . . 7
- 2.4 From Observables to Parameter Inference . . . 8

3 Bayesian Inference and Sampling Techniques for EoR 21-cm Observations 10
- 3.1 Bayesian Framework for EoR Parameter Inference . . . 10
- 3.2 Fundamentals of Bayesian Inference . . . 10
- 3.3 Different Sampling techniques . . . 12
  - 3.3.1 Metropolis-Hastings Algorithm . . . 12
  - 3.3.2 Hamiltonian Monte Carlo . . . 13
  - 3.3.3 Gibbs Sampling . . . 14
    - 3.3.3.1 Metropolis-within-Gibbs . . . 16
  - 3.3.4 Diffusive Gibbs Sampling . . . 16
- 3.4 Why Gibbs Sampling? . . . 18

4 Inference Strategies for EoR 21-cm Signal Under Foreground Contamination 20
- 4.1 Introduction . . . 20
- 4.2 Inference via Foreground Avoidance . . . 20
- 4.3 Inference through Foreground Subtraction . . . 21
- 4.4 Foreground Modelling and Joint Inference . . . 21
- 4.5 Reconstruction-Based Approaches and Missing Data Inference . . . 22
- 4.6 Challenges in Inference with Incomplete Data . . . 23

5 Signal Reconstruction using Gaussian Constrained Realizations 24
- 5.1 Reconstruction-Based Inference: Overview . . . 24
- 5.2 Gaussian Constrained Realizations . . . 25
  - 5.2.1 Conceptual Framework . . . 25
  - 5.2.2 Mathematical Formulation . . . 25
  - 5.2.3 Interpretation . . . 25
  - 5.2.4 Role in Cosmological Signal Reconstruction . . . 26
- 5.3 Signal Reconstruction Framework in This Work . . . 26

5.3.1 Objective and Scope . . . 26
5.3.2 Reduced Representation as a Controlled Testbed . . . 27
5.3.3 Reconstruction of Missing Power Spectrum Modes . . . 27
5.3.4 Inference Pipeline for the Reconstructed Power Spectrum . . . 27
5.3.5 Implementation of Gaussian Constrained Realizations . . . 28
5.4 Discussion and Outlook . . . 28

6 Towards a Unified Inference Framework for EoR 21-cm Signal **29**
6.1 Motivation and Overview . . . 29
6.2 Inference Framework and Methodology . . . 30
6.3 Staged Evaluation of the Inference Framework . . . 31
6.3.1 Stage I: Sampler Validation in an Idealized Setting . . . 31
6.3.1.1 Likelihood Model and Error Covariance . . . 31
6.3.1.2 The Bayesian Neural Network . . . 32
6.3.1.3 Sampling Techniques . . . 33
6.3.1.4 Test Parameter Sets . . . 34
6.3.2 Stage II: Reconstruction with Incomplete Data . . . 34
6.4 Stage I: Validation of Gibbs Sampling Against Metropolis–Hastings . . . 35
6.4.1 Parameter Set 1: (176.32, 38.95, 33.40) . . . 35
6.4.2 Parameter Set 2: (675.26, 48.42, 17.26) . . . 36
6.4.3 Parameter Set 3: (467.37, 181.05, 12.65) . . . 37
6.4.4 Parameter Recovery and Sampling Efficiency . . . 37
6.5 Stage II: Reconstruction of the 2D Signal Power Spectrum . . . 38
6.5.1 Signal Reconstruction: Early Stages of Reionization . . . 39
6.5.2 Signal Reconstruction: Intermediate Stage of Reionization . . . 42
6.5.3 Signal Reconstruction: Late Stage of Reionization . . . 45
6.5.4 Interpretation . . . 48
6.6 Development of Emulator for the Reconstructed Signal and Future Directions 48

7 Summary and Discussion **50**

8 Future Work **52**
8.1 Completion of Emulator-Based Inference . . . 52
8.2 Beyond the Power Spectrum: Incorporating Higher-Order Statistics . . . 52
8.3 Towards Full Signal Reconstruction . . . 53
8.4 Application to Realistic Observations . . . 53
8.5 Outlook . . . 54

List of Figures

1.1 Evolution of the 21-cm brightness temperature from the pre-stellar era to the end of reionization, showing simulated fluctuations and the expected global signal shaped by early galaxy formation, heating, and ionization of the IGM. Adapted from [Pritchard and Loeb, 2010]. 2

1.2 Lyman-α forest spectral region for three quasars spanning a wide range of redshifts. Each spectrum shows numerous absorption features produced by intervening neutral hydrogen clouds along the line of sight, illustrating how the density of Lyman-α absorption evolves with redshift [McQuinn, 2016]. . 3

2.1 Global 21-cm signal during Cosmic Dawn and the Epoch of Reionization. The figure shows (i) the CMB temperature $T_{\rm CMB}$, gas kinetic temperature T_K, and spin temperature T_S (top), (ii) ionized fractions in galaxies x_i and diffuse regions x_e (middle), and (iii) the mean 21-cm brightness temperature δT_b for different astrophysical models (bottom) [Pritchard and Loeb, 2012]. 6

2.2 The power spectrum of the 21-cm signal at high redshifts, including all fluctuations in the spin temperature. Reproduced from [Santos et al., 2010]. 8

3.1 Converging views in Bayesian inference. Two scientists with different prior beliefs, $p(\theta|I_i)$ (panel a, red and green PDFs), observe a single datum with likelihood $L(\theta)$ (panel b). Their posterior distributions $p(\theta|m_1)$ (panel c), obtained via Bayes' theorem , represent updated knowledge of the parameter θ. After observing many data points, the posteriors become essentially indistinguishable, illustrating convergence of Bayesian inference. Adapted from Trotta [Trotta, 2008]. 12

3.2 A trace plot (left) and histogram (right) of 5000 Metropolis chain values of μ. The target posterior PDF is superimposed in blue.[Johnson et al., 2022] 13

3.3 Comparison of sampling efficiency for a 100-dimensional Gaussian. Shown are the sampled values of the variable with the largest standard deviation using a random-walk Metropolis run (left) and an HMC run with $L = 150$ (right). HMC demonstrates reduced correlation and faster exploration of the parameter space, illustrating its superior mixing compared to Metropolis [Neal, 2011]. 14

3.4 Trace of Gibbs sampling for two correlated parameters. The stair-step pattern illustrates sequential updates of θ_1 and θ_2. 15

3.5 Visualization of 10^4 samples for the 2D MoG-40 synthetic problem [Chen et al., 2024] generated by each sampler. All samplers are initialized at the origin. HMC and MALA fail to explore distant modes, PT covers all modes but with uneven sample distribution, and DiGS successfully captures all modes with balanced samples. 18

6.1 BNN prediction of the 21-cm power spectrum for a representative test case. 33

6.2 Trace plots for parameter set 1 35

6.3 Corner plots for parameter set 1 . . . 35
6.4 Trace plots for parameter set 2 . . . 36
6.5 Corner plots for parameter set 2 . . . 36
6.6 Trace plots for parameter set 3 . . . 37
6.7 Corner plots for parameter set 3 . . . 37
6.8 Early stage under optimistic masking. Top: true, masked, posterior mean, and sampled reconstruction. Bottom: relative error. The reconstruction remains accurate across all accessible modes. . . . 39
6.9 Early stage under intermediate masking. The global structure is recovered, though errors increase within the masked region, particularly at low-k modes. . . . 40
6.10 Early stage under pessimistic masking. Significant deviations appear in the masked region, particularly at low-k modes. . . . 41
6.11 Intermediate stage under optimistic masking. Both large- and intermediate-scale features are accurately recovered with low relative error. . . . 42
6.12 Intermediate stage under intermediate masking. Errors increase within the masked wedge but remain controlled. . . . 43
6.13 Intermediate stage under pessimistic masking. Small-scale features are increasingly smoothed, and errors grow in regions far from observed modes. 44
6.14 Late stage under optimistic masking. Large-scale modes are well reconstructed with low error, while small-scale deviations begin to appear. . . . 45
6.15 Late stage under intermediate masking. Errors remain low at large scales but increase at smaller scales within the masked region. . . . 46
6.16 Late stage under pessimistic masking. Large-scale structure is retained, while small-scale modes exhibit increased uncertainty. . . . 47

List of Tables

6.1 Architecture of the BNN-based emulator developed in this work. The network takes three reionization parameters ($M_{\rm min}, N_{\rm ion}, R_{\rm mfp}$) as input and predicts the 21-cm power spectrum across 7 k-bins [Mahida et al., 2025]. . 32
6.2 Three test parameter sets used for comparing BNN-based Gibbs sampling with MCMC. The corresponding neutral hydrogen fractions $x_{\rm H}$ indicate different stages of cosmic reionization: early (set 1), intermediate (set 2), and late (set 3). 34
6.3 Recovered parameter values for the three test parameter sets using MCMC and Gibbs sampling. 38
6.4 Acceptance ratios and number of sampling steps for MCMC and Gibbs sampling. 38

1 Cosmic Dawn and Epoch of Reionization (EoR)

1.1 Introduction

After a period of recombination and the ensuing "Dark Ages," the Universe entered two transformative epochs — the Cosmic Dawn (CD) and the Epoch of Reionization (EoR) — which mark the period when the first stars, galaxies, and black hole seeds formed, and when the intergalactic medium (IGM) transitioned from neutral to ionized [Barkana and Loeb, 2001, Loeb and Furlanetto, 2013].

The Cosmic Dawn refers to the era when the first generation of stars formed within small dark matter halos at redshifts $z \sim 30$–15, roughly 100–300 Myr after the Big Bang [Barkana and Loeb, 2001, Pritchard and Loeb, 2012]. These primordial stars were extremely massive, metal-free, and short-lived, producing intense ultraviolet (UV) and X-ray radiation that illuminated the early Universe. The emergence of the earliest galaxies and black hole seeds established the first astrophysical sources of ionizing photons.

As the Cosmic Dawn progressed, the Universe transitioned into the Epoch of Reionization, during which UV photons emitted by the first stars and galaxies gradually ionized the neutral hydrogen filling the IGM. Photons with energies above 13.6 eV were responsible for this transition [Loeb and Furlanetto, 2013]. While both stars and quasars contributed, dwarf galaxies were far more abundant at $z \gtrsim 6$, making stellar radiation the dominant driver of reionization [Fan et al., 2006].

Ionizing photons escaped their host galaxies and produced ionized bubbles in the surrounding neutral hydrogen. These bubbles expanded, merged, and eventually reionized the entire IGM by $z \sim 6$, marking the end of the EoR [Barkana and Loeb, 2001, Loeb and Furlanetto, 2013, Semelin and Iliev, 2015].

Together, the Cosmic Dawn and Epoch of Reionization represent the transformation of the early Universe into one where large-scale structures like galaxies began to form. Studying these epochs helps us understand when the first light sources emerged and how they reshaped cosmic evolution. The 21 cm signal provides a powerful probe of these eras, enabling direct observations of the thermal and ionization history of the early Universe [Pritchard and Loeb, 2012, Furlanetto et al., 2006, Koopmans et al., 2015].

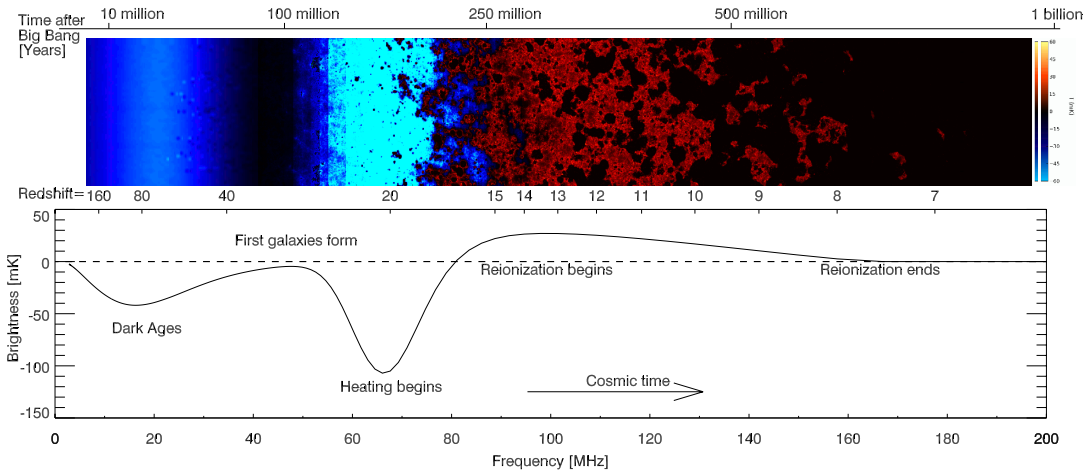


Figure 1.1: Evolution of the 21-cm brightness temperature from the pre-stellar era to the end of reionization, showing simulated fluctuations and the expected global signal shaped by early galaxy formation, heating, and ionization of the IGM. Adapted from [Pritchard and Loeb, 2010].

1.2 Key Observables: Lyman-α, CMB, and 21 cm Line

The intergalactic medium (IGM), which makes up most of the ordinary matter in the Universe, serves as a key record of cosmic evolution. During the Epoch of Reionization (EoR), the IGM transitioned from mostly neutral to highly ionized as ultraviolet radiation from the first stars and galaxies spread through space. Studying this transformation is essential for understanding when and how reionization occurred.

Neutral hydrogen plays a central role in tracing the IGM, as it was the dominant element in the early Universe. Once it became ionized during the EoR, it left behind observable signatures that remain detectable today. The most important among these are the Lyman-α forest, the anisotropies of the Cosmic Microwave Background (CMB), and the redshifted 21 cm signal.

1.2.1 The Lyman-α Forest

The Lyman-α forest is a series of absorption lines seen in the spectra of distant quasars and galaxies. These features arise when light from these sources passes through numerous clouds of neutral hydrogen along the line of sight. Each hydrogen cloud selectively absorbs light at the Lyman-α wavelength, imprinting distinct absorption features.

Together, these absorption lines form the Lyman-α forest, with each line corresponding to a different hydrogen cloud at a different redshift. By studying the distribution and depth of these lines, one can trace the evolution of the neutral fraction of the IGM. This makes the Lyman-α forest a powerful probe of the EoR and the thermal history of the early Universe [McQuinn, 2016].

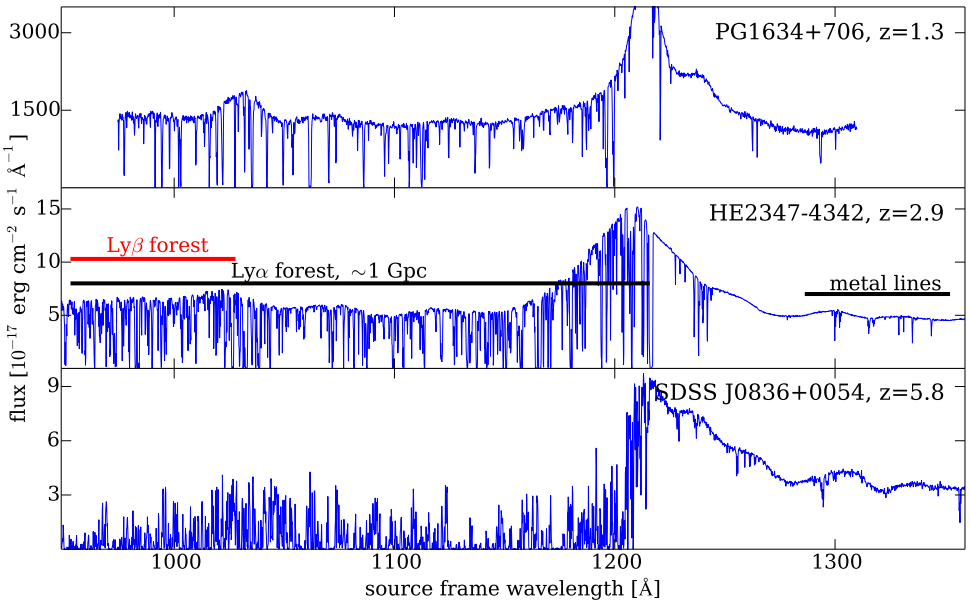


Figure 1.2: Lyman-α forest spectral region for three quasars spanning a wide range of redshifts. Each spectrum shows numerous absorption features produced by intervening neutral hydrogen clouds along the line of sight, illustrating how the density of Lyman-α absorption evolves with redshift [McQuinn, 2016].

1.2.2 CMB Anisotropies

After recombination, photons decoupled from matter and formed the Cosmic Microwave Background (CMB). During the Epoch of Reionization, free electrons produced by the reionized gas scattered some of these photons through Thomson scattering. This interaction altered the anisotropy and polarization patterns of the CMB.

These scattering-induced signatures provide key constraints on the timing and duration of reionization. In particular, measurements of the optical depth to Thomson scattering by the *Planck* satellite offer one of the most precise global constraints on when reionization occurred [Planck Collaboration, 2020].

1.2.3 21 cm Signal

In the early Universe, neutral hydrogen atoms emitted radiation at a wavelength of 21 cm via the hyperfine spin-flip transition. As the Universe expanded, this radiation became redshifted, allowing the 21 cm line to map hydrogen across a wide range of cosmic epochs.

The intensity of the 21 cm signal depends on the temperature, density, and ionization state of the hydrogen gas. As the first stars and galaxies formed, their radiation heated and ionized the surrounding medium, imprinting spatial fluctuations in the 21 cm brightness temperature.

By observing these fluctuations, astronomers can trace the thermal and ionization history of the Universe during the Cosmic Dawn and the EoR. Current and upcoming radio telescopes aim to detect this faint but highly informative signal, offering a powerful

probe of early structure formation [Furlanetto, 2016, Pritchard and Loeb, 2012, Mesinger et al., 2015].

2 The 21 cm Signal

2.1 Observable

The observable quantity of the 21 cm signal is the differential brightness temperature. The differential brightness temperature, denoted as δT_b, is the difference between the brightness temperature of the 21 cm radiation from neutral hydrogen and the background temperature of the Cosmic Microwave Background (CMB). It represents the observed intensity of the 21 cm line relative to the CMB at a given redshift.

It is expressed as :

$$\delta T_b(\mathbf{x}, z) \approx 27\, x_{\rm HI}(1+\delta_b)\left(1-\frac{T_{\rm CMB}(z)}{T_S}\right)\left(\frac{1+z}{10}\frac{0.15}{\Omega_m h^2}\right)^{1/2}\left(\frac{\Omega_b h^2}{0.023}\right)\ \text{mK}, \tag{2.1}$$

where $x_{\rm HI}$ is the neutral hydrogen fraction, δ_b is the baryon overdensity, and Ω_m and Ω_b are the total matter and baryon density parameters, respectively.[Pritchard and Loeb, 2012, Furlanetto et al., 2006, Mesinger et al., 2015]

The evolution of δT_b traces distinct phases of the early Universe. When $T_S < T_{\rm CMB}$, the 21 cm signal appears in absorption, as during the Cosmic Dawn the gas cooled faster than the background radiation field. As the first stars and galaxies formed, the intergalactic medium (IGM) was ionized and heated, driving $T_S > T_{\rm CMB}$ and making the signal appear in emission. Eventually, as reionization progresses, $x_{\rm HI}$ decreases and the signal fades.

Spatial fluctuations in δT_b arise due to variations in density, temperature, and ionization fraction across different regions of the IGM. These fluctuations record how matter and radiation interacted during the early stages of structure formation. Some regions collapsed first, forming early stars and galaxies that emitted UV and X-ray radiation, heating and ionizing the surrounding neutral hydrogen while distant regions remained neutral and cold. This led to spatial variations in the spin temperature and hence in the brightness contrast. Mapping δT_b reveals how early cosmic structures influenced their environments during the Cosmic Dawn and the Epoch of Reionization (EoR).

From these fluctuations, several statistical observables can be derived. The power spectrum quantifies the variance of brightness temperature fluctuations over different spatial scales, providing a primary tool for connecting simulations to observations. Higher-order statistics such as the bispectrum offer additional insight into the non-Gaussian nature of reionization. Together, these observables provide a statistical framework to interpret the 21 cm signal [Furlanetto et al., 2006].

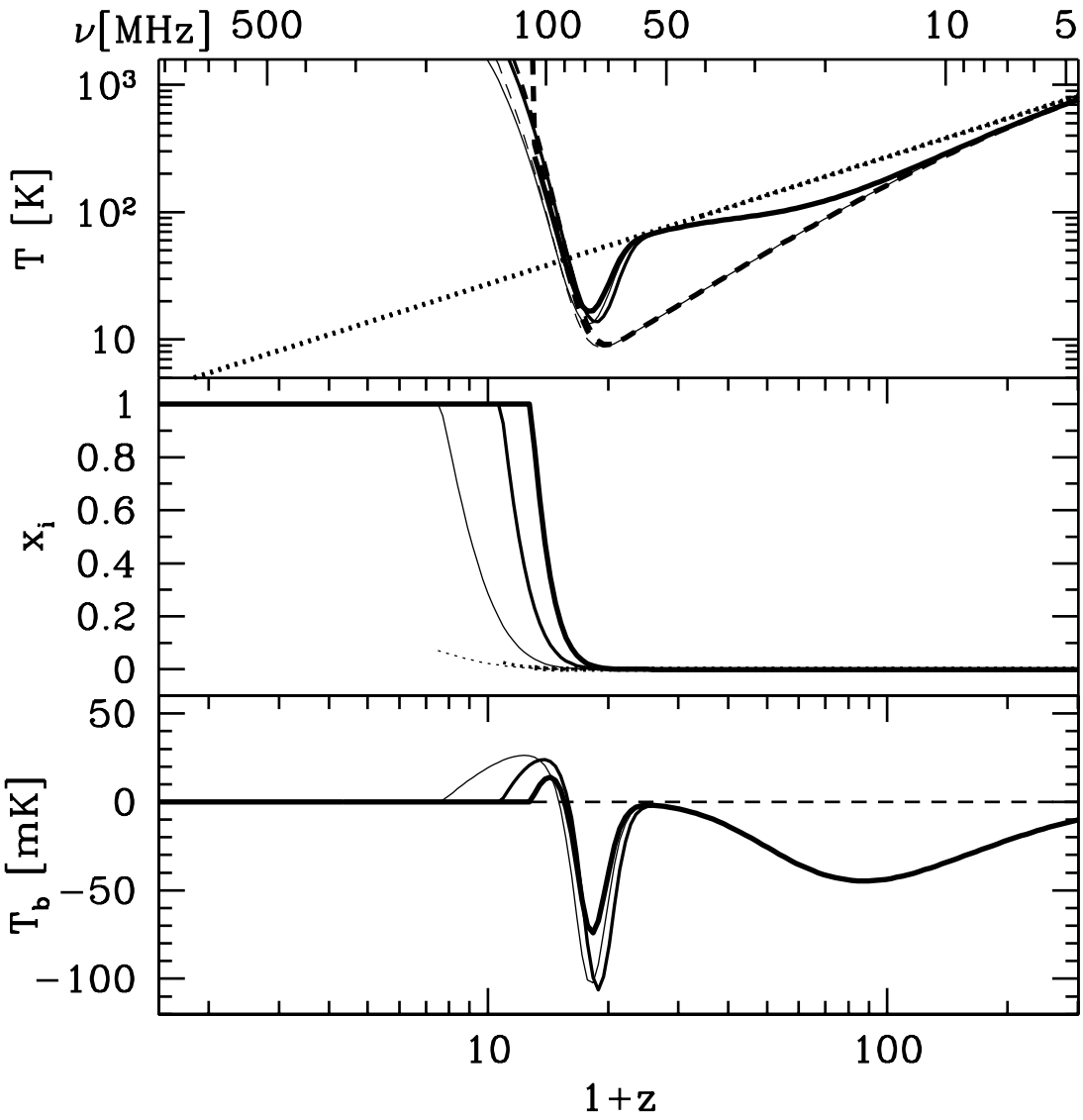


Figure 2.1: Global 21-cm signal during Cosmic Dawn and the Epoch of Reionization. The figure shows (i) the CMB temperature $T_{\rm CMB}$, gas kinetic temperature T_K, and spin temperature T_S (top), (ii) ionized fractions in galaxies x_i and diffuse regions x_e (middle), and (iii) the mean 21-cm brightness temperature δT_b for different astrophysical models (bottom) [Pritchard and Loeb, 2012].

2.2 Challenges: Foregrounds and Systematics

Detecting the 21 cm signal is a challenging prospect due to contamination from much stronger astrophysical foregrounds and radio frequency interference (RFI). These foregrounds arise mainly from Galactic synchrotron radiation, free-free emission, and extragalactic radio sources like galaxies and quasars. These foregrounds are typically 10^4 to 10^5 times stronger than the expected signal, making accurate subtraction and modelling essential [Barry et al., 2019, Liu and Tegmark, 2011, Wolz et al., 2015].

While these foregrounds vary smoothly with frequency, small calibration errors or systematic effects like RFI can cause the bright foreground to leak into the cosmological window. This leads to a contamination region in the two-dimensional power spectrum

known as the "foreground wedge", which limits the scales that can be used for analysis [Liu and Tegmark, 2011].

In addition, instrumental effects such as beam chromaticity, polarization leakage, ionospheric distortions and gain errors complicate measurements. Thermal noise and RFI at low frequencies add to the challenge. Overcoming these issues requires accurate calibration, careful instrument design, and advanced analysis techniques like foreground removal and statistical recovery [Barry et al., 2019].

2.3 Statistical Quantification through Power Spectrum

The 21 cm brightness temperature fluctuations, $\delta T_b(\mathbf{x}, z)$, encode the spatial and temporal evolution of the IGM during the Cosmic Dawn and the Epoch of Reionization (EoR). To analyze the statistical behaviour of these fluctuations, higher-order statistics such as the power spectrum and bispectrum are employed .

The 21 cm power spectrum quantifies the variance of the brightness temperature as a function of spatial scale [Mesinger et al., 2015]. In Fourier space, the 3D power spectrum of the 21 cm brightness temperature fluctuations is defined by

$$\left\langle \delta\tilde{T}_b(\mathbf{k})\, \delta\tilde{T}_b^*(\mathbf{k}') \right\rangle = (2\pi)^3\, \delta_D(\mathbf{k}-\mathbf{k}')\, P_{21}(k,z), \tag{2.2}$$

where $\delta\tilde{T}_b(\mathbf{k})$ is the Fourier transform of the brightness temperature field $\delta T_b(\mathbf{x})$, δ_D denotes the Dirac delta, and $P_{21}(k,z)$ is the 3D power spectrum at wavenumber $k = |\mathbf{k}|$ and redshift z [Mellema et al., 2013].

The dimensionless power spectrum is often used,

$$\Delta^2_{21}(k) \;=\; \frac{k^3}{2\pi^2}\, P_{21}(k,z), \tag{2.3}$$

which gives the contribution to the variance per logarithmic interval in k [Mellema et al., 2013].

Physically, $P_{21}(k)$ provides a measure of the amplitude of 21 cm fluctuations at a given scale k [Mellema et al., 2013]. During the early stages of reionization, the signal is nearly Gaussian and thus the power spectrum alone captures most of the information. As reionization progresses, the appearance of ionized bubbles introduces strong non-linearities and correlations between different regions, making it increasingly non-Gaussian. The power spectrum, being a two-point statistic, measures only the variance of fluctuations and assumes random, uncorrelated phases among Fourier modes. But once the field becomes non-Gaussian, these assumptions break down. Different ionization topologies can yield identical power spectra despite different spatial distributions of ionized and neutral regions. Thus, the power spectrum alone becomes insufficient for a complete statistical description of the 21 cm signal [Majumdar et al., 2018, Mesinger et al., 2015].

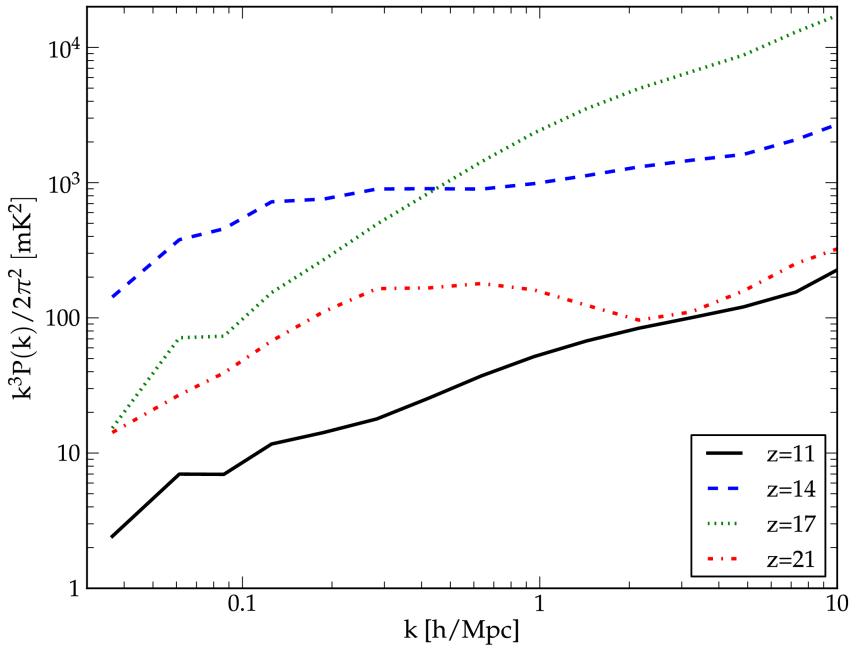


Figure 2.2: The power spectrum of the 21-cm signal at high redshifts, including all fluctuations in the spin temperature. Reproduced from [Santos et al., 2010].

To capture these non-Gaussian features, one may make use of the bispectrum. The bispectrum is the Fourier equivalent of a three-point correlation function, which measures the correlation between triplets of Fourier modes. The bispectrum encodes information about the phase relation between different modes, making it particularly sensitive to non-Gaussian features in the 21 cm brightness temperature field.

2.4 From Observables to Parameter Inference

The statistical quantities discussed above, particularly the 21-cm power spectrum, provide a compact summary of the fluctuations in the brightness temperature field. However, the ultimate goal of 21-cm cosmology is not merely to characterize these fluctuations, but to infer the underlying astrophysical parameters that govern the process of reionization, such as $M_{\rm min}$, $N_{\rm ion}$, and $R_{\rm mfp}$ [Pritchard and Loeb, 2012].

This inference problem is inherently challenging. The mapping between astrophysical parameters and the observed power spectrum is highly non-linear and often degenerate, with multiple parameter combinations producing similar statistical signatures [Furlanetto et al., 2006, Greig and Mesinger, 2017, Iliev et al., 2015]. Additionally, observational uncertainties, foreground contamination, and instrumental effects further complicate the extraction of reliable constraints [Liu and Tegmark, 2011].

To address these challenges, a probabilistic framework is required—one that can incorporate prior knowledge, handle uncertainties in both the data and the model, and provide a full characterization of parameter degeneracies. Bayesian inference offers a natural and powerful approach for this purpose [Gelman et al., 2013].

In the following chapter, we introduce the Bayesian framework and discuss a range of sampling techniques that enable efficient exploration of complex posterior distributions arising in 21-cm parameter inference.

3 Bayesian Inference and Sampling Techniques for EoR 21-cm Observations

3.1 Bayesian Framework for EoR Parameter Inference

In the previous chapter, we discussed how the 21-cm signal can be statistically characterized using quantities such as the power spectrum. While these observables provide a compact description of the signal, the central goal of this work is to infer the underlying Epoch of Reionnization(EoR) parameters that govern the reionization process.

This inverse problem is inherently complex. The relationship between the model parameters and the observed signal is highly non-linear, often degenerate, and subject to uncertainties arising from both observations and modelling. As a result, a deterministic approach to parameter estimation is insufficient, and a probabilistic framework is required to fully characterize the range of parameter values consistent with the data.

Bayesian inference provides a natural framework for this task. It enables the incorporation of prior knowledge, systematic treatment of uncertainties, and direct estimation of posterior probability distributions for the parameters of interest. However, in most realistic scenarios, the posterior distribution cannot be evaluated analytically due to the high dimensionality and complexity of the model.

To address this, sampling-based techniques such as Markov Chain Monte Carlo (MCMC) are employed to explore the posterior distribution efficiently. In this chapter, we introduce the fundamentals of Bayesian inference and discuss a range of sampling techniques, with particular emphasis on methods suited to the 21-cm parameter inference problem.

3.2 Fundamentals of Bayesian Inference

In Bayesian Statistics, probability is defined as the degree of belief or confidence in a particular hypothesis rather than a long-run frequency of occurrence. This belief is continuously updated based on the new available information and the prior knowledge. This framework acknowledges and quantifies uncertainty arising from incomplete data or the inherent randomness in observations. This is particularly useful in fields like astronomy where conducting multiple trials is not feasible due to singular or unique events. Thus, Bayesian statistics allow a more flexible approach to understanding uncertainty in scenarios where traditional frequentist methods fall short.

The foundation of Bayesian inference is the Bayes Theory. It mathematically expresses how prior beliefs are updated in light of new evidence:

$$p(\theta|D) = \frac{p(D|\theta)\, p(\theta)}{p(D)} \tag{3.1}$$

where:

- $p(\theta|D)$ is the **posterior probability** of the parameter θ given the observed data D,
- $p(D|\theta)$ is the **likelihood**, representing the probability of observing the data under the model defined by θ,
- $p(\theta)$ is the **prior**, encapsulating our knowledge about the parameter before seeing the data, and
- $p(D)$ is the **evidence**, a normalization constant defined as

$$p(D) = \int p(D|\theta)\, p(\theta)\, d\theta \tag{3.2}$$

If the dataset consists of independent measurements x_i, the likelihood function becomes

$$L(D|\theta) = \prod_{i=1}^{n} p(x_i|\theta) \tag{3.3}$$

In many practical problems, especially those focused on parameter estimation, the evidence term $p(D)$ can be treated as a constant since it does not depend on θ [Gelman et al., 2013]. Consequently, the posterior distribution can be expressed as being proportional to the product of the likelihood and the prior:

$$p(\theta|D) \propto p(D|\theta)\, p(\theta) \tag{3.4}$$

where $P(\theta \mid D)$ is the posterior distribution of parameters θ given the data D, $P(D \mid \theta)$ is the likelihood, $P(\theta)$ is the prior, and $P(D)$ is the evidence (or marginal likelihood) that ensures normalization [Gelman et al., 2013].

In practical applications such as cosmology and astrophysics, the evidence term P(D) is often computationally intractable due to high dimensional integrals. Therefore, approximate methods such as Markov Chain Monte Carlo (MCMC) are employed to sample from the posterior distribution without directly evaluating the evidence[Foreman-Mackey et al., 2013]. This allows one to estimate parameter uncertainties, correlations and credible ranges effectively. A crucial aspect for analyzing complex, multi parameter models such as those describing the 21 cm signal.

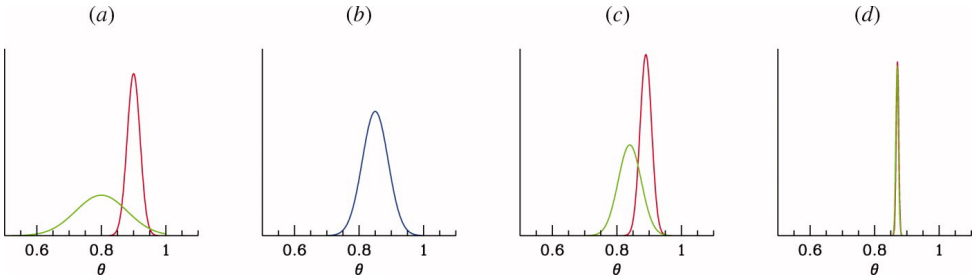


Figure 3.1: Converging views in Bayesian inference. Two scientists with different prior beliefs, $p(\theta|I_i)$ (panel a, red and green PDFs), observe a single datum with likelihood $L(\theta)$ (panel b). Their posterior distributions $p(\theta|m_1)$ (panel c), obtained via Bayes' theorem , represent updated knowledge of the parameter θ. After observing many data points, the posteriors become essentially indistinguishable, illustrating convergence of Bayesian inference. Adapted from Trotta [Trotta, 2008].

3.3 Different Sampling techniques

Bayesian inference in modern cosmology and machine learning frequently require sampling from posterior distribution that have no closed form solution. Markov Chain Monte Carlo(MCMC) methods provide a systematic way to explore such distributions by constructing a Markov Chain whose stationary distribution is the target posterior. Because posteriors arising in astrophysical analysis can be highly non-Gaussian, curved or multimodal, the efficiency of the MCMC method becomes critical. Each sampler uses a different strategy to propose new states and evaluate their plausibility, with the choice of method impacting covergence, mixing and computational cost. The following subsections discuss several basic and advanced sampling strategies used in contemporary Bayesian analysis.

3.3.1 Metropolis-Hastings Algorithm

The Metropolis-Hastings (MH) algorithm is a foundational method in Markov Chain Monte Carlo (MCMC), designed to draw samples from complex posterior distributions that cannot be sampled analytically [Metropolis et al., 1953, Hastings, 1970]. Its central idea is to explore the parameter space by proposing a new candidate state and deciding whether to accept it based on an acceptance probability:

$$\alpha = \min\left(1, \frac{P(D \mid \theta')P(\theta')q(\theta \mid \theta')}{P(D \mid \theta)P(\theta)q(\theta' \mid \theta)}\right), \tag{3.5}$$

where:

- $P(D \mid \theta)$ is the likelihood of the data D given the current parameter θ,
- $P(\theta)$ is the prior distribution of the parameter θ,
- $q(\theta' \mid \theta)$ is the proposal distribution for proposing a new state θ' from the current state θ.

Algorithm:

1. Choose an initial point $\theta^{(0)}$.
2. At step t, propose a candidate parameter $\theta' \sim q(\theta' \mid \theta^{(t)})$.
3. Compute the acceptance ratio

$$r = \frac{P(D \mid \theta')P(\theta')q(\theta^{(t)} \mid \theta')}{P(D \mid \theta^{(t)})P(\theta^{(t)})q(\theta' \mid \theta^{(t)})}.$$

4. Accept or reject the candidate:
   - Generate a uniform random number $u \in [0, 1]$.
   - If $u \leq r$, accept the candidate: $\theta^{(t+1)} = \theta'$.
   - If $u > r$, reject the candidate: $\theta^{(t+1)} = \theta^{(t)}$.
5. Repeat until a sufficient number of samples is obtained.

The MH algorithm is valued for its simplicity, as it only requires the posterior up to a constant normalization. Its main drawback is the random-walk behavior, which can lead to slow mixing in high-dimensional, strongly correlated, or ill-scaled parameter spaces. Efficient tuning of the proposal distribution is often necessary, making the method less automatic and more problem dependent [Metropolis et al., 1953, Hastings, 1970].

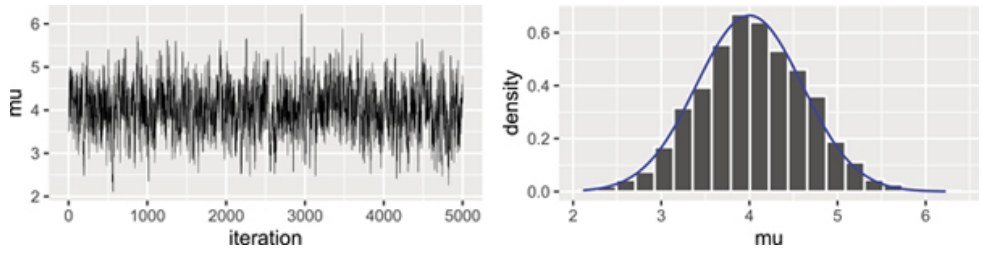


Figure 3.2: A trace plot (left) and histogram (right) of 5000 Metropolis chain values of μ. The target posterior PDF is superimposed in blue.[Johnson et al., 2022]

3.3.2 Hamiltonian Monte Carlo

Hamiltonian Monte Carlo(HMC) was developed to address the inefficiency of random-walk exploration in MH. By incorporating ideas from classical mechanics, HMC uses gradients of the log-posterior to propose distant,highly probable states while maintaining high acceptance probabilty. This dramatically reduces the random walk behavior of MH and improves mixing in high dimensional problems [Neal, 2011, Betancourt, 2017].

Algorithm:

1. Introduce momentum variables

$$p \sim \mathcal{N}(0, M).$$

2. Construct the Hamiltonian

$$H(\theta, p) = -\log P(\theta \mid D) + \frac{1}{2} p^T M^{-1} p.$$

3. Simulate Hamiltonian dynamics using leapfrog integration:

   (a) Update momentum by a half step:

   $$p \leftarrow p - \frac{\varepsilon}{2} \nabla_\theta \log P(\theta \mid D).$$

   (b) Update position (parameters) by a full step:

   $$\theta \leftarrow \theta + \varepsilon \, M^{-1} p.$$

   (c) Update momentum by another half step:

   $$p \leftarrow p - \frac{\varepsilon}{2} \nabla_\theta \log P(\theta \mid D).$$

4. Use Metropolis acceptance based on the Hamiltonian difference:

   $$\alpha = \min\bigl(1, \exp\bigl(-H(\theta', p') + H(\theta, p)\bigr)\bigr).$$

5. Repeat the process to obtain samples.

HMC achieves dramatically improved sampling efficiency and effective sampling size relative to to MH, particularly in high dimensional differentiable models [Neal, 2011, Betancourt, 2017]. However it relies on gradient information, requiring posterior to be continuous and the mechanism to calculate derivatives efficiently and quickly. Careful tuning of integration step size and trajectory length is essential and the method can struggle with multimodal or discontinuous posteriors.

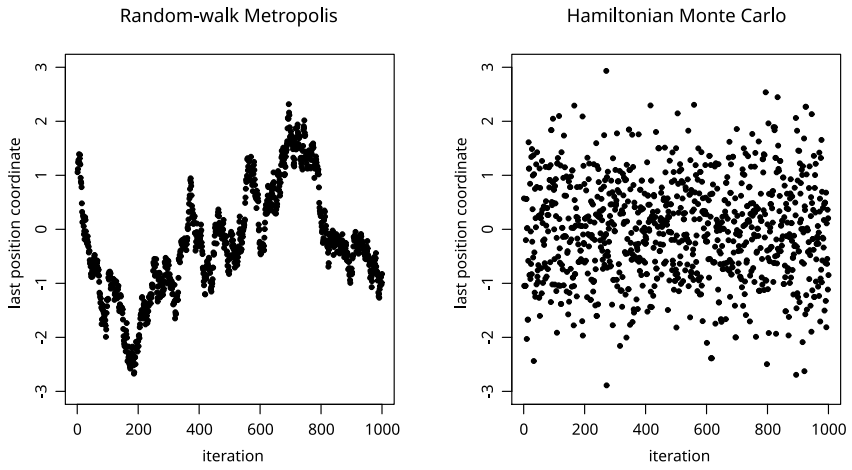


Figure 3.3: Comparison of sampling efficiency for a 100-dimensional Gaussian. Shown are the sampled values of the variable with the largest standard deviation using a random-walk Metropolis run (left) and an HMC run with $L = 150$ (right). HMC demonstrates reduced correlation and faster exploration of the parameter space, illustrating its superior mixing compared to Metropolis [Neal, 2011].

3.3.3 Gibbs Sampling

Gibbs sampling simplifies sampling in multivariate systems by updating one parameter at a time while conditioning on all others. When conditional distributions are available, Gibbs sampling avoids rejection steps entirely and can provide rapid convergence for hierarchical or conjugate Bayesian models [Gelman et al., 2013].

Algorithm

1. Start with an initial vector

$$\theta^{(0)} = (\theta_1, \theta_2, \ldots, \theta_d).$$

2. For iteration t, update each component sequentially:

   (a) Sample the first component:

$$\theta_1^{(t+1)} \sim P\Big(\theta_1 \mid \theta_2^{(t)}, \ldots, \theta_d^{(t)}, D\Big).$$

   (b) Sample the second component:

$$\theta_2^{(t+1)} \sim P\Big(\theta_2 \mid \theta_1^{(t+1)}, \theta_3^{(t)}, \ldots, D\Big).$$

   (c) Continue this process for all components:

$$\theta_i^{(t+1)} \sim P\Big(\theta_i \mid \theta_1^{(t+1)}, \ldots, \theta_{i-1}^{(t+1)}, \theta_{i+1}^{(t)}, \ldots, \theta_d^{(t)}, D\Big),$$

   until all d components have been updated.

3. Repeat the full update cycle for the desired number of iterations to generate Gibbs samples.

Gibbs sampling is extremely efficient when full conditionals are available, enabling high high acceptance and often strong mixing[Gelman et al., 2013]. In cosmology, it has been crucial for CMB map reconstruction and hierarchical modeling. Its limitation is structural, if even one conditional posterior lacks a closed form or is expensive to sample, pure Gibbs sampling becomes infeasible. Strong parameter correlation may also cause slow movement along narrow posterior ridges.

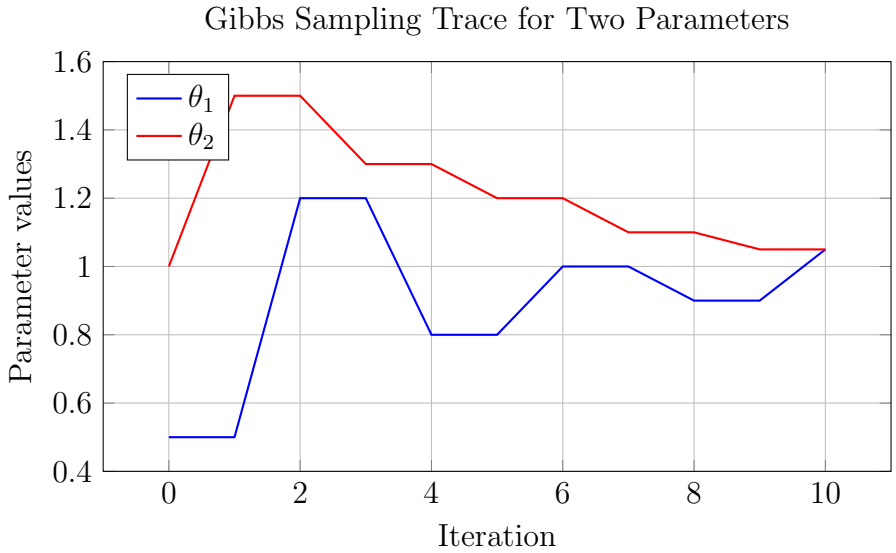


Figure 3.4: Trace of Gibbs sampling for two correlated parameters. The stair-step pattern illustrates sequential updates of θ_1 and θ_2.

3.3.3.1 Metropolis-within-Gibbs

Metropolis-within-Gibbs extends Gibbs sampling to more complex models in which only some conditional posteriors are analytically tractable. It preserves the efficient component-wise updating structure while enabling approximate updates for difficult parameters [Gilks et al., 1995].

Algorithm:

1. For each parameter θ_i:

   (a) If the full conditional distribution

   $$P(\theta_i \mid \theta_{-i}, D)$$

   is available analytically, then sample directly:

   $$\theta_i^{(t+1)} \sim P\Big(\theta_i \mid \theta_{-i}^{(t)}, D\Big).$$

   (b) Else (no closed-form conditional):

   i. Propose a candidate

   $$\theta_i' \sim q(\theta_i' \mid \theta_i^{(t)}).$$

   ii. Compute the Metropolis–Hastings acceptance ratio restricted to the conditional posterior:

   $$r = \frac{P(\theta_i' \mid \theta_{-i}, D)\, q(\theta_i^{(t)} \mid \theta_i')}{P(\theta_i^{(t)} \mid \theta_{-i}, D)\, q(\theta_i' \mid \theta_i^{(t)})}.$$

   iii. Accept or reject:

   $$\theta_i^{(t+1)} = \begin{cases} \theta_i', & \text{with probability } \min(1, r), \\ \theta_i^{(t)}, & \text{otherwise.} \end{cases}$$

2. Move to the next parameter and continue cycling through all coordinates.

3. Repeat the cycle until convergence.

This hybrid approach maintains the modular structure of Gibbs sampling while adding the flexibility of Metropolis-Hastings algorithm for complex calculations. Its performance depends on the tuning of the embedded MH step, and correlated parameters can still lead to slow mixing unless parameters are blocked or reparameterized.

3.3.4 Diffusive Gibbs Sampling

Diffusive Gibbs Sampling modifies standard Gibbs sampling by injecting controlled randomness into each update. This extra diffusion helps the sampler move through narrow, highly correlated corridors of posterior where classical Gibbs sampler may become trapped [Chen et al., 2024].

Algorithm:

1. **Initialize particles** Draw N particles $\{\theta_i\}_{i=1}^N$ from the prior distribution $\pi(\theta)$.
2. **Assign likelihood levels** Set the initial likelihood level $L_0 = -\infty$. A sequence of levels $\{L_k\}$ will be created adaptively during sampling.
3. **Initialize level indices** For each particle i, set its current likelihood-level index to
$$k_i = 0.$$
4. **Define diffusive mixture distribution** Particles explore a mixture of constrained distributions:
$$p_k(\theta) \propto \pi(\theta)\, \mathbb{I}[L(\theta) > L_k]\,,$$
and may diffuse up and down the hierarchy of levels.
5. **Iterate MCMC exploration for each particle**
   (a) Select a particle at random.
   (b) Propose a new state θ' using a symmetric proposal $q(\theta' \mid \theta)$.
   (c) Accept θ' if and only if
   $$L(\theta') > L_{k_i},$$
   otherwise reject.
   (d) With some probability, propose a move to a different level (either $k_i + 1$ or $k_i - 1$) and accept based on the relative level weights.
6. **Create new likelihood levels adaptively** As particles explore higher likelihood regions, define a new level L_{k+1} such that approximately an $\exp(-1)$ fraction of samples at level k satisfy
$$L(\theta) > L_{k+1}.$$
Append levels sequentially as needed.
7. **Continue diffusive exploration** Particles repeatedly diffuse:
   - *horizontally*: MCMC moves in parameter space θ,
   - *vertically*: moves across likelihood levels.

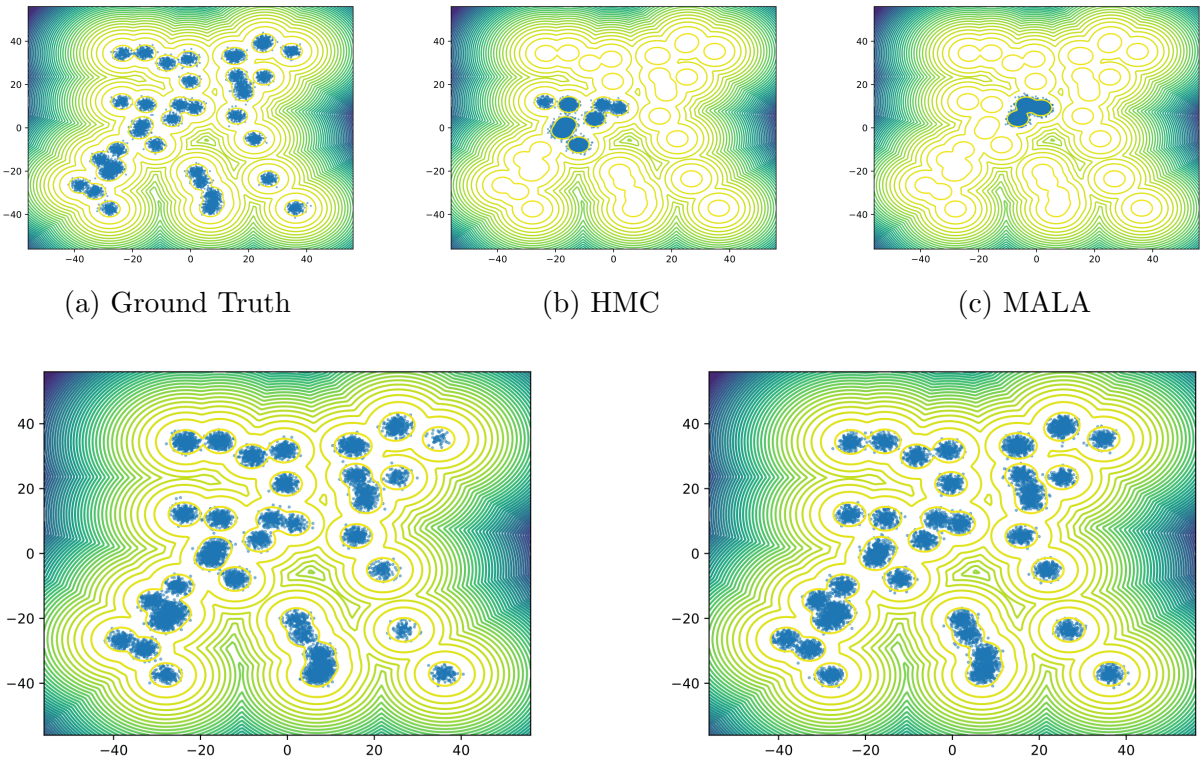
   This enables efficient traversal of the nested likelihood contours.
8. **Estimate compression factors** For each level transition $k \to k+1$, estimate the compression factor
$$t_k = \frac{X_{k+1}}{X_k},$$
where X_k is the surviving prior mass above level L_k.
9. **Assign weights to saved particle states** As particles visit each level, store their states θ and assign weights proportional to the prior-mass difference:
$$w_k \propto X_k - X_{k+1}.$$
10. **Stopping criterion** Terminate when the remaining evidence contribution from unexplored levels becomes negligible.

11. **Compute the Bayesian evidence** Estimate the marginal likelihood using

$$Z \approx \sum_i w_i \, L(\theta_i).$$

12. **Construct posterior samples** Normalize the weights and resample to obtain draws from the posterior distribution $p(\theta \mid D)$.

Diffusive Gibbs sampling improves mixing in models with strong parameter correlations or narrow conditional distributions. However the perturbation makes it an approximate method unless paired with metropolis corrections. Tuning the magnitude of diffusion can also be model dependent [Chen et al., 2024].



(a) Ground Truth (b) HMC (c) MALA

(d) PT (e) DiGS

Figure 3.5: Visualization of 10^4 samples for the 2D MoG-40 synthetic problem [Chen et al., 2024] generated by each sampler. All samplers are initialized at the origin. HMC and MALA fail to explore distant modes, PT covers all modes but with uneven sample distribution, and DiGS successfully captures all modes with balanced samples.

3.4 Why Gibbs Sampling?

Gibbs sampling, implemented through a Metropolis-within-Gibbs (MWG) scheme, was chosen for this project because it is well suited to the structure of the 21 cm inference problem and the Bayesian Neural Network (BNN) emulator used to model the power spectrum. The astrophysical parameters $M_{\min}$, N_{ion}, and R_{mfp} are strongly correlated, meaning that a full three-dimensional Metropolis–Hastings update would mix slowly due to random-walk behavior [Mesinger et al., 2015, Iliev et al., 2015]. Gibbs sampling avoids

this by updating one parameter at a time while keeping the others fixed, allowing the chain to move efficiently along parameter degeneracies.

Gradient-based samplers such as Hamiltonian Monte Carlo cannot be used because the likelihood depends on BNN predictions, which are non-differentiable and do not permit analytical gradients. Pure Gibbs sampling is also not feasible because the conditional posteriors do not have closed-form expressions. For this reason, the MWG scheme is adopted: each parameter is updated with a one-dimensional Metropolis step using the conditional posterior, while the remaining parameters remain fixed.

This greatly reduces tuning difficulty, increases acceptance rates, and lowers computational cost by minimizing expensive BNN evaluations. Furthermore, Gibbs sampling provides a flexible and modular framework that can naturally be extended to more complex inference tasks. In particular, future work aims to incorporate realistic foregrounds, instrumental effects, and gaps in the frequency spectrum into the inference pipeline. Techniques such as Gaussian Constrained Realisations(GCR) and Gibbs-based joint inference can handle gappy data and uncertain foregrounds within a fully Bayesian framework. The choice of Gibbs sampling in the the current study therefore not only ensures efficient and robust parameter recovery for the present test cases but also establishes a foundation for these more advanced, joint inference scenarios.

Overall, the resulting sampler is robust, efficient, and well suited for highly related cosmological parameter inference, while being readily extendable to tackle realistic observational challenges in future work.

Extension of the Inference Framework

The choice of a Metropolis-within-Gibbs sampling strategy, as motivated in the previous section, provides an efficient and flexible approach for exploring the posterior distribution of the astrophysical parameters. Its component-wise update structure is particularly well suited to problems involving correlated parameters and complex likelihood evaluations.

The formulation presented thus far assumes that the observed 21-cm signal can be directly compared with model predictions. However, in realistic settings, the data entering the inference pipeline is often incomplete due to foreground contamination and the associated masking of regions in Fourier space. This modifies the structure of the likelihood and necessitates an extension of the inference framework.

In the following chapter, we extend the inference framework to account for foreground contamination within the statistical model. We then introduce a reconstruction-based extension using Gaussian Constrained Realizations, which enables the recovery of missing modes and allows inference to be performed on incomplete data.

These developments extend the core Gibbs-based inference methodology into a more general framework applicable to realistic 21-cm observations.

4 Inference Strategies for EoR 21-cm Signal Under Foreground Contamination

4.1 Introduction

The redshifted 21-cm from neutral hydrogen provides a unique probe of the Cosmic Dawn and the Epoch of Reionization. However, its detection and interpretation are severely challenged by the presence of foregrounds, which are several orders of magnitude brighter than the cosmological signal. These foregrounds, dominated by Galactic synchrotron emission and extragalactic radio sources, exceed the expected 21-cm signal by factors of 10^4–10^5, making accurate inference difficult.

Foreground contamination does not merely introduce noise, but fundamentally alters the structure of the observed data. Due to instrumental chromaticity and calibration imperfections, spectrally smooth foregrounds leak into higher Fourier modes, producing a contaminated region in cylindrical k-space known as the foreground wedge [Liu and Tegmark, 2011]. As a result, a significant fraction of the Fourier modes becomes unreliable or inaccessible for cosmological analysis.

Therefore, parameter inference from 21-cm observations must be treated as a problem of inference with incomplete and contaminated data. Rather than being solely a foreground removal problem, it is fundamentally a statistical inference problem in which part of the information is missing or corrupted. In this chapter, we review the major classes of approaches developed to address this challenge, highlighting their assumptions and limitations, and motivating the reconstruction-based framework discussed in the next chapter.

4.2 Inference via Foreground Avoidance

Foreground avoidance is one of the most widely used strategies used in early 21-cm analyses. This approach exploits the fact that foregrounds are spectrally smooth and primarily contaminate a well-defined region in the Fourier space- the foreground wedge- while leaving a relatively clean region known as the EoR window [Wolz et al., 2015].

In avoidance-based methods, all modes within the contaminated wedge are excluded from analysis, and parameter inference is performed using only the remaining contaminated modes [Liu and Tegmark, 2011]. This approach has the advantage of being simple, robust, and largely independent of detailed foreground modelling.

However, the primary limitation of foreground avoidance is the substantial loss of information. By discarding a large fraction of the Fourier modes, particularly at low $k_{||}$, the sensitivity to large-scale features of the signal is significantly reduced. This leads to broader posterior distributions and weaker constraints on parameters. Additionally, the anisotropic removal of modes complicates the interpretation of the recovered power spectrum.

Thus, while avoidance provides a reliable baseline, it is inherently limited by its inability to exploit the available data fully.

4.3 Inference through Foreground Subtraction

An alternative class of methods seeks to model and subtract foreground contamination from the observed data. These approaches exploit the spectral smoothness of foregrounds to separate them from the cosmological signal, which contains additional small-scale structure along the line-of-sight direction, corresponding to higher $k_{||}$ modes.

Common techniques include polynomial fitting, principal component analysis (PCA), independent component analysis (ICA), and Gaussian process regression [Santos et al., 2005, Chapman et al., 2013, Wolz et al., 2015]. These methods attempt to identify the dominant smooth components associated with foregrounds and remove them from the data.

Although foreground removal methods can, in principle, recover more Fourier modes than avoidance, they suffer from some important limitations. Imperfect modelling can lead to residual foreground contamination, which biases the inferred parameters. Conversely, aggressively fitting can remove part of the cosmological signal itself, resulting in signal loss. Furthermore, these methods are highly sensitive to instrumental systematics, such as calibration errors and beam chromaticity, which can introduce non-smooth features that are difficult to distinguish from the signal.

As a result, foreground removal involves a trade-off between minimizing residual contamination and preserving the cosmological signal, making it challenging to achieve optimal performance [Wolz et al., 2015].

4.4 Foreground Modelling and Joint Inference

A more statistically consistent approach is to model the observed data as a combination of cosmological signal, foregrounds, and noise, and to perform joint inference over all components. In this framework, the data model can be written as:

$$D = S(\theta) + F(\phi) + N, \tag{4.1}$$

where $S(\theta)$ represents the cosmological signal parameterized by θ, $F(\phi)$ denotes the foreground contribution with parameters ϕ, and N represents noise.

The goal is then to infer the joint posterior distribution:

$$P(\theta, \phi \mid D), \tag{4.2}$$

which allows uncertainties in foreground modelling to be propagated into the cosmological parameter estimates [Jacobs et al., 2016]

This approach has the advantage of being statistically rigorous and capable of incorporating realistic instrument models. However, it is computationally demanding, as it involves a high-dimensional parameter space and repeated evaluations of complex forward models. In addition, degeneracies between signal and foreground parameters can lead to slow convergence and difficulties in sampling.

4.5 Reconstruction-Based Approaches and Missing Data Inference

An alternate perspective is to interpret foreground contamination as a problem of incomplete or missing data in Fourier space. Instead of discarding contaminated modes or attempting to subtract foregrounds, one may treat these modes as latent variables and attempt to reconstruct them within the inference process.

In this framework, the goal is to infer the joint posterior distribution:

$$P(\theta, D_{\text{missing}} \mid D_{\text{observed}}), \tag{4.3}$$

where D_{observed} denotes the uncontaminated modes and D_{missing} represents the contaminated or inaccessible modes.

This viewpoint is closely related to reconstruction techniques developed in other areas of cosmology, particularly in the context of Cosmic Microwave Background (CMB) analysis and large-scale structure inference. Methods such as Gaussian Constrained Realizations (GCR) have been successfully used to reconstruct missing or masked regions of cosmological fields by sampling from the conditional distribution of the signal given the observed data [Wandelt et al., 2004, Jasche and Wandelt, 2013]. These approaches typically assume a Gaussian prior on the underlying field and generate statistically consistent realizations of the missing data.

In these frameworks, the inference is performed at the level of the full cosmological field, enabling the joint reconstruction of the signal and the estimation of model parameters. This represents the most complete formulation of the inference problem in the presence of incomplete or contaminated data.

In contrast, the present work adopts a reduced representation of the data, namely the 21-cm power spectrum, as a computationally tractable intermediate step. Rather than being the final objective, this reduced data space serves as a controlled testbed for developing and validating reconstruction-based inference techniques.

Moreover, most existing 21-cm analysis pipelines rely on either foreground avoidance or foreground subtraction and do not explicitly incorporate the reconstruction of missing

modes within the inference process. While forward-modelling approaches can, in principle, account for foregrounds, they generally do not treat missing data as explicit latent variables to be sampled during inference.

This highlights a gap between existing 21-cm inference approaches and fully Bayesian reconstruction frameworks developed in other areas of cosmology. The approach developed in this work aims to bridge this gap by formulating the inference problem in a way that can naturally extend from reduced summary statistics to full field-level reconstruction, ultimately enabling joint inference of the underlying HI signal and astrophysical parameters.

4.6 Challenges in Inference with Incomplete Data

The approaches discussed above represent different strategies for performing inference in the presence of foreground-contaminated and incomplete data, each with distinct advantages and limitations.

- Avoidance-based methods are robust but discard a significant fraction of the data, limiting the available information for inference [Liu and Tegmark, 2011].
- Foreground removal techniques attempt to recover additional information by modelling and subtracting foregrounds, but are susceptible to residual contamination and signal loss [Chapman et al., 2013].
- Forward-modelling approaches provide a statistically consistent framework by jointly modelling signal, foregrounds, and noise, but are computationally expensive and sensitive to modelling assumptions [Greig and Mesinger, 2017].
- Reconstruction-based methods offer a principled framework for handling incomplete or contaminated data and have been successfully applied in cosmological analyses such as CMB map reconstruction and large-scale structure inference [Wandelt et al., 2004, Jasche and Wandelt, 2013]. However, in the context of 21-cm cosmology, their application has largely been confined to data-level or field-based approaches, and they have not yet been systematically integrated into parameter inference frameworks based on reduced observables such as the power spectrum.

A common limitation across these approaches is that they either discard contaminated modes or do not fully incorporate the missing information into the inference process. This motivates the development of frameworks that explicitly treat foreground contamination as a missing-data problem.

In the next chapter, we build upon this idea and develop a reconstruction-inspired inference framework using Gaussian Constrained Realization(GCR), tailored to emulator-based likelihoods for the 21-cm power spectrum. This approach represents a step toward integrating reconstruction and parameter inference within a unified Bayesian framework.

5 Signal Reconstruction using Gaussian Constrained Realizations

5.1 Reconstruction-Based Inference: Overview

Foreground contamination in 21-cm observations leads to incomplete measurements in Fourier space, where a subset of modes is either contaminated or entirely unavailable [Liu and Tegmark, 2011, Iliev et al., 2015]. As discussed in the previous chapter, traditional approaches, such as foreground avoidance, discard these modes, while foreground removal techniques attempt to model and subtract foregrounds [Liu and Tegmark, 2011]. Both approaches suffer from important limitations, either through loss of information or the introduction of bias due to imperfect modelling.

An alternative perspective is to treat the inference problem as one of incomplete data, in which the missing or contaminated modes are incorporated directly within the statistical framework [Jasche and Wandelt, 2013, Wandelt et al., 2004]. In this view, foreground contamination does not simply degrade the data, but fundamentally removes information that must be accounted for during inference. Reconstruction-based methods naturally arise from this perspective. Rather than discarding or subtracting contaminated modes, these methods aim to infer the missing information by exploiting the statistical correlations present in the underlying signal. This allows for a more complete use of the available data and provides a principled way to propagate uncertainties associated with missing modes.

Such approaches have been successfully developed in other areas of cosmology, including Cosmic Microwave Background (CMB) analysis and large-scale structure inference, where incomplete sky coverage and survey masks are common [Jasche and Wandelt, 2013, Wandelt et al., 2004]. In these contexts, reconstruction techniques enable the recovery of missing regions in a statistically consistent manner.

The primary objective of this work is to develop a reconstruction-based inference framework for 21-cm cosmology, with the long-term goal of reconstructing the full three-dimensional HI signal in the presence of foreground contamination and observational systematics.

5.2 Gaussian Constrained Realizations

5.2.1 Conceptual Framework

Gaussian Constrained Realizations (GCR) provide a statistically rigorous method for reconstructing missing or unobserved components of a signal under the assumption that the underlying field can be described as a Gaussian random process [Wandelt et al., 2004, Hoffman and Ribak, 1991, Iliev et al., 2015].

In this framework, the signal is fully characterized by its covariance structure, which encodes correlations between different modes. In cosmological applications, this covariance is typically determined by the power spectrum, linking the reconstruction directly to the underlying physics of the signal.

We consider a data vector that can be partitioned into observed and missing components. The central idea is that these components are statistically correlated, allowing the missing modes to be inferred in a manner consistent with both the observed data and the assumed statistical properties of the field.

5.2.2 Mathematical Formulation

Under the Gaussian assumption, the full data vector follows a multivariate Gaussian distribution. The conditional distribution of the missing data given the observed data is also Gaussian [Wandelt et al., 2004, Jasche and Wandelt, 2013]:

$$P(\mathbf{d}_{\text{miss}} \mid \mathbf{d}_{\text{obs}}) = \mathcal{N}(\boldsymbol{\mu}_{\text{cond}}, \mathbf{C}_{\text{cond}}) \tag{5.1}$$

with

$$\boxed{\boldsymbol{\mu}_{\text{cond}} = \mathbf{C}_{\text{miss,obs}}\, \mathbf{C}_{\text{obs,obs}}^{-1}\, \mathbf{d}_{\text{obs}}} \tag{5.2}$$

$$\boxed{\mathbf{C}_{\text{cond}} = \mathbf{C}_{\text{miss,miss}} - \mathbf{C}_{\text{miss,obs}}\, \mathbf{C}_{\text{obs,obs}}^{-1}\, \mathbf{C}_{\text{obs,miss}}} \tag{5.3}$$

5.2.3 Interpretation

Here, $\mathbf{d}_{\text{obs}}$ represents the observed modes and $\mathbf{d}_{\text{miss}}$ the missing or contaminated modes. The matrices $\mathbf{C}_{\text{obs,obs}}$, $\mathbf{C}_{\text{miss,miss}}$, and $\mathbf{C}_{\text{miss,obs}}$ describe correlations within observed modes, within missing modes, and between the two, respectively.

The conditional mean,

$$\boldsymbol{\mu}_{\text{cond}} = \mathbf{C}_{\text{miss,obs}}\, \mathbf{C}_{\text{obs,obs}}^{-1}\, \mathbf{d}_{\text{obs}},$$

provides the best linear estimate of the missing modes given the observed data. It describes how information propagates from observed to missing regions through the covariance structure. Modes that are strongly correlated with the observed data are reconstructed with higher accuracy, while weakly correlated modes remain less constrained.

The conditional covariance,

$$\mathbf{C}_{\text{cond}} = \mathbf{C}_{\text{miss,miss}} - \mathbf{C}_{\text{miss,obs}}\, \mathbf{C}^{-1}_{\text{obs,obs}}\, \mathbf{C}_{\text{obs,miss}},$$

quantifies the uncertainty remaining after conditioning on the observed data. The first term represents the prior uncertainty, while the second term captures the reduction in uncertainty due to information from the observed modes.

In cosmological applications, this implies that reconstruction accuracy is governed by the strength and scale of correlations in the signal [Jasche and Wandelt, 2013]. Large-scale correlations enable efficient propagation of information into masked regions, whereas small-scale or weakly correlated structures are reconstructed with greater uncertainty.

5.2.4 Role in Cosmological Signal Reconstruction

Cosmological signals exhibit correlations across scales due to the underlying physics of structure formation. As a result, observed modes contain indirect information about missing modes, making reconstruction possible even in the presence of incomplete data [Liu and Tegmark, 2011].

In the context of 21-cm cosmology, foreground contamination removes modes in a structured region of Fourier space (the wedge), leaving gaps in the $(k_\perp, k_\parallel)$ plane [Liu and Tegmark, 2011]. The reconstruction problem, therefore, reduces to inferring these missing modes using their statistical correlations with the observed modes outside the wedge.

GCR has been widely applied in cosmology, particularly in CMB analysis and large-scale structure inference, where incomplete observations are common [Hoffman and Ribak, 1991, Wandelt et al., 2004]. In these contexts, it enables the reconstruction of missing regions while consistently propagating uncertainties.

5.3 Signal Reconstruction Framework in This Work

5.3.1 Objective and Scope

The central objective of this work is to develop a framework for reconstructing the cosmological 21-cm signal in the presence of foreground contamination and incomplete data. In its most general form, this requires inference at the level of the full three-dimensional HI signal, where missing or contaminated modes are treated as unknown quantities and reconstructed jointly with model parameters.

Such a formulation represents the most complete description of the inference problem, but is computationally demanding and requires detailed modelling of observational systematics.

5.3.2 Reduced Representation as a Controlled Testbed

To make the problem tractable, we adopt a reduced representation of the cosmological 21-cm signal in the form of the two-dimensional 21-cm power spectrum, the preferred summary statistics of the signal, which is being used for its first detection. This provides a simplified yet physically meaningful description of the signal while significantly reducing computational complexity [Iliev et al., 2015].

The power spectrum is therefore used as a controlled testbed for developing reconstruction-based inference techniques, rather than as the final objective.

5.3.3 Reconstruction of Missing Power Spectrum Modes

A key component of this work is the explicit reconstruction of missing modes in the two-dimensional power spectrum. These missing modes arise due to foreground contamination and data selection effects. Instead of discarding these modes, we reconstruct them by exploiting statistical correlations between different Fourier modes. The observed modes act as constraints, while the covariance structure determines how information propagates into missing regions.

This results in a reconstructed power spectrum in which all modes are populated with statistically consistent values, while retaining uncertainty associated with the missing information. The effectiveness of this reconstruction is therefore determined by how information propagates from observed modes into the masked region. This propagation is governed by the covariance structure, which encodes both the strength and scale of correlations.

As will be demonstrated in the following chapter, this leads to a clear scale-dependent reconstruction behaviour: modes that are strongly correlated with observed data are reconstructed with low error, while modes that are weakly connected or distant from observed regions exhibit higher uncertainty.

5.3.4 Inference Pipeline for the Reconstructed Power Spectrum

The reconstructed power spectrum forms the input to the inference pipeline. The main components are:

- **Observed Data:** Measured modes in Fourier space after foreground masking.
- **Reconstructed Data:** Combination of observed and reconstructed modes.
- **Emulator-Based Model:** A Bayesian Neural Network (BNN) emulator predicts the power spectrum for given parameters.
- **Likelihood Construction:** Comparison between reconstructed data and emulator predictions.
- **Parameter Estimation:** Exploration of parameter space to obtain posterior distributions.

At present, the reconstruction stage has been implemented, while the emulator-based inference is under development.

5.3.5 Implementation of Gaussian Constrained Realizations

The reconstruction of missing modes follows the principles of Gaussian Constrained Realizations [Hoffman and Ribak, 1991]. The power spectrum is treated as a random field with a covariance structure derived from simulations.

The data are divided into observed and missing components, with the missing modes treated as latent variables and inferred from their conditional distribution given the observed modes, ensuring statistical consistency. In this framework, the covariance matrix acts as a prior, encoding correlations between modes, while the observed data provide the constraints. The reconstructed power spectrum, therefore, represents a single realization drawn from the distribution of all spectra consistent with the observations, and this probabilistic interpretation naturally ensures proper uncertainty propagation.

It is important to note that the Gaussian assumption represents an approximation, particularly for the 21-cm signal at later stages of reionization where the field becomes increasingly non-Gaussian. In this work, the Gaussian model is treated as an effective description that captures the dominant correlation structure relevant for reconstruction.

5.4 Discussion and Outlook

This chapter has introduced Gaussian Constrained Realizations as a principled framework for handling incomplete data in cosmological inference. By leveraging statistical correlations, GCR enables the reconstruction of missing information in a manner consistent with both the observed data and the assumed statistical model.

The formulation developed here provides a clear interpretation of reconstruction performance in terms of covariance-driven information propagation. In particular, the ability to reconstruct missing modes depends not only on the amount of missing data, but also on the strength and scale of correlations in the signal. This framework forms the foundation for the results presented in the next chapter, where reconstruction performance is evaluated across different stages of reionization and masking scenarios. These results demonstrate that reconstruction accuracy is inherently scale-dependent and closely linked to the structure of the underlying signal.

Future work will focus on extending this approach to full three-dimensional HI signal reconstruction and incorporating more flexible models capable of capturing non-Gaussian features of the signal.

6 Towards a Unified Inference Framework for EoR 21-cm Signal

6.1 Motivation and Overview

Inference of the 21-cm signal from the epoch of reionization presents a combination of theoretical and computational challenges. On the observational side, foreground contamination removes a significant fraction of Fourier modes, leading to incomplete data in the $(k_\perp, k_\parallel)$ plane. On the modelling side, accurate predictions of the signal require expensive numerical simulations, making repeated likelihood evaluations computationally demanding. Together, these issues complicate both parameter estimation and signal reconstruction.

A comprehensive solution to this problem requires a framework that can simultaneously address three key aspects of the inference process. First, an efficient sampling strategy is needed to explore the high-dimensional and often correlated parameter space. Second, a fast and reliable forward model is required to predict the 21-cm signal for a given set of Epoch of Reionization(EoR) parameters. Third, a principled method is needed to account for missing or contaminated modes in the observed data.

In this work, we develop a unified inference framework that integrates these components within a modular pipeline. The sampling strategy is based on a Metropolis-within-Gibbs approach, designed to improve convergence in correlated parameter spaces. The forward model is provided by a Bayesian Neural Network (BNN) emulator, enabling rapid and uncertainty-aware predictions of the 21-cm power spectrum. To address the problem of incomplete data, we employ a reconstruction framework based on Gaussian Constrained Realizations (GCR), which infers missing modes using statistical correlations in the signal.

Rather than introducing all components simultaneously, we adopt a staged approach. We first validate the performance of the sampling strategy in an idealized, foreground-free setting, where the behaviour of the inference algorithm can be studied in isolation. We then investigate the reconstruction of missing modes in the two-dimensional power spectrum under realistic masking scenarios. Finally, we discuss the current status of the emulator and outline the challenges involved in extending it to higher-dimensional observables, particularly the two-dimensional power spectrum.

This structure allows us to build toward a fully integrated inference pipeline, while clearly isolating and understanding the role of each component.

6.2 Inference Framework and Methodology

The inference framework developed in this work is composed of three core components: a sampling strategy for exploring the posterior distribution, a forward model for predicting the observable signal, and a reconstruction framework for handling incomplete data. While each component addresses a distinct aspect of the problem, they are designed to operate together within a unified pipeline.

The first component is the sampling strategy, which is responsible for exploring the posterior distribution of astrophysical parameters. In high-dimensional inference problems, especially those involving correlated parameters, standard sampling methods can suffer from slow convergence and inefficient exploration. To address this, we adopt a Metropolis-within-Gibbs (MwG) approach, in which parameters are updated sequentially rather than jointly. By decomposing the multi-dimensional proposal into a sequence of one-dimensional updates, this method improves mixing and reduces the impact of parameter correlations, leading to more efficient exploration of the posterior.

The second component is the forward model, which maps astrophysical parameters to observable quantities. In the context of 21-cm cosmology, this typically requires computationally expensive simulations of the signal. To enable efficient inference, we replace these simulations with a Bayesian Neural Network (BNN) emulator. The emulator is trained on simulated data and provides fast predictions of the 21-cm power spectrum for a given parameter set, along with an estimate of predictive uncertainty. This uncertainty is incorporated directly into the likelihood, ensuring a consistent probabilistic treatment of model errors during inference.

The third component is the reconstruction framework, which addresses the problem of missing data arising from foreground contamination. In Fourier space, foregrounds remove a structured region of modes, commonly referred to as the wedge, leaving gaps in the observed data. Rather than discarding these modes or attempting to subtract foregrounds, we treat them as missing variables and reconstruct them using Gaussian Constrained Realizations (GCR). This approach exploits the statistical correlations of the signal to generate realizations of the missing modes that are consistent with the observed data and the assumed covariance structure.

A key feature of this framework is its modular design. The sampling, forward modelling, and reconstruction components are implemented independently, allowing each to be validated in isolation before being combined into a full inference pipeline. This modularity not only simplifies development and testing, but also provides flexibility for future extensions, such as improving the emulator or incorporating more sophisticated reconstruction techniques.

In the following sections, we first focus on validating the sampling strategy in an idealized setting. We then examine the performance of the reconstruction framework in the presence of incomplete data, before discussing how these components can be integrated into a unified inference pipeline.

6.3 Staged Evaluation of the Inference Framework

To assess the performance of the proposed inference framework, we adopt a staged evaluation strategy. Rather than introducing all complexities simultaneously, we validate individual components of the pipeline in controlled settings before progressing to more realistic scenarios.

In the first stage, we focus on the sampling strategy in an idealized setting where foreground contamination is absent. This allows us to isolate the performance of the inference algorithm and evaluate its ability to recover model parameters using a well-defined likelihood. In the second stage, we introduce incomplete data by masking regions of Fourier space, mimicking the effect of foreground contamination, and assess the performance of the reconstruction framework in recovering missing modes of the 21-cm power spectrum.

6.3.1 Stage I: Sampler Validation in an Idealized Setting

In this test case, we aim to infer the three reionization parameters

$$\boldsymbol{\theta} = (M_{\rm min}, N_{\rm ion}, R_{\rm mfp})$$

given an observed 21-cm power spectrum D. Following the framework of [Mahida et al., 2025], the posterior distribution is written as

$$p(\boldsymbol{\theta} \mid D) \propto p(D \mid \boldsymbol{\theta})\, \Pi(\boldsymbol{\theta}), \tag{6.1}$$

where we assume uniform priors over physically motivated parameter ranges.

6.3.1.1 Likelihood Model and Error Covariance

The likelihood is modelled as a multivariate Gaussian distribution, requiring a forward model to predict the power spectrum for a given parameter set. In this work, the forward model is provided by a Bayesian Neural Network (BNN) emulator, which produces both a mean prediction $\boldsymbol{\mu}_\theta$ and a predictive uncertainty $\sigma_{\rm BNN}(i)$ for each k-bin.

The observational variance in each bin includes contributions from both sample variance and thermal noise,

$$\sigma^2_{\rm obs}(i) = \sigma^2_{\rm SV}(i) + \sigma^2_{\rm N}(i), \tag{6.2}$$

where sample variance arises from the finite number of Fourier modes per bin and thermal noise depends on instrumental properties and observing time.

The total variance used in the likelihood further incorporates the emulator predictive uncertainty,

$$\sigma^2_t(i) = \sigma^2_{\rm obs}(i) + \sigma^2_{\rm BNN}(i). \tag{6.3}$$

This leads to a diagonal covariance matrix

$$\Sigma = \mathrm{diag}\left[\sigma^2_t(1), \ldots, \sigma^2_t(n)\right], \tag{6.4}$$

and the corresponding log-likelihood is

$$\ln \mathcal{L}(\boldsymbol{\theta}) = -\frac{1}{2}\left[(\mathbf{P}_{\rm ref} - \boldsymbol{\mu}_\theta)^T \Sigma^{-1} (\mathbf{P}_{\rm ref} - \boldsymbol{\mu}_\theta) + \ln\det(2\pi\Sigma)\right]. \tag{6.5}$$

Foreground and calibration systematics are assumed to be perfectly removed in this idealized setup, allowing us to focus purely on the performance of the inference algorithm.

6.3.1.2 The Bayesian Neural Network

The Bayesian Neural Network (BNN) emulator is constructed to provide fast and uncertainty-aware predictions of the 21-cm power spectrum. The training dataset consists of simulated 21-cm maps generated for 7200 parameter sets. From these, 50 samples are reserved as a test set, while 80% of the remaining data is used for training and 20% for validation.

The emulator is implemented using the `PYRO` probabilistic programming library based on PyTorch. It takes the three reionization parameters $(M_{\rm min}, N_{\rm ion}, R_{\rm mfp})$ as input and predicts the power spectrum across 7 k-bins. The network architecture consists of three hidden layers with 50, 100, and 50 nodes respectively, using ELU activation functions.

Layer	**Nodes**	**Activation**
Input Layer	3	-
Hidden Layer 1	50	ELU
Hidden Layer 2	100	ELU
Hidden Layer 3	50	ELU
Output Layer	7	-

Table 6.1: Architecture of the BNN-based emulator developed in this work. The network takes three reionization parameters $(M_{\rm min}, N_{\rm ion}, R_{\rm mfp})$ as input and predicts the 21-cm power spectrum across 7 k-bins [Mahida et al., 2025].

Training is performed using Bayesian inference with a multivariate Gaussian likelihood, running an MCMC chain for 1000 steps with 500 warm-up iterations. This setup provides a balance between convergence and computational efficiency.

As an illustrative example, for the parameter set $M_{\rm min} = 467.37 \times 10^8\, M_\odot$, $N_{\rm ion} = 181.05$, $R_{\rm mfp} = 12.65$ Mpc, and a mass-averaged neutral fraction $\bar{x}_{\rm HI} = 0.377$, the BNN accurately predicts the 21-cm power spectrum across all 7 k-bins, with the true values lying within the 3σ confidence interval.

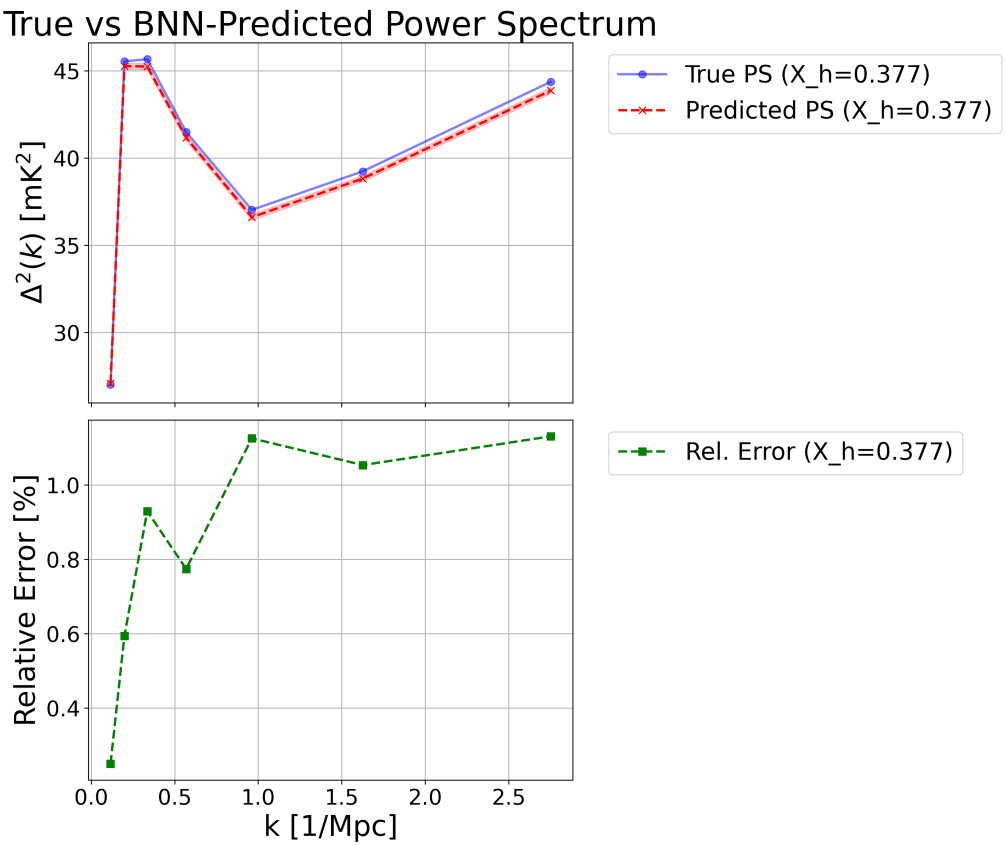


Figure 6.1: BNN prediction of the 21-cm power spectrum for a representative test case.

6.3.1.3 Sampling Techniques

We now consider two sampling strategies for exploring the posterior distribution.

Metropolis-Hastings (MH): In this approach, all parameters are jointly proposed at each step. While effective for general exploration, it can suffer from slow mixing in the presence of strong parameter correlations.

Metropolis-within-Gibbs (MwG): To improve efficiency, we implement a Gibbs sampling strategy in which parameters are updated sequentially:

1. update $M_{\rm min}$ while keeping $N_{\rm ion}$ and $R_{\rm mfp}$ fixed,
2. update $N_{\rm ion}$ while keeping $M_{\rm min}$ and $R_{\rm mfp}$ fixed,
3. update $R_{\rm mfp}$ while keeping $M_{\rm min}$ and $N_{\rm ion}$ fixed.

Since the conditional distributions are not analytically available, each update is performed using a Metropolis step. This decomposes the three-dimensional exploration problem into a sequence of one-dimensional updates, improving convergence and reducing correlations between samples.

6.3.1.4 Test Parameter Sets

For a direct comparison with the baseline BNN–MCMC results, we adopt the same three test parameter sets:

Test Set	$M_{\min}$	N_{ion}	R_{mfp}	x_{H}
1	675.26	48.42	17.26	0.844
2	176.32	38.95	33.40	0.627
3	467.37	181.05	12.65	0.377

Table 6.2: Three test parameter sets used for comparing BNN-based Gibbs sampling with MCMC. The corresponding neutral hydrogen fractions x_{H} indicate different stages of cosmic reionization: early (set 1), intermediate (set 2), and late (set 3).

These three parameter sets correspond to different stages of cosmic reionization, as indicated by their neutral hydrogen fractions (x_{H}). Test set 1 represents an early phase with a high neutral fraction ($x_{\text{H}} \sim 0.84$), set 2 probes an intermediate phase ($x_{\text{H}} \sim 0.63$), and set 3 corresponds to a late phase ($x_{\text{H}} \sim 0.38$).

For each test set, three independent chains are run for both samplers (MCMC and Gibbs), initialized at different points in parameter space. This setup enables a robust comparison of convergence behaviour and sampling performance across different stages of reionization.

6.3.2 Stage II: Reconstruction with Incomplete Data

Having validated the sampling strategy in an idealized setting, we now introduce incomplete data to evaluate the reconstruction component of the framework under more realistic observational conditions. Foreground contamination is modelled by masking regions of Fourier space, producing a wedge-like exclusion region in the $(k_\perp, k_\parallel)$ plane. The reconstruction problem is therefore formulated as inferring the missing modes from their statistical correlations with the observed modes.

This inference is carried out using Gaussian Constrained Realizations (GCR), which generate realizations of the missing data consistent with both the observed modes and the assumed covariance structure. In contrast to Stage I, where inference is performed on complete data, this stage explicitly incorporates the presence of missing modes into the analysis, allowing for a more realistic assessment of reconstruction performance.

To enable a controlled comparison, we consider the same three parameter sets introduced in Stage I, corresponding to early, intermediate, and late stages of reionization. For each parameter set, three masking configurations—optimistic, intermediate, and pessimistic—are applied, representing increasing levels of data loss in Fourier space. The reconstructed power spectrum is then compared to the corresponding true signal for each combination of parameter set and masking scenario, providing a systematic framework for evaluating reconstruction performance across different signal regimes and levels of incompleteness.

6.4 Stage I: Validation of Gibbs Sampling Against Metropolis–Hastings

In this section, we validate the performance of the Metropolis-within-Gibbs sampler against the standard Metropolis–Hastings (MH) algorithm for inferring the reionization parameters $(M_{\rm min}, N_{\rm ion}, R_{\rm mfp})$ using BNN-emulated 21-cm power spectra.

The comparison is carried out using the three test parameter sets introduced earlier, which span early, intermediate, and late stages of reionization. For each parameter set, we present trace plots and corner plots for both sampling methods, allowing a direct comparison of convergence behaviour, mixing efficiency, and posterior structure.

6.4.1 Parameter Set 1: (176.32, 38.95, 33.40)

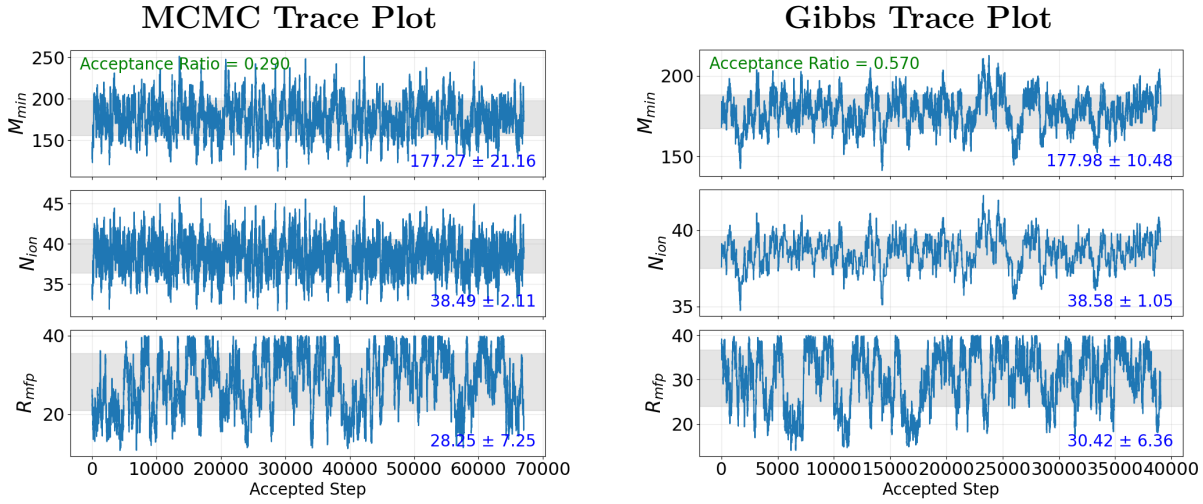


Figure 6.2: Trace plots for parameter set 1

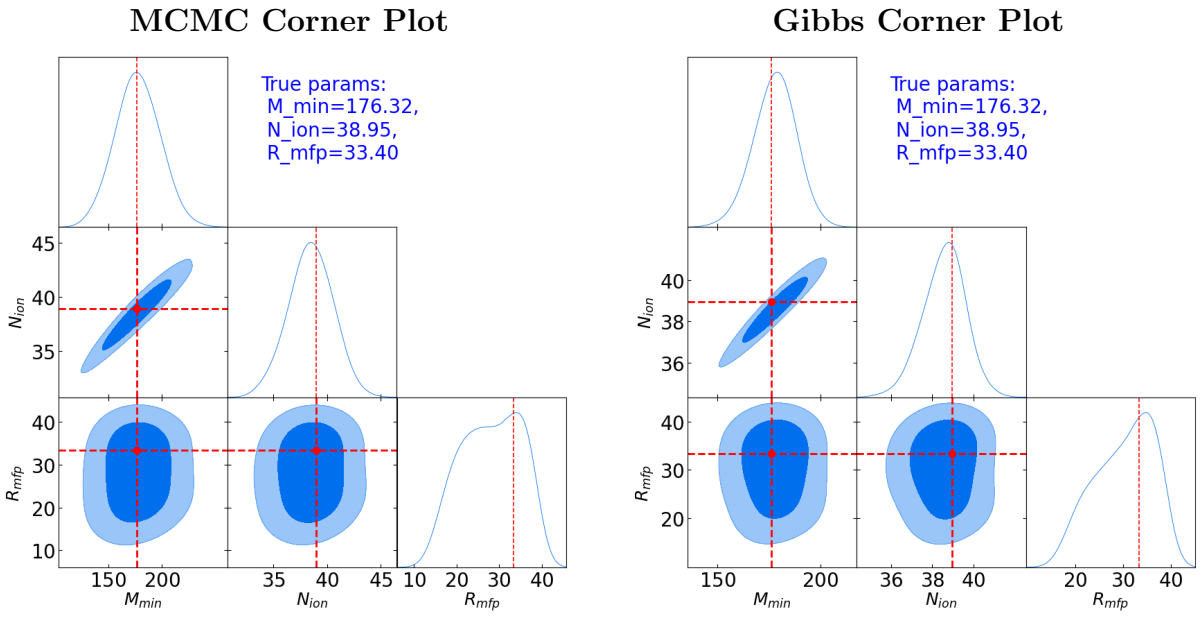


Figure 6.3: Corner plots for parameter set 1

6.4.2 Parameter Set 2: (675.26, 48.42, 17.26)

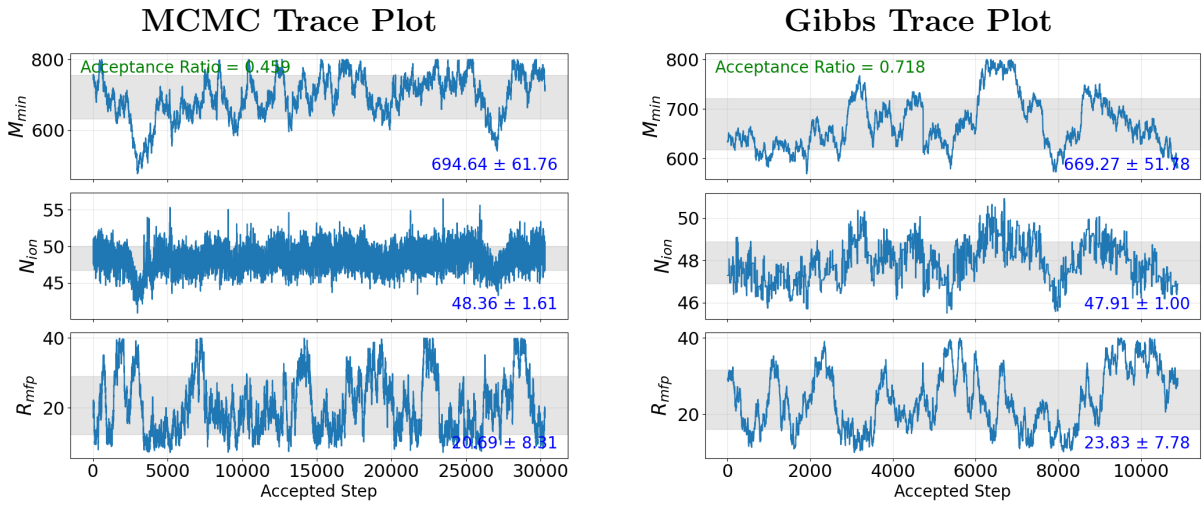


Figure 6.4: Trace plots for parameter set 2

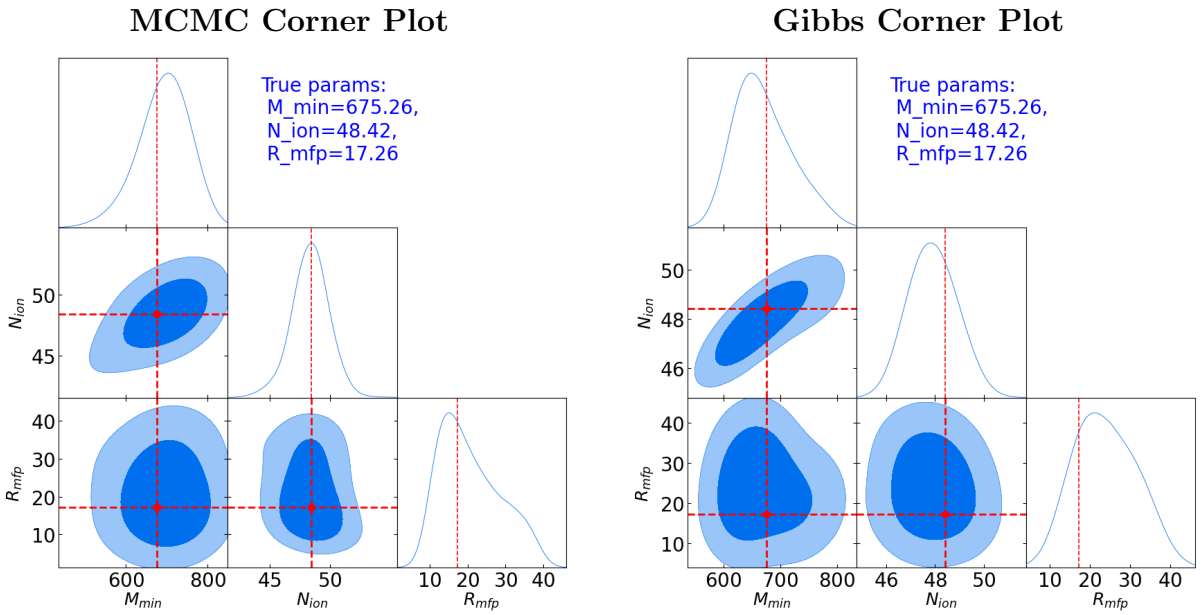


Figure 6.5: Corner plots for parameter set 2

6.4.3 Parameter Set 3: (467.37, 181.05, 12.65)

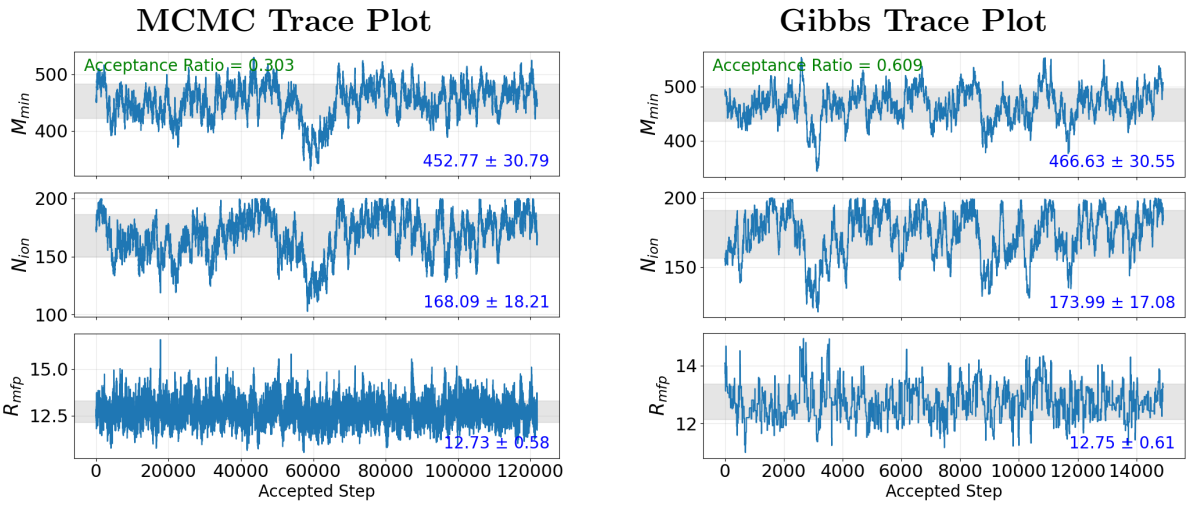


Figure 6.6: Trace plots for parameter set 3

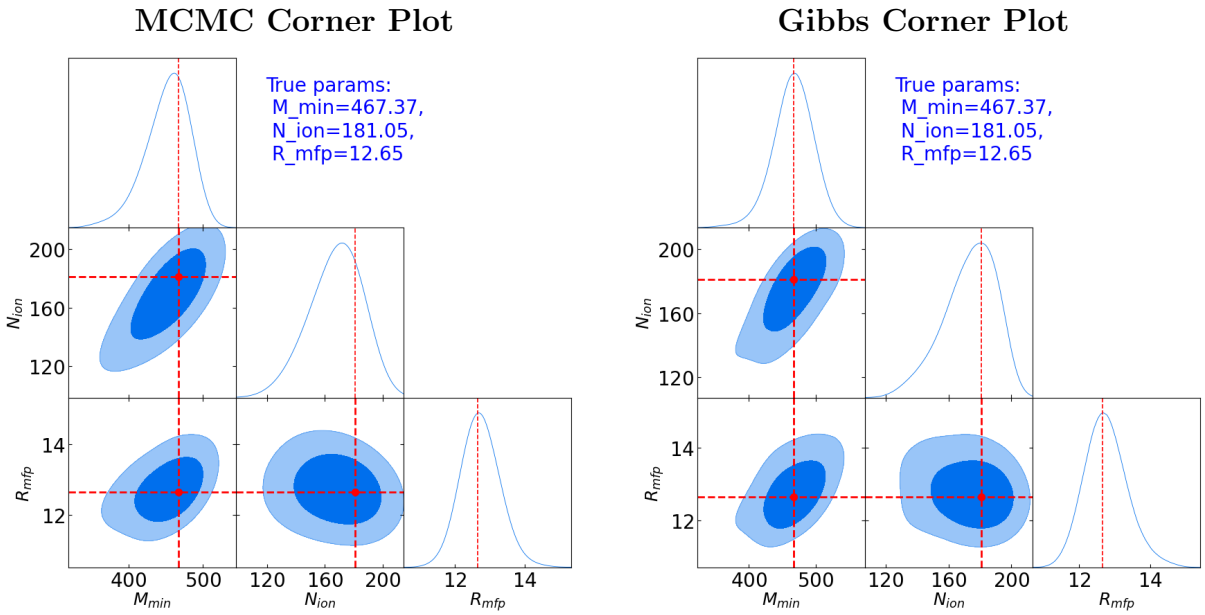


Figure 6.7: Corner plots for parameter set 3

6.4.4 Parameter Recovery and Sampling Efficiency

The recovered parameter values for both sampling methods are summarised in Table 6.3. Across all three test cases, the Gibbs sampler produces posterior estimates that are consistent with those obtained using the MH algorithm, while generally yielding reduced uncertainties.

Parameter Set	Sampler	$\mathbf{M}_{\mathrm{min}}$	$\mathbf{N}_{\mathrm{ion}}$	$\mathbf{R}_{\mathrm{mfp}}$
(176.32, 38.95, 33.40)	MCMC	177.17 ± 21.10	38.48 ± 2.10	28.12 ± 7.30
	Gibbs	177.98 ± 10.48	38.58 ± 1.05	30.42 ± 6.36
(675.26, 48.42, 17.26)	MCMC	702.65 ± 60.02	48.48 ± 1.59	20.77 ± 8.26
	Gibbs	670.73 ± 52.53	47.93 ± 1.01	23.85 ± 7.88
(467.37, 181.05, 12.65)	MCMC	452.77 ± 30.79	168.09 ± 18.21	12.73 ± 0.58
	Gibbs	466.63 ± 30.55	173.99 ± 17.08	12.75 ± 0.61

Table 6.3: Recovered parameter values for the three test parameter sets using MCMC and Gibbs sampling.

The relatively broad posterior distributions observed for R_{mfp} reflect known degeneracies with other parameters, rather than limitations of the sampling methods.

Sampling efficiency metrics are presented in Table 6.4. The Gibbs sampler consistently achieves higher acceptance ratios while requiring significantly fewer sampling steps.

Parameter Set	Sampler	Acceptance Ratio	Steps
(176.32, 38.95, 33.40)	MCMC	0.290	70000
	Gibbs	0.570	35000
(675.26, 48.42, 17.26)	MCMC	0.459	40000
	Gibbs	0.718	15000
(467.37, 181.05, 12.65)	MCMC	0.303	120000
	Gibbs	0.609	15000

Table 6.4: Acceptance ratios and number of sampling steps for MCMC and Gibbs sampling.

Overall, the Gibbs sampler demonstrates improved sampling efficiency, achieving faster convergence and smoother exploration of the posterior while maintaining consistency with the MH results.

6.5 Stage II: Reconstruction of the 2D Signal Power Spectrum

In this section, we present the results of the reconstruction framework applied to the same three parameter sets introduced in Stage I. These correspond to early, intermediate, and late stages of reionization, allowing a direct comparison between the idealized inference setting and the case with incomplete data.

Foreground contamination is modelled by masking regions of Fourier space, producing a wedge-like exclusion region in the $(k_\perp, k_\parallel)$ plane. For each parameter set, we consider three masking configurations—optimistic, intermediate, and pessimistic—representing increasing levels of data loss. This results in a 3×3 set of test cases.

For each configuration, we present both the reconstructed power spectrum and the corresponding relative error. This allows a direct assessment of reconstruction quality across different regions of Fourier space.

6.5.1 Signal Reconstruction: Early Stages of Reionization

At early times, the 21-cm signal is relatively weak and close to Gaussian, with limited contrast across scales. While this simplifies the statistical description of the signal, the low amplitude and weak correlations reduce the information available to constrain missing modes.

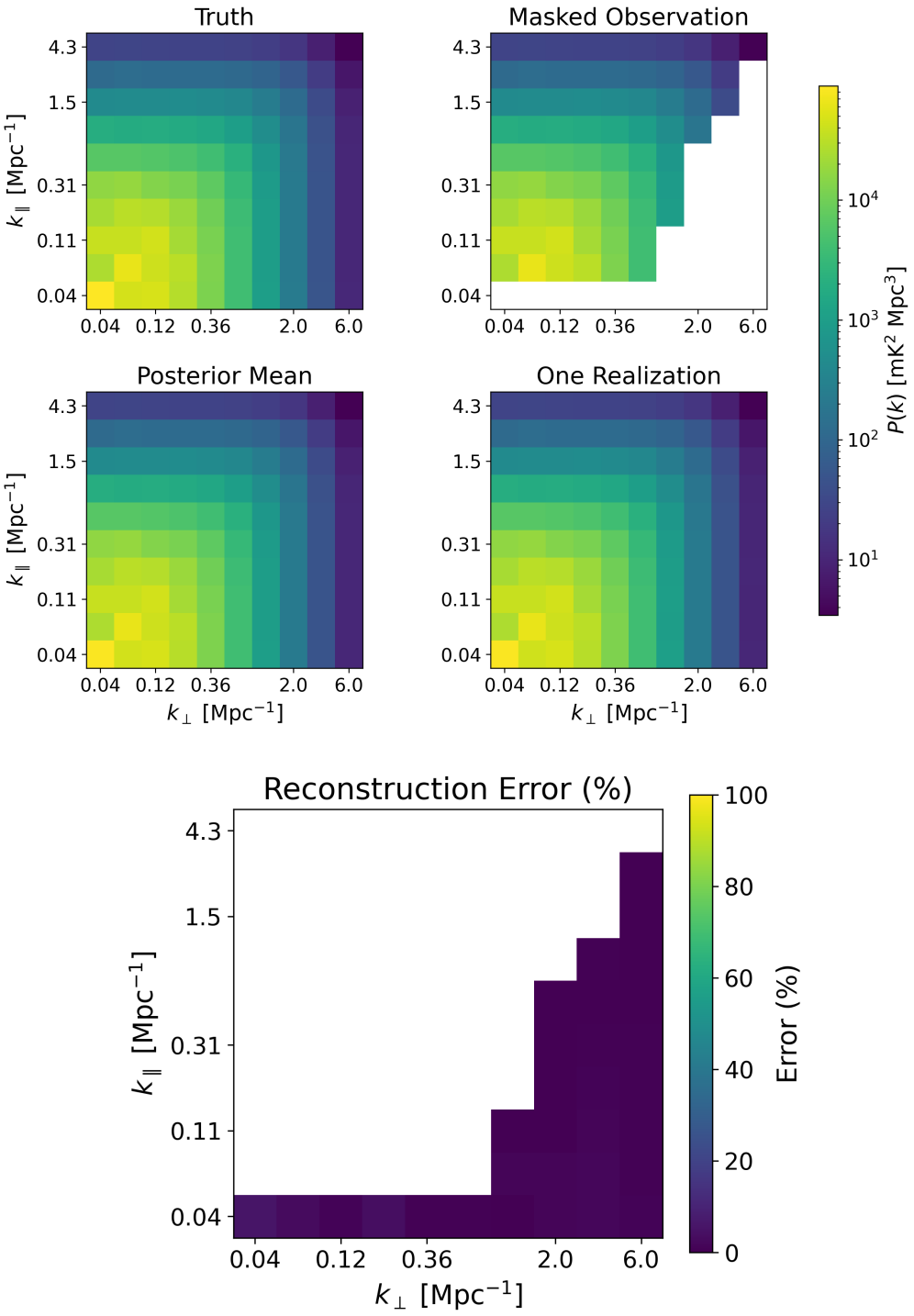


Figure 6.8: Early stage under optimistic masking. Top: true, masked, posterior mean, and sampled reconstruction. Bottom: relative error. The reconstruction remains accurate across all accessible modes.

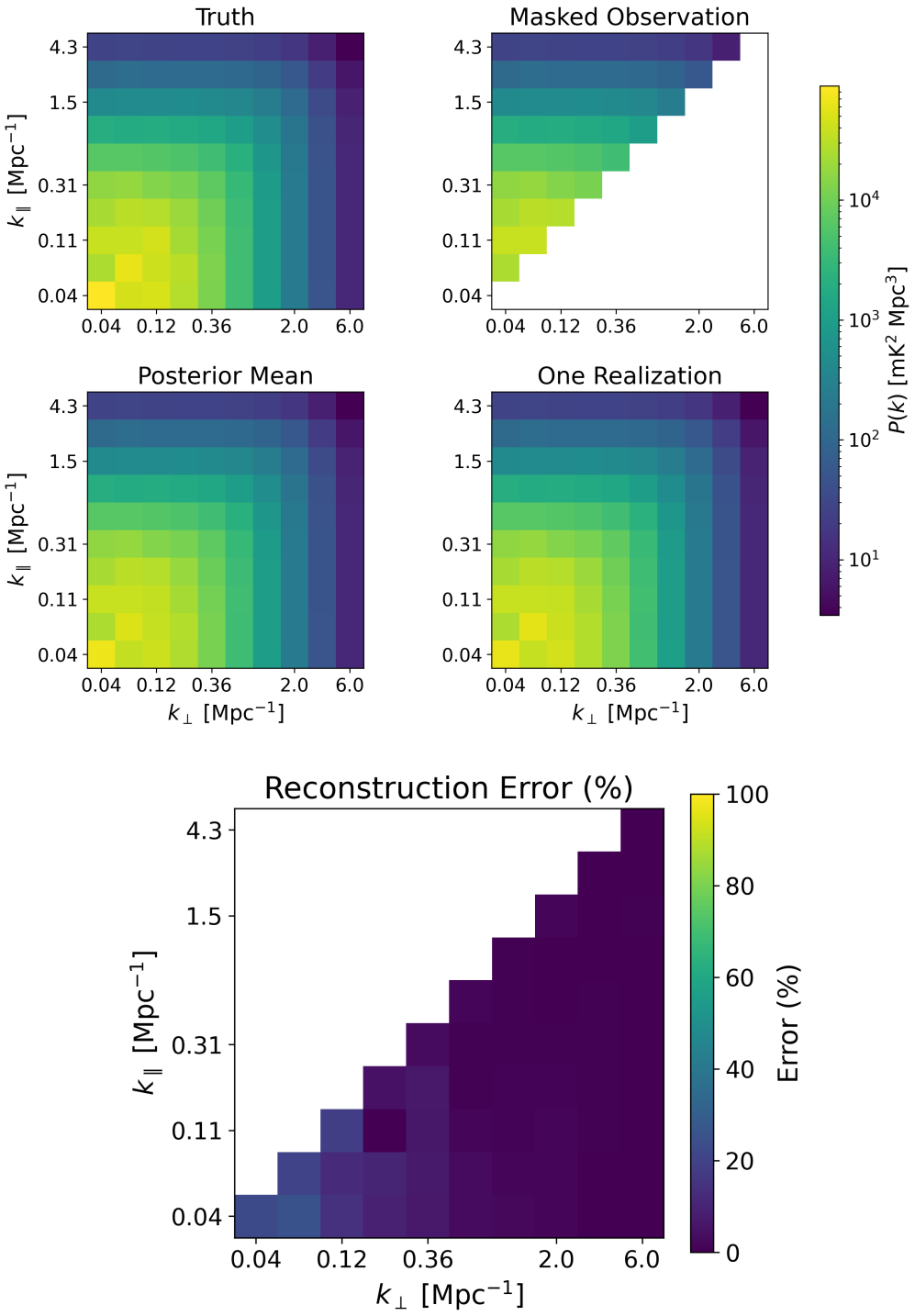


Figure 6.9: Early stage under intermediate masking. The global structure is recovered, though errors increase within the masked region, particularly at low-k modes.

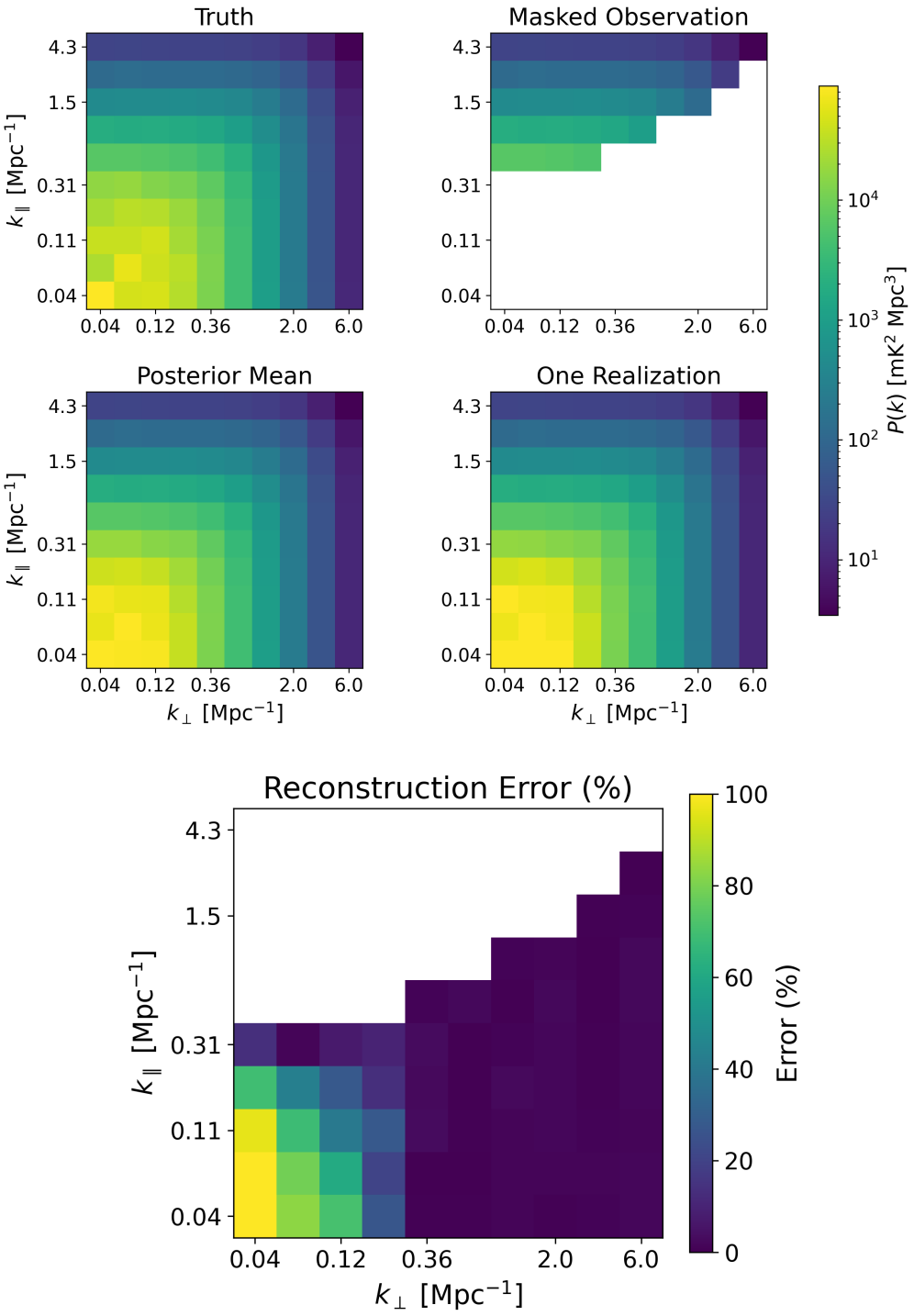


Figure 6.10: Early stage under pessimistic masking. Significant deviations appear in the masked region, particularly at low-k modes.

The reconstruction in this regime is limited by the low signal amplitude and weak correlations, which reduce the ability to propagate information into masked regions.

6.5.2 Signal Reconstruction: Intermediate Stage of Reionization

At intermediate stages, the signal develops stronger structure across multiple scales, leading to improved correlations and enhanced reconstruction performance.

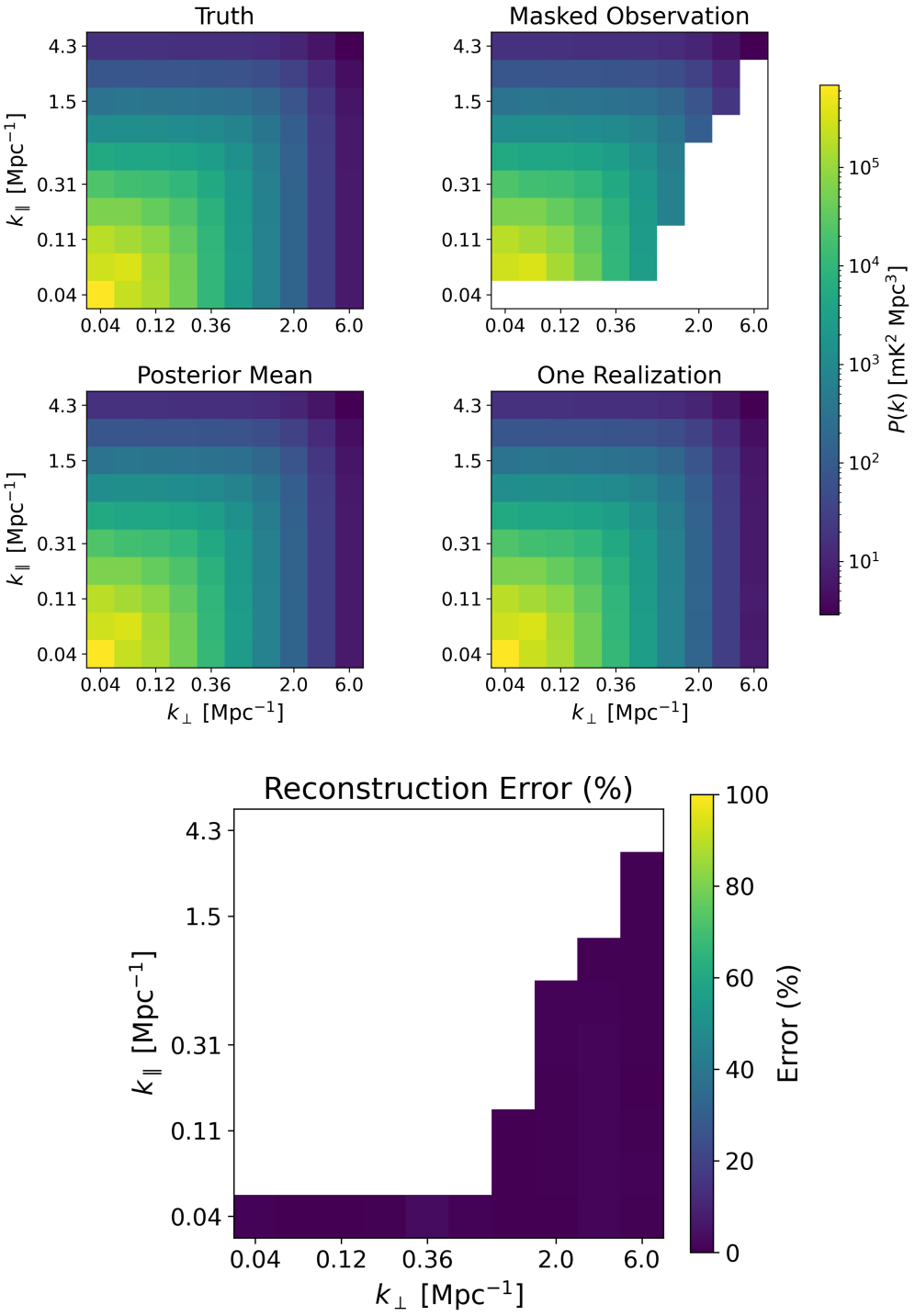


Figure 6.11: Intermediate stage under optimistic masking. Both large- and intermediate-scale features are accurately recovered with low relative error.

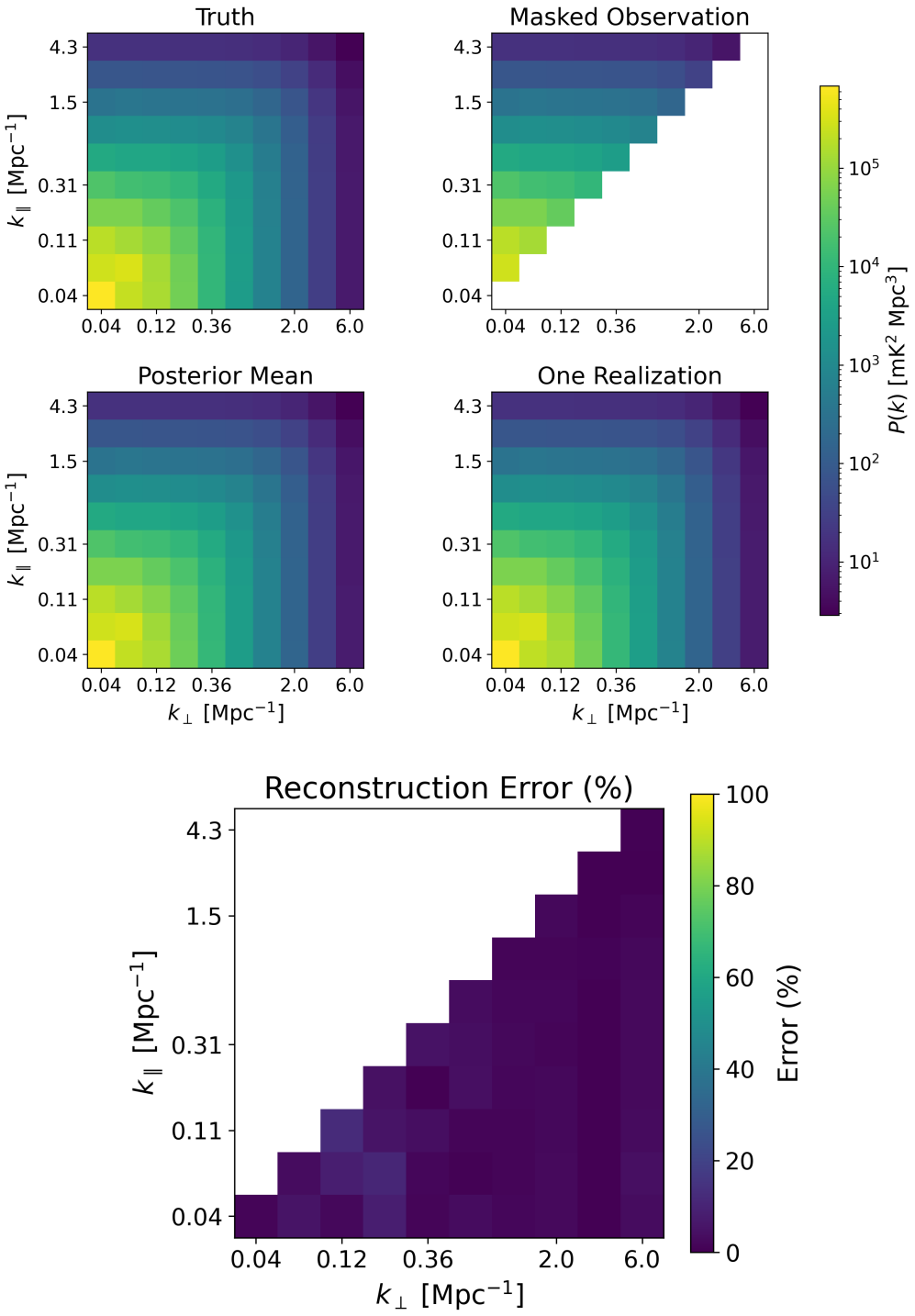


Figure 6.12: Intermediate stage under intermediate masking. Errors increase within the masked wedge but remain controlled.

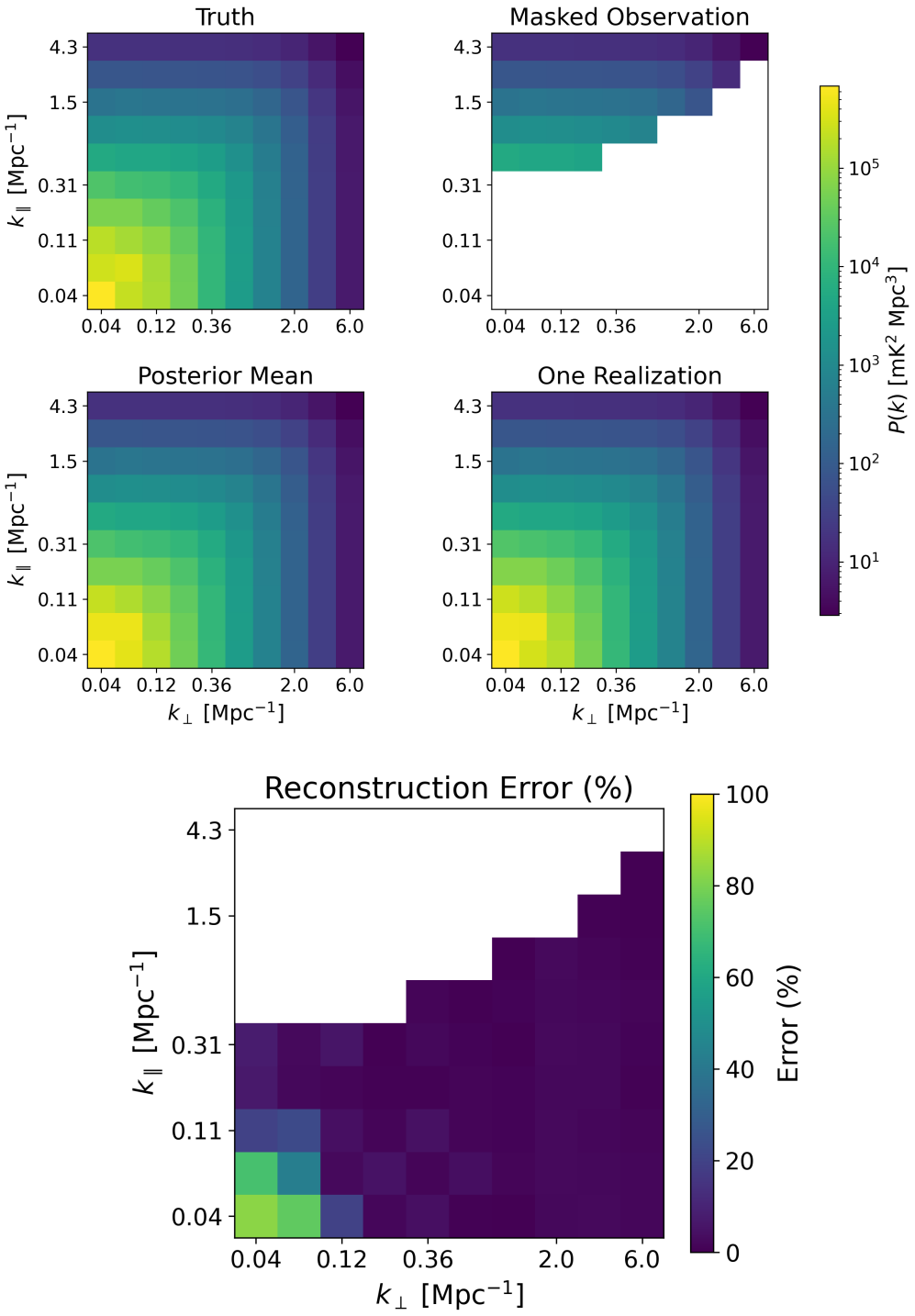


Figure 6.13: Intermediate stage under pessimistic masking. Small-scale features are increasingly smoothed, and errors grow in regions far from observed modes.

Compared to the early stage, reconstruction improves due to stronger signal contrast, although performance degrades as masking becomes more severe.

6.5.3 Signal Reconstruction: Late Stage of Reionization

At late times, the signal exhibits strong large-scale contrast and pronounced structure due to ionized regions. This enhances correlations at large scales while introducing increased complexity at smaller scales.

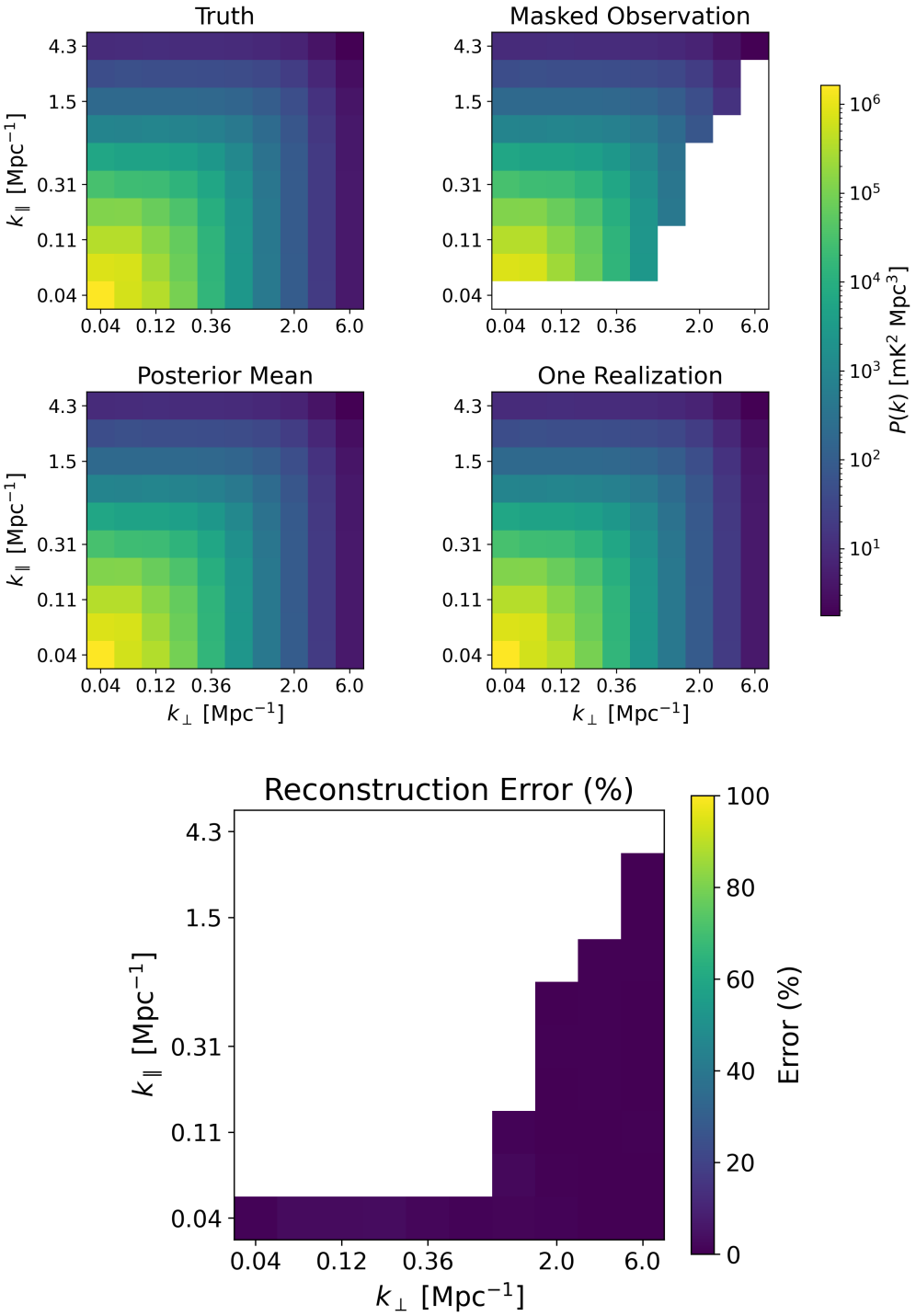


Figure 6.14: Late stage under optimistic masking. Large-scale modes are well reconstructed with low error, while small-scale deviations begin to appear.

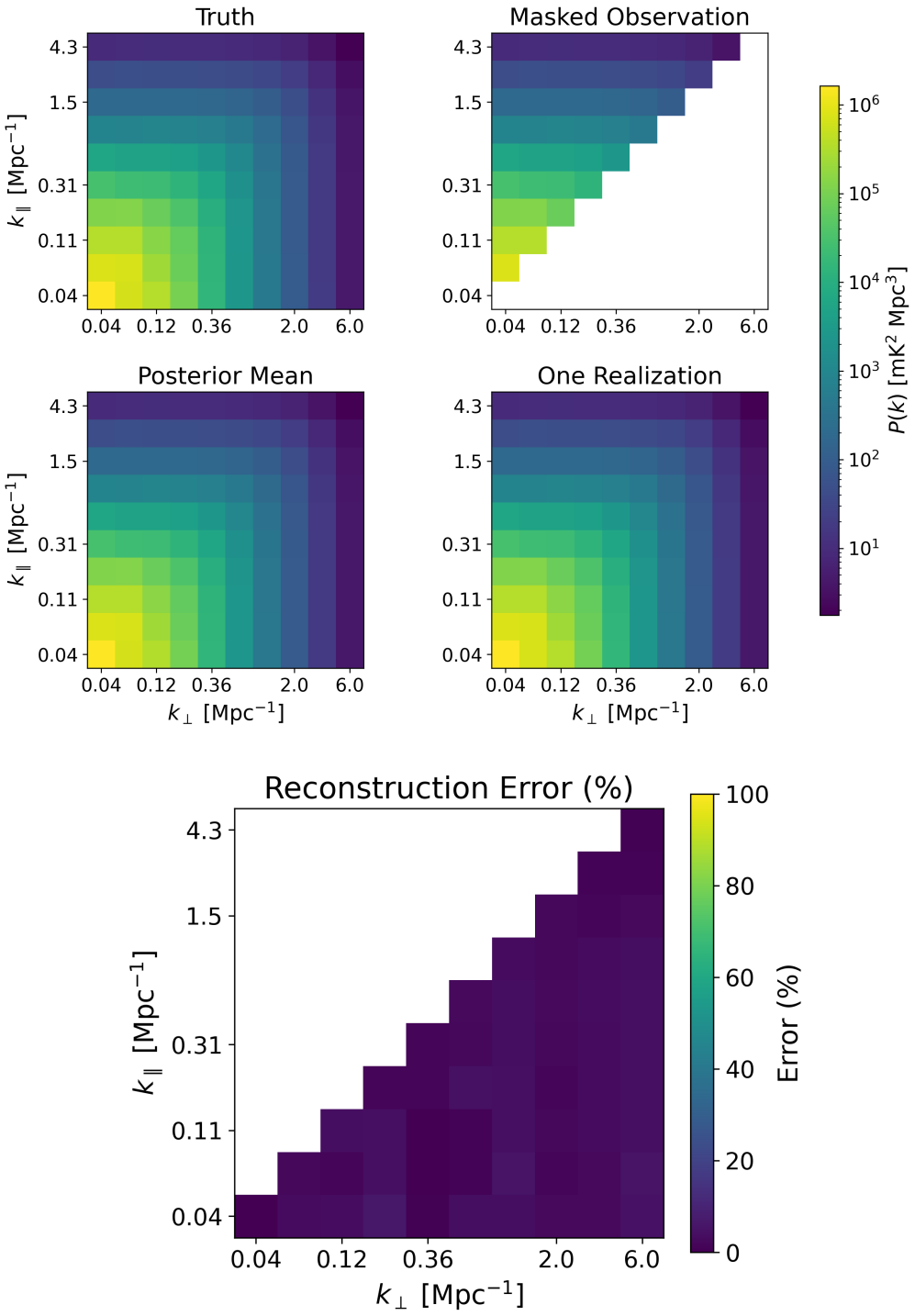


Figure 6.15: Late stage under intermediate masking. Errors remain low at large scales but increase at smaller scales within the masked region.

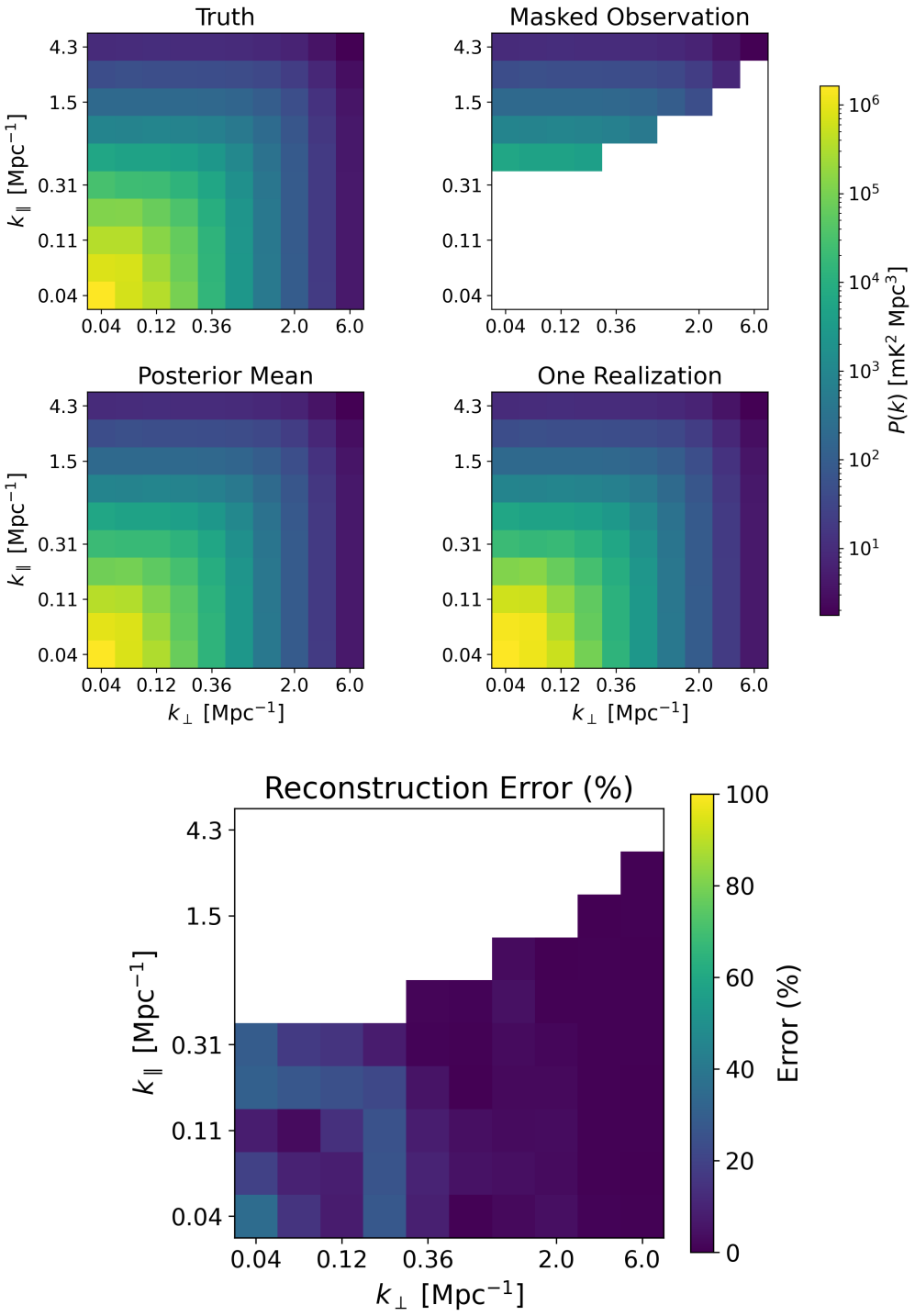


Figure 6.16: Late stage under pessimistic masking. Large-scale structure is retained, while small-scale modes exhibit increased uncertainty.

This stage demonstrates improved reconstruction at large scales compared to earlier stages, driven by stronger signal contrast and enhanced correlations.

6.5.4 Interpretation

The reconstruction performance depends on both the extent of masking and the intrinsic structure of the signal, exhibiting a clear scale dependence. In general, the accuracy is governed by the strength of correlations and the information content available in the observed modes, rather than the degree of Gaussianity alone. As a result, errors are systematically larger in regions of Fourier space that lie far from observed modes, highlighting the critical role of data availability in constraining the reconstruction.

This behaviour is reflected in the comparison across different stages of reionization. The early-stage signal, although closer to Gaussian, has low amplitude and weak fluctuations, providing limited information and leading to larger reconstruction errors, particularly on large scales. In contrast, the late-stage signal exhibits stronger contrast and enhanced large-scale power, which improves reconstruction accuracy in low-k modes despite its increased non-Gaussianity.

6.6 Development of Emulator for the Reconstructed Signal and Future Directions

The inference framework developed in this work relies on a fast forward model to enable efficient exploration of the parameter space. In the present implementation, this role is fulfilled by a Bayesian Neural Network (BNN) emulator trained to predict the one-dimensional 21-cm power spectrum.

This emulator has been successfully integrated into the inference pipeline and validated in Stage I, where it enables rapid and uncertainty-aware likelihood evaluation. Its performance demonstrates that emulator-based inference provides a viable alternative to direct simulation-based approaches, significantly reducing computational cost while retaining statistical consistency. However, the reconstruction framework introduced in Stage II operates on the two-dimensional power spectrum, which contains substantially more information than its one-dimensional counterpart. Extending the emulator to this higher-dimensional observable is therefore a necessary step toward a fully unified inference pipeline.

This extension is not a straightforward generalization. In the one-dimensional case, the emulator predicts a small number of k-bins, allowing for relatively simple modelling of the output space. In contrast, the two-dimensional power spectrum involves a significantly larger number of bins across the $(k_\perp, k_\parallel)$ plane, leading to a high-dimensional and highly correlated output space. Accurately capturing these correlations while maintaining reliable uncertainty estimates presents a non-trivial challenge.

Furthermore, the increased dimensionality of the output requires a corresponding expansion in the training dataset and model capacity, which can introduce additional computational and statistical complexities. In particular, ensuring stable training and meaningful uncertainty quantification becomes increasingly difficult as the dimensionality of the prediction space grows.

As a result, the development of a robust emulator for the two-dimensional power spectrum is an ongoing effort. Future work will focus on designing architectures capa-

ble of handling structured, high-dimensional outputs, as well as exploring strategies for dimensionality reduction and improved uncertainty modelling.

The successful integration of such an emulator would enable the full realization of the inference framework presented in this work, allowing for joint parameter estimation and reconstruction within a single, coherent Bayesian pipeline.

7 Summary and Discussion

This work presents two complementary investigations aimed at improving parameter inference in 21-cm cosmology, addressing both the efficiency of Bayesian sampling methods and the challenges posed by incomplete observational data in the presence of foreground contamination.

The first part of the work focused on evaluating Bayesian sampling strategies in an idealized setting, where the full two-dimensional 21-cm power spectrum is assumed to be available without foreground contamination. This setup provides a controlled environment in which the performance of different sampling methods can be directly compared without the additional complexity introduced by missing data. In this context, we compared the Metropolis–Hastings (MH) MCMC sampler with a Gibbs sampling approach for inferring the astrophysical parameters $(M_{\rm min}, N_{\rm ion}, R_{\rm mfp})$.

Both samplers were coupled to the same Bayesian Neural Network (BNN) emulator for likelihood evaluation, ensuring that the comparison reflects differences in sampling strategy rather than model accuracy. The results show that both methods successfully recover the correct posterior distributions across multiple test cases, demonstrating the validity of the emulator-based inference framework. However, their computational efficiency differs significantly. The MH sampler, which relies on joint proposals in parameter space, exhibits lower acceptance rates and slower mixing, particularly in the presence of parameter correlations. In contrast, the Gibbs sampler performs conditional updates of individual parameters, allowing it to better adapt to the structure of the posterior distribution. This results in improved mixing behavior, higher acceptance rates, and larger effective sample sizes. These findings highlight the importance of structured sampling strategies in correlated inference problems and demonstrate that Gibbs sampling provides a more efficient approach in this setting.

The second part of the work extends the inference problem to a more realistic observational scenario by incorporating foreground contamination and incomplete data. In 21-cm observations, foregrounds introduce a structured region of missing modes in Fourier space, commonly referred to as the wedge, leading to incomplete coverage in the $(k_\perp, k_\parallel)$ plane. Instead of discarding these contaminated modes, we approach the problem from the perspective of incomplete data, treating the missing modes as latent variables to be inferred within a statistical framework.

To address this, a reconstruction-based approach was developed to recover the missing modes of the two-dimensional 21-cm power spectrum. This was implemented using a Gaussian Constrained Realization-inspired framework, which leverages the covariance structure of the signal to propagate information from observed modes into masked regions. In this formulation, the reconstruction is not a deterministic interpolation, but a probabilistic inference of the missing data conditioned on the observed modes.

The reconstruction was evaluated across multiple masking scenarios, ranging from optimistic to pessimistic foreground contamination. The results demonstrate that the reconstructed power spectra are consistent with the true signal within expected uncertainties. In particular, large-scale modes are recovered with high accuracy, reflecting the strong correlations present in the signal, while uncertainty increases in regions that are weakly constrained by the data. This behavior is consistent with the covariance-driven nature of the reconstruction, in which the effectiveness of information propagation depends on the strength and scale of correlations in Fourier space.

Taken together, these two components form a coherent framework for 21-cm inference. The first establishes efficient and reliable sampling strategies under ideal conditions, while the second addresses the practical challenges posed by incomplete and contaminated data. The natural next step is to combine these elements into a unified inference pipeline in which reconstruction and parameter estimation are performed jointly, allowing for a fully consistent treatment of uncertainty.

A key component required to complete this pipeline is the extension of the emulator from the one-dimensional (spherically averaged) power spectrum to the full two-dimensional power spectrum. This transition introduces additional complexity, as the emulator must capture the anisotropic structure of the signal in $(k_\perp, k_\parallel)$ space while operating in a higher-dimensional setting.

At the same time, it is important to consistently account for uncertainties associated with the Bayesian Neural Network (BNN) emulator, ensuring that model errors are properly propagated through the inference process. Work is currently in progress to develop such an emulator, enabling a fully consistent pipeline in which reconstruction and parameter estimation can be combined.

Although the present implementation operates in a reduced data space using the two-dimensional power spectrum, the methodology developed here is designed with scalability in mind. A natural extension of this work is to move beyond summary statistics and perform reconstruction at the level of the full three-dimensional HI signal. This would allow for a more complete and physically consistent treatment of the signal, including foregrounds, missing data, and parameter inference within a unified Bayesian framework.

More broadly, this work demonstrates the importance of combining statistical inference techniques with physically motivated models and machine learning tools. The integration of reconstruction methods, efficient sampling strategies, and emulator-based inference provides a promising pathway toward robust and scalable analysis pipelines for next-generation 21-cm experiments.

8 Future Work

The work presented in this thesis naturally leads to two key directions for future development: completing the current inference framework and extending it toward full signal reconstruction. These directions are closely connected, with the former providing a necessary intermediate step toward the latter.

8.1 Completion of Emulator-Based Inference

The immediate next step is to develop a robust emulator for the two-dimensional 21-cm power spectrum and complete the inference pipeline. While the reconstruction of missing modes has been successfully demonstrated, parameter inference on the reconstructed power spectrum remains an open task.

In the current implementation, the emulator is trained on the one-dimensional (spherically averaged) power spectrum. Extending this to the full two-dimensional power spectrum requires modeling the anisotropic structure of the signal in $(k_\perp, k_\parallel)$ space, significantly increasing the dimensionality of the problem. This demands careful design of the emulator architecture to ensure both accuracy and computational efficiency.

A Bayesian Neural Network (BNN) emulator will be developed to predict the two-dimensional power spectrum across the relevant parameter space, while also providing uncertainty estimates. A key requirement in this stage is the consistent propagation of emulator uncertainty into the likelihood evaluation, ensuring that model errors are not misinterpreted as physical signal variations.

Once developed, the emulator will be integrated with the reconstruction framework to enable full parameter inference on reconstructed data. This will allow for a quantitative assessment of how reconstruction affects parameter constraints under realistic observational conditions.

8.2 Beyond the Power Spectrum: Incorporating Higher-Order Statistics

While the power spectrum captures the two-point statistical properties of the signal, it is insensitive to non-Gaussian features that become increasingly important during the later stages of reionization. To address this limitation, a natural extension of this work is to incorporate higher-order statistics, such as the bispectrum.

The bispectrum provides access to phase correlations and non-Gaussian structure in the signal, offering complementary information to the power spectrum. Future work will focus on developing an emulator-based framework for the bispectrum, enabling joint inference using both two-point and three-point statistics.

A key challenge in this direction is the increased dimensionality and computational cost associated with the bispectrum. Efficient representations and emulator architectures will be required to make this approach tractable within an MCMC-based inference pipeline.

8.3 Towards Full Signal Reconstruction

The broader goal of this work is to move beyond summary statistics and perform reconstruction at the level of the full three-dimensional HI signal. In this framework, missing or contaminated modes would be treated as latent variables and jointly inferred along with astrophysical parameters.

Building on the current approach, a key planned direction is to incorporate morphology-aware priors derived from multi-scale representations of the signal. By decomposing the HI field into components that capture structure across different spatial scales, it becomes possible to characterize features such as ionized bubbles and their boundaries in a physically meaningful way. These representations can then be used to define region-dependent priors that reflect the local structure of the signal.

Such a framework would allow the reconstruction to adapt spatially, rather than relying on a single global statistical model. In particular, different regions of the data—such as ionized interiors, neutral regions, and boundary zones—can be treated with priors that are tailored to their underlying morphology. This provides a more flexible and physically motivated description of the signal, and offers a natural pathway toward full-field reconstruction in the presence of foreground contamination.

8.4 Application to Realistic Observations

An essential step toward practical application is testing the framework on realistic observational data. Upcoming experiments such as the Square Kilometre Array (SKA) will provide high-sensitivity measurements of the 21-cm signal, but will also introduce additional challenges.

These include instrumental noise, calibration uncertainties, beam effects, and more complex foreground contamination. Incorporating these effects into the reconstruction and inference framework will be necessary for realistic applications.

Simulated SKA-like datasets will be used as an intermediate step to validate the methodology. This will allow us to assess the robustness of the pipeline under realistic noise and systematics before applying it to actual observations.

8.5 Outlook

The current work establishes the foundation for reconstruction-based inference in 21-cm cosmology by combining statistical reconstruction, efficient sampling strategies, and emulator-based modeling.

Future developments aim to integrate these components into a unified framework capable of handling incomplete data, foreground contamination, and parameter estimation in a fully consistent manner. The inclusion of higher-order statistics, more expressive priors, and realistic observational effects will further enhance the physical fidelity of the analysis.

Together, these efforts will contribute toward the development of scalable and robust inference pipelines for next-generation 21-cm experiments.

Bibliography

Rennan Barkana and Abraham Loeb. In the beginning: The first sources of light and the reionization of the universe. *Physics Reports*, 349(2):125–238, 2001. doi: 10.1016/S0370-1573(00)00064-3.

Nichole Barry, Bryna Hazelton, Ian Sullivan, and Miguel F. Morales. Calibration requirements for detecting the 21 cm signal from the epoch of reionization. *The Astrophysical Journal*, 884(1):1–16, 2019. doi: 10.3847/1538-4357/ab3e57.

Michael Betancourt. A conceptual introduction to Hamiltonian Monte Carlo. *arXiv preprint*, arXiv:1701.02434, 2017. doi: 10.48550/arXiv.1701.02434.

Emma Chapman, Filipe B. Abdalla, Geraint Harker, Vibor Jelić, Panagiotis Labropoulos, Saleem Zaroubi, Michiel A. Brentjens, and Léon V. E. Koopmans. Foreground removal using FastICA: a showcase of LOFAR-EoR. *Monthly Notices of the Royal Astronomical Society*, 429(2):165–176, 2013. doi: 10.1093/mnras/stu2601.

Wenlin Chen et al. Diffusive Gibbs sampling. *arXiv preprint*, arXiv:2402.03008, 2024. doi: 10.48550/arXiv.2402.03008. Accepted at ICML 2024.

Xiaohui Fan, Chris L. Carilli, and Brian Keating. Observational constraints on cosmic reionization. *Annual Review of Astronomy and Astrophysics*, 44:415–462, 2006. doi: 10.1146/annurev.astro.44.051905.092514.

Daniel Foreman-Mackey, David W. Hogg, Dustin Lang, and Jonathan Goodman. EMCEE: The MCMC hammer. *Publications of the Astronomical Society of the Pacific*, 125(925):306–312, 2013. doi: 10.1086/670067.

Steven R. Furlanetto. The 21-cm line as a probe of reionization. In *The First Galaxies*, volume 43 of *Saas-Fee Advanced Course*, pages 247–356. Springer, Berlin, 2016. doi: 10.1007/978-3-319-15504-2_5.

Steven R. Furlanetto, S. Peng Oh, and Frank H. Briggs. Cosmology at low frequencies: The 21 cm transition and the high-redshift universe. *Physics Reports*, 433(4):181–301, 2006. doi: 10.1016/j.physrep.2006.08.002.

Andrew Gelman, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin. *Bayesian Data Analysis*. CRC Press, Boca Raton, FL, 3rd edition, 2013.

Walter R. Gilks, Nicky G. Best, and Kee K. C. Tan. Adaptive rejection Metropolis sampling within Gibbs sampling. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 44(4):455–472, 1995. doi: 10.2307/2986138.

Bradley Greig and Andrei Mesinger. Simultaneously constraining the astrophysics of reionisation and the epoch of heating with 21CMMC. *Monthly Notices of the Royal Astronomical Society*, 472(3):2651–2669, 2017. doi: 10.1093/mnras/stx2118.

W. Keith Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, 1970. doi: 10.1093/biomet/57.1.97.

Yehuda Hoffman and Eldad N. Ribak. Constrained realizations of Gaussian fields: A simple algorithm. *The Astrophysical Journal Letters*, 380:L5–L8, 1991. doi: 10.1086/186160.

Ilian T. Iliev, Mario G. Santos, Andrei Mesinger, Suman Majumdar, and Garrelt Mellema. Epoch of reionization modelling and simulations for SKA. *Proceedings of Science*, AASKA14:007, 2015. doi: 10.48550/arXiv.1501.04213.

Daniel C. Jacobs, Bryna J. Hazelton, Cathryn M. Trott, Joshua S. Dillon, Bart Pindor, Ian S. Sullivan, Jonathan C. Pober, Nichole Barry, Adam P. Beardsley, Gianni Bernardi, Judd D. Bowman, Frank Briggs, et al. The Murchison Widefield Array 21 cm power spectrum analysis methodology. *The Astrophysical Journal*, 825(2):114, 2016. doi: 10.3847/0004-637X/825/2/114.

Jens Jasche and Benjamin D. Wandelt. Bayesian physical reconstruction of initial conditions from large-scale structure surveys. *Monthly Notices of the Royal Astronomical Society*, 432(2):894–913, 2013. doi: 10.1093/mnras/stt449.

Alicia A. Johnson, Miles Q. Ott, and Mine Dogucu. *Bayes Rules!: An Introduction to Applied Bayesian Modeling*. Chapman & Hall/CRC, Boca Raton, FL, 2022. doi: 10.1201/9780429288340.

Léon V. E. Koopmans, Jonathan Pritchard, Garrelt Mellema, Filipe Abdalla, et al. The cosmic dawn and epoch of reionization with the square kilometre array. *Proceedings of Science*, AASKA14:001, 2015. doi: 10.22323/1.215.0001.

Adrian Liu and Max Tegmark. A method for 21 cm power spectrum estimation in the presence of foregrounds. *Physical Review D*, 83(10):103006, 2011. doi: 10.1103/PhysRevD.83.103006.

Abraham Loeb and Steven R. Furlanetto. *The First Galaxies in the Universe*. Princeton University Press, Princeton, NJ, 2013.

Yashrajsinh Mahida, Sanjay Kumar Yadav, Suman Majumdar, Leon Noble, Chandra Shekhar Murmu, Saswata Dasgupta, Sohini Dutta, Himanshu Tiwari, and Abinash Kumar Shaw. From ANN to BNN: Inferring reionization parameters using uncertainty-aware emulators of 21-cm summaries. *arXiv preprint*, arXiv:2508.13261, 2025. doi: 10.48550/arXiv.2508.13261.

Suman Majumdar, Tirthankar Roy Choudhury, and Somnath Bharadwaj. The bispectrum as a probe of reionization. *Monthly Notices of the Royal Astronomical Society*, 476(4):4007–4026, 2018. doi: 10.1093/mnras/sty511.

Matthew McQuinn. The evolution of the intergalactic medium. *Annual Review of Astronomy and Astrophysics*, 54:313–362, 2016. doi: 10.1146/annurev-astro-082214-122355.

Garrelt Mellema, Léon V. E. Koopmans, Himanshu Shukla, et al. Reionization and the cosmic dawn with the square kilometre array. *Experimental Astronomy*, 36:235–318, 2013. doi: 10.1007/s10686-013-9334-5.

Andrei Mesinger, Andrea Ferrara, Bradley Greig, Ilian Iliev, Garrelt Mellema, Jonathan Pritchard, and Mario G. Santos. Constraining the astrophysics of the cosmic dawn and the epoch of reionization with the SKA. *Proceedings of Science*, AASKA14:010, 2015. doi: 10.48550/arXiv.1501.04106.

Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 1953. doi: 10.1063/1.1699114.

Radford M. Neal. MCMC using Hamiltonian dynamics. In Steve Brooks, Andrew Gelman, Galin L. Jones, and Xiao-Li Meng, editors, *Handbook of Markov Chain Monte Carlo*, pages 113–162. Chapman & Hall/CRC, Boca Raton, FL, 2011.

Planck Collaboration. Planck 2018 results. VI. cosmological parameters. *Astronomy & Astrophysics*, 641:A6, 2020. doi: 10.1051/0004-6361/201833910.

Jonathan R. Pritchard and Abraham Loeb. Evolution of the global 21-cm signal before reionization. *Physical Review D*, 82(2):023006, 2010. doi: 10.1103/PhysRevD.82.023006.

Jonathan R. Pritchard and Abraham Loeb. 21-cm cosmology in the 21st century. *Reports on Progress in Physics*, 75(8):086901, 2012. doi: 10.1088/0034-4885/75/8/086901.

Mario G. Santos, Asantha Cooray, and Lloyd Knox. Multifrequency analysis of 21 cm fluctuations from the era of reionization. *The Astrophysical Journal*, 625(2):575–587, 2005. doi: 10.1086/429857.

Mario G. Santos, Luis Ferramacho, Marta B. Silva, Alexandre Amblard, and Asantha Cooray. Fast large-volume simulations of the 21-cm signal from the reionization and pre-reionization epochs. *Monthly Notices of the Royal Astronomical Society*, 406(4):2421–2432, 2010. doi: 10.1111/j.1365-2966.2010.16898.x.

Benoît Semelin and Ilian T. Iliev. The physics of reionization: processes relevant for SKA observations. *Proceedings of Science*, AASKA14:019, 2015. doi: 10.48550/arXiv.1501.04459.

Roberto Trotta. Bayes in the sky: Bayesian inference and model selection in cosmology. *Contemporary Physics*, 49(2):71–104, 2008. doi: 10.1080/00107510802066753.

Benjamin D. Wandelt, David L. Larson, and Arjun Lakshminarayanan. Global, exact cosmic microwave background data analysis using Gibbs sampling. *Physical Review D*, 70(8):083511, 2004. doi: 10.1103/PhysRevD.70.083511.

Laura Wolz, Filipe B. Abdalla, David Alonso, Chris Blake, Philip Bull, et al. Foreground subtraction in intensity mapping with the SKA. *Proceedings of Science*, AASKA14:035, 2015. doi: 10.48550/arXiv.1501.03823.