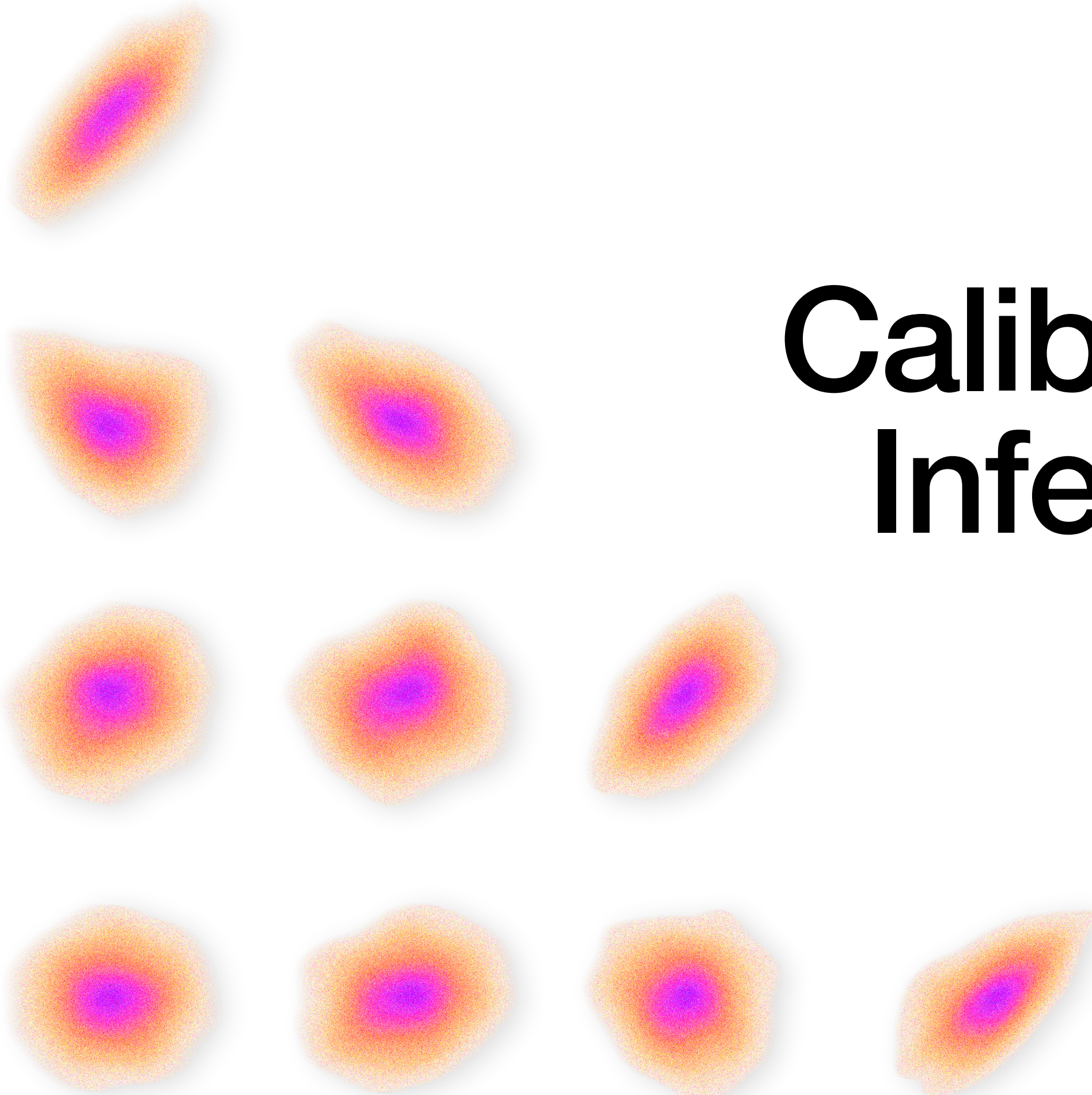



Calibrated Simulation-Based Inference for the Lya Forest

Using Transformer Ratio Estimators

About Me



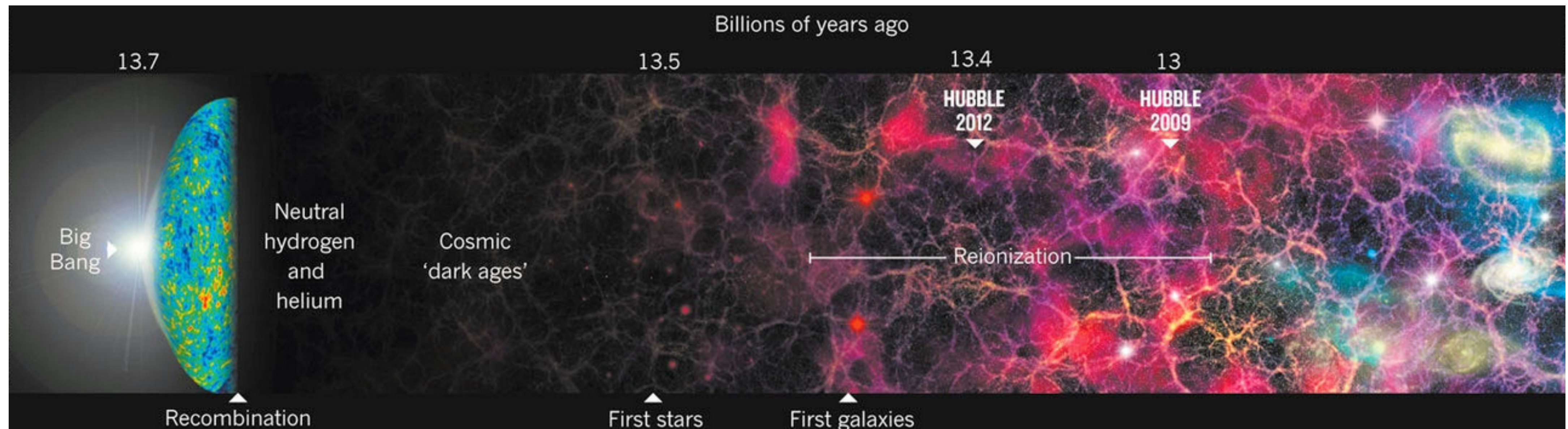
PhD Candidate at UC Santa Barbara
Advisor: Prof. Joseph Hennawi
Research: Parameter Estimation for the EoR



Intern at Lawrence Berkeley National Lab
Supervisor: Dr. Zarija Lukić
Research: Foundational ML Model for the IGM

The Epoch of Reionization (EoR)

The Evolution of the Intergalactic Medium (IGM)

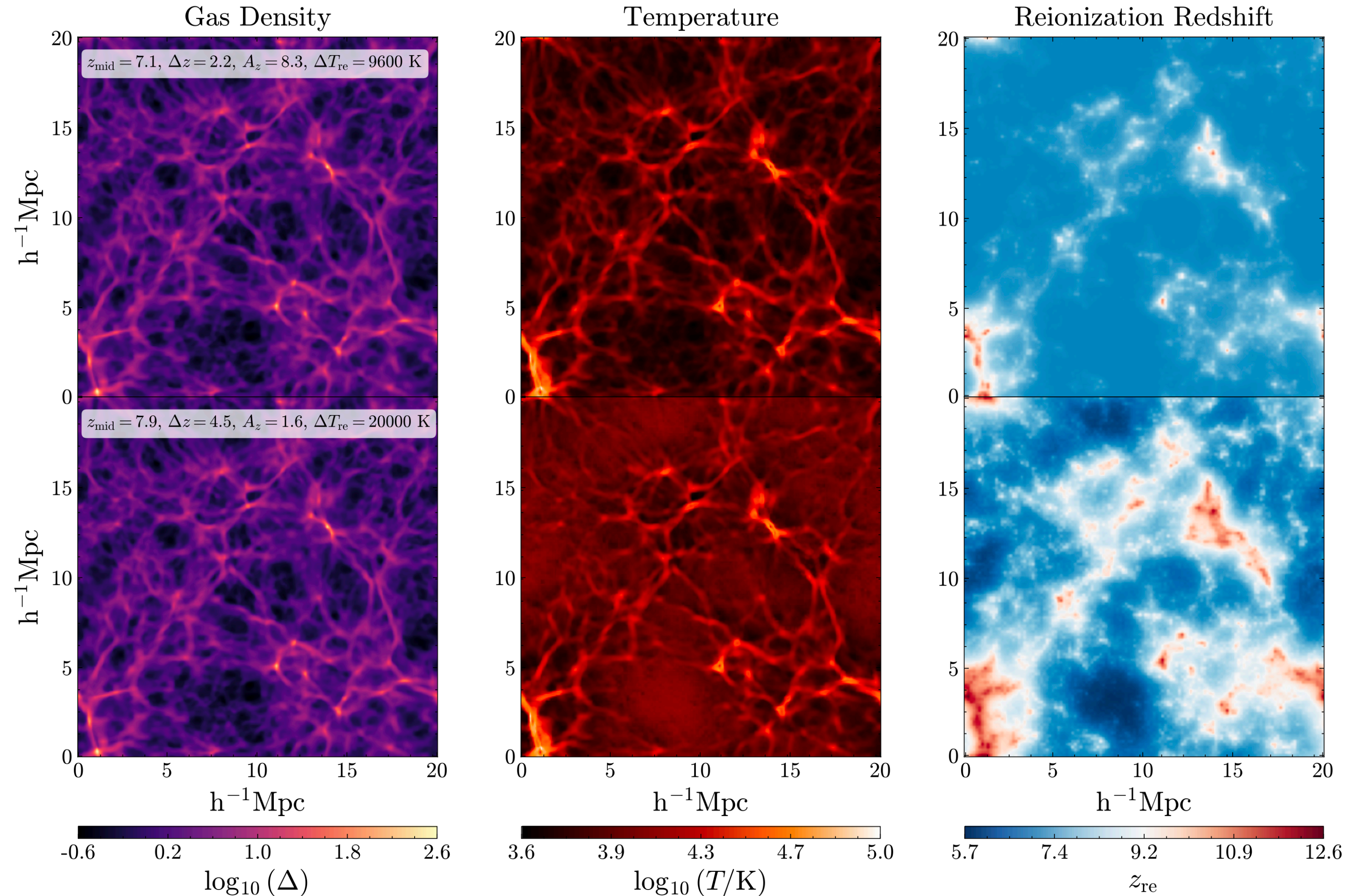


* Adapted from **Brant E. Robertson, Richard S. Ellis, James S. Dunlop, Ross J. McLure, and Daniel P. Stark**, Early star-forming galaxies and the reionization of the Universe, *Nature* 468, 49 (2010).

Modeling the Ly α forest

During the EoR

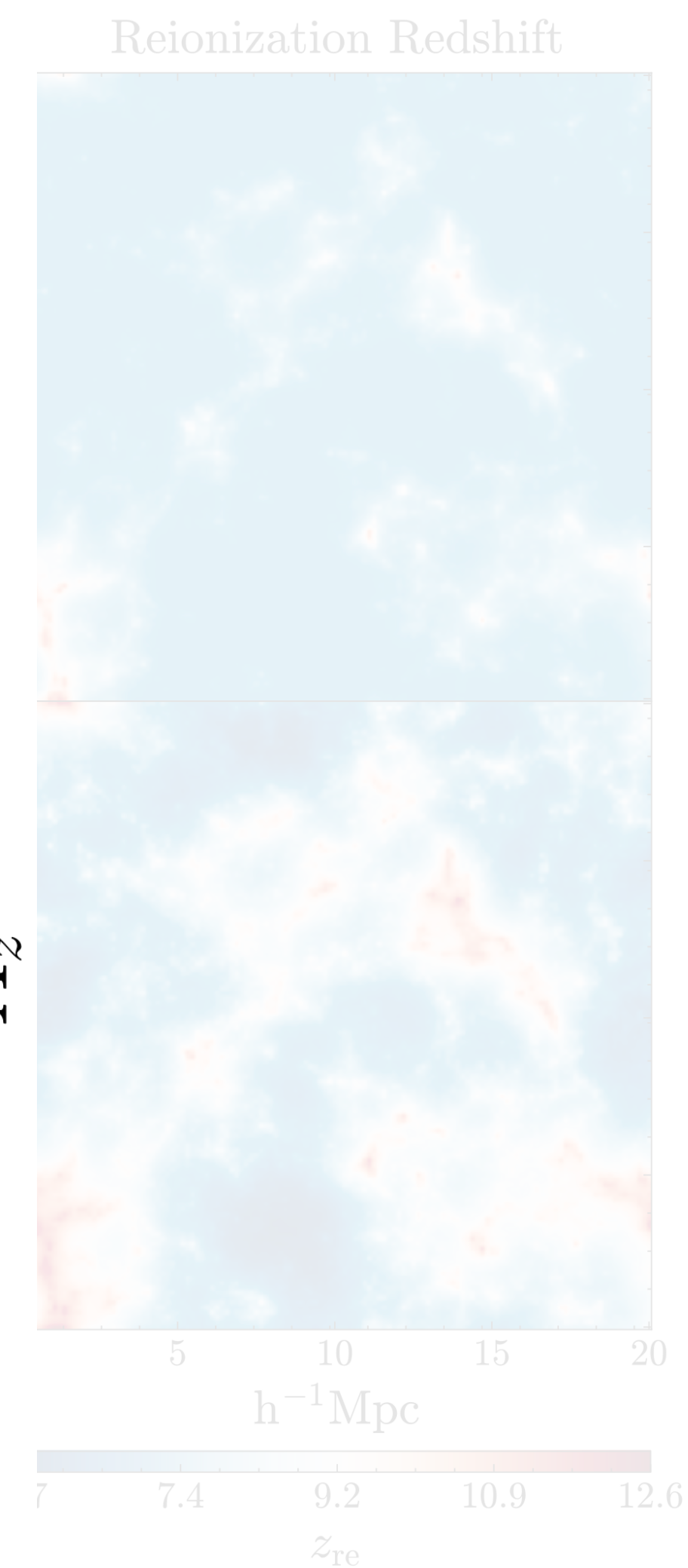
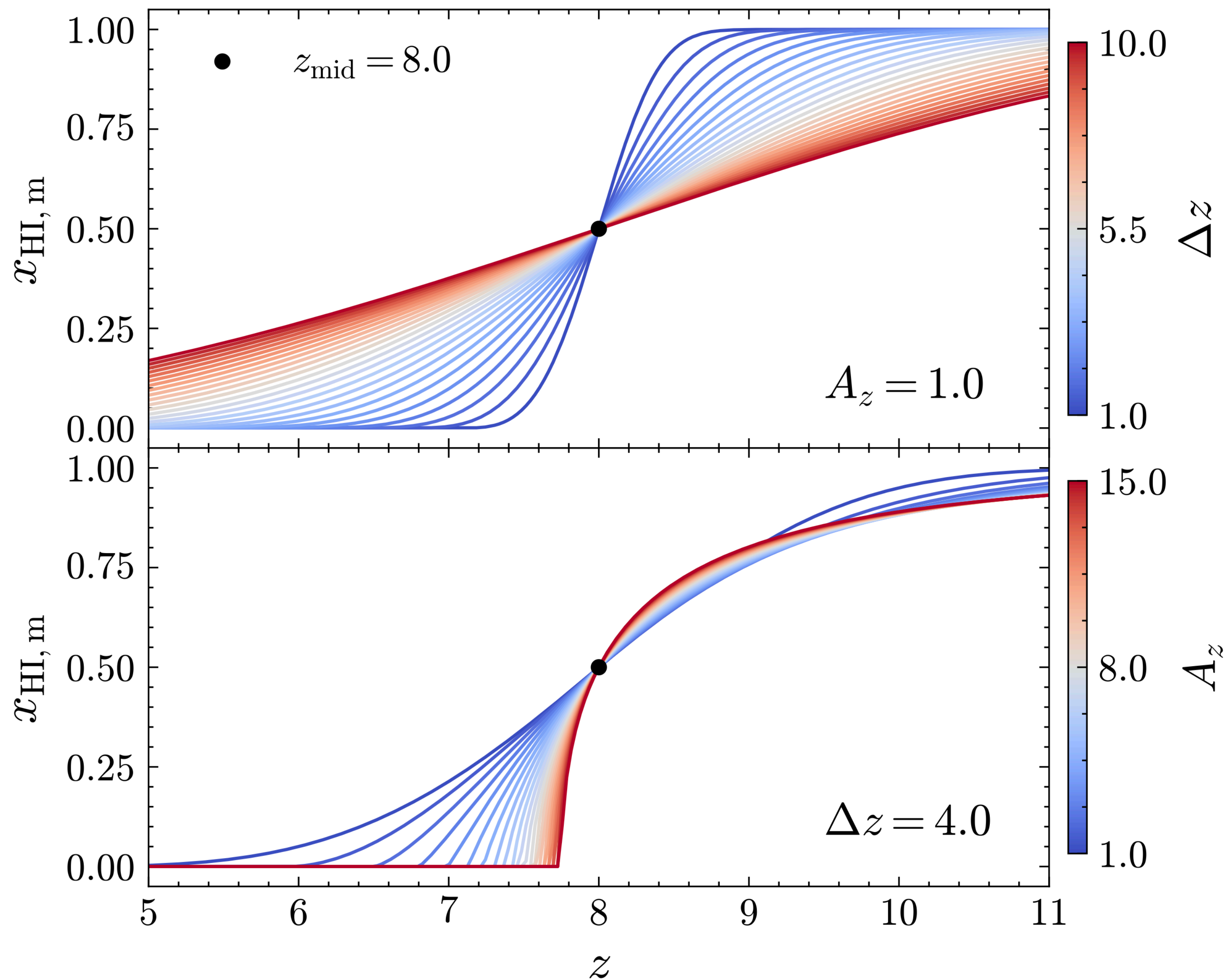
- Abundance Matching Box for the EoR (AMBER; Trac et al. 2022) creates user-defined reionization fields.
- Nyx (Almgren et al. 2013; Lukić et al. 2015) to model the IGM.
- Simulation outputs can be used to model observables.



Modelin

During the Ec

- Abundance $\mathcal{N}$ for the EoR ($\mathcal{N}$ al. 2022) crea defined reioni
- Nyx (Almgren Lukić et al. 20 the IGM.
- Simulation ou used to mode

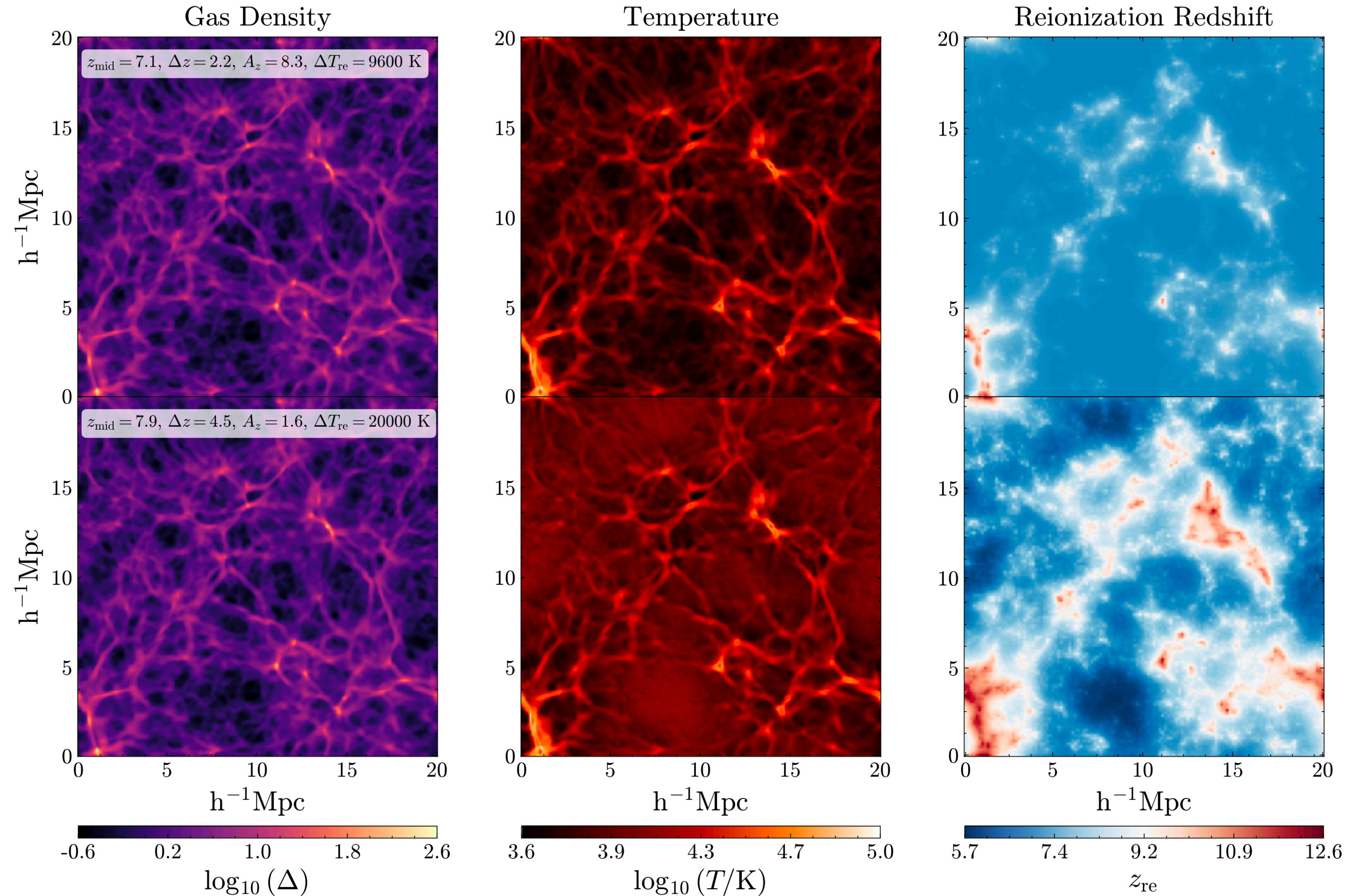


Modeling the Ly α forest

During the EoR

- Abundance Matching Box for the EoR (AMBER; Trac et al. 2022) creates user-defined reionization fields.
- Nyx (Almgren et al. 2013; Lukić et al. 2015) to model the IGM.
- Simulation outputs can be used to model observables.

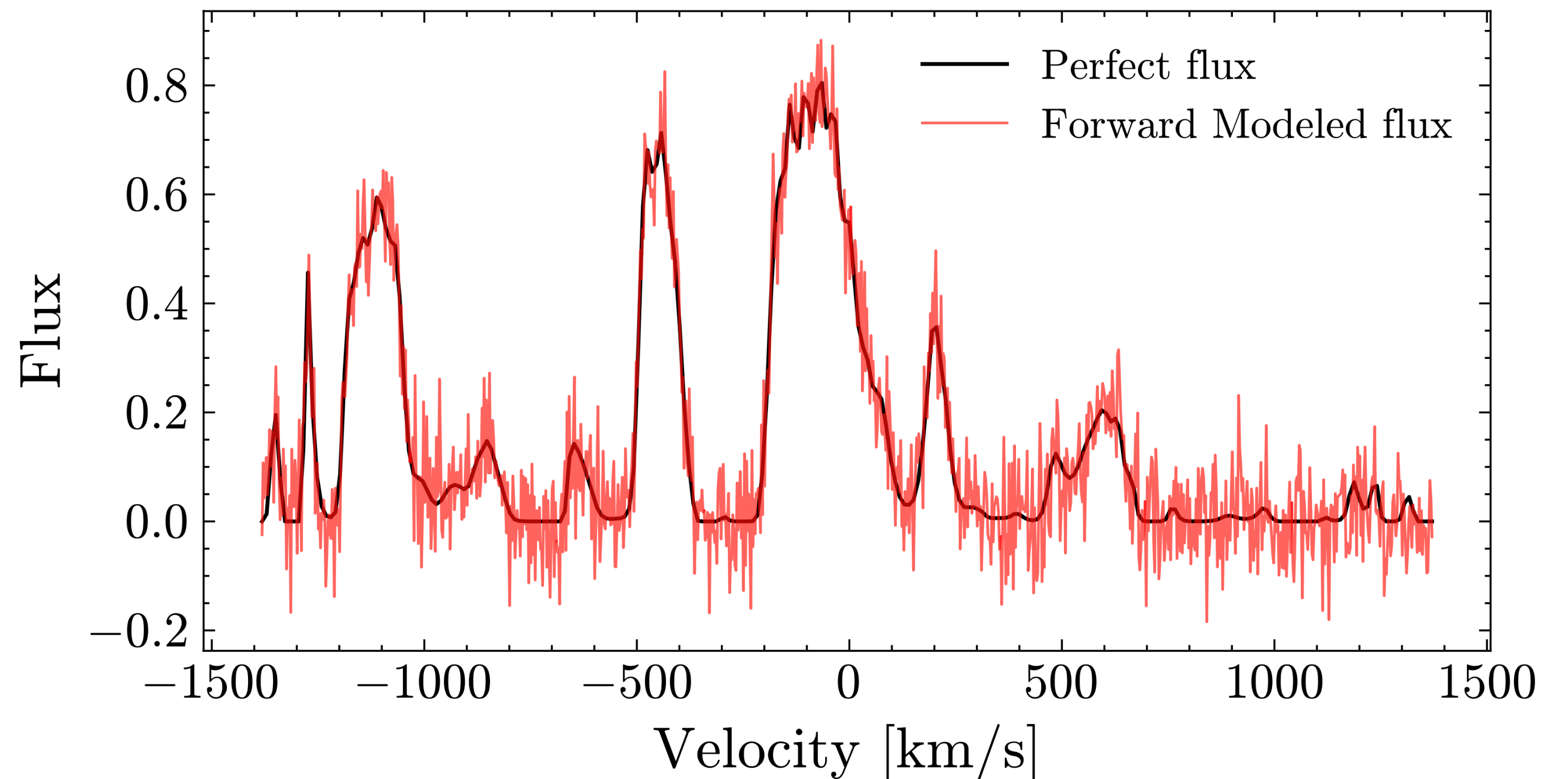
$$\theta = \{z_{\text{mid}}, \Delta z, A_z, \Delta T_{\text{re}}, \langle F \rangle\}$$



Parameter Estimation

With a Gaussian Likelihood

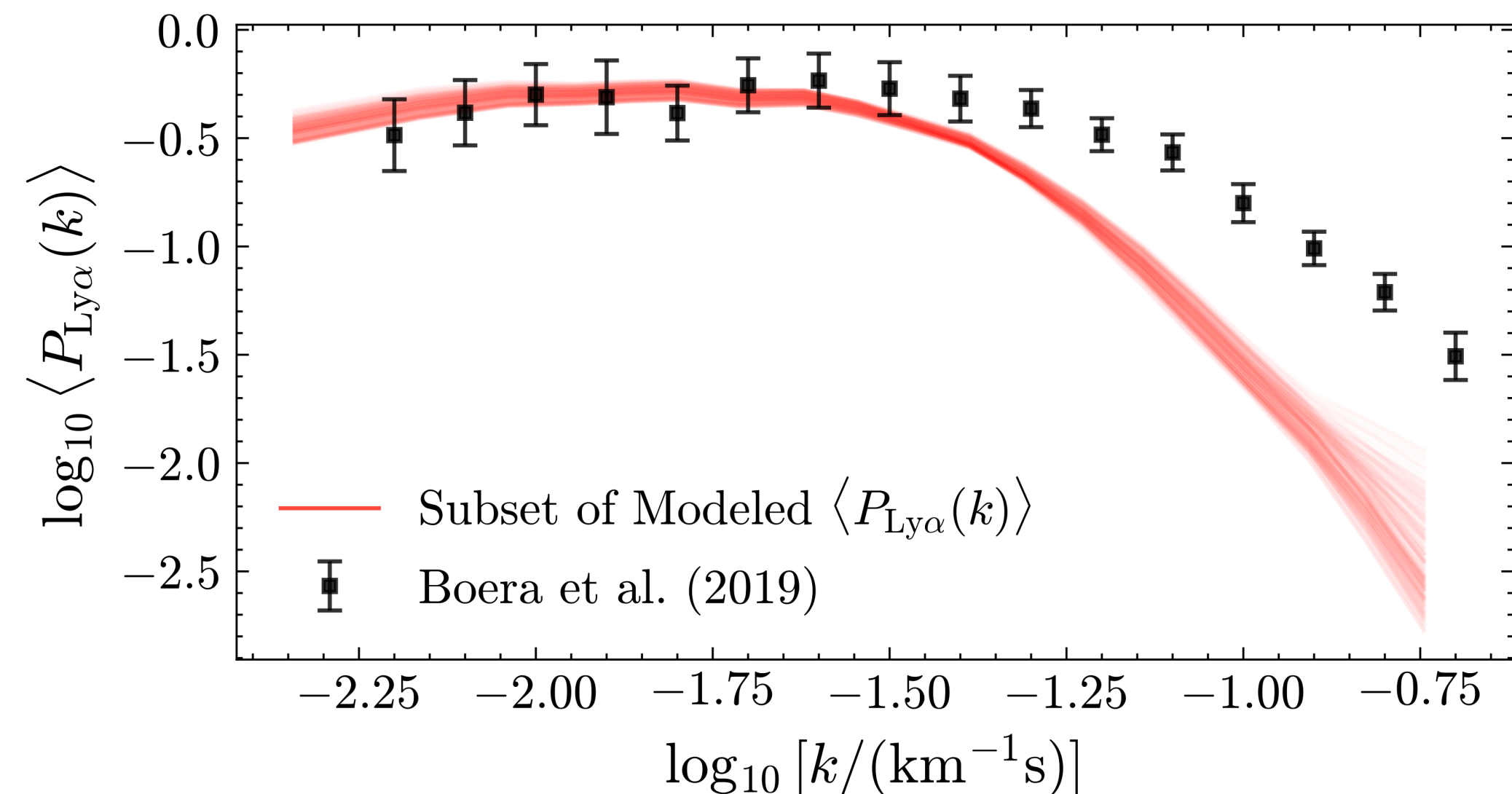
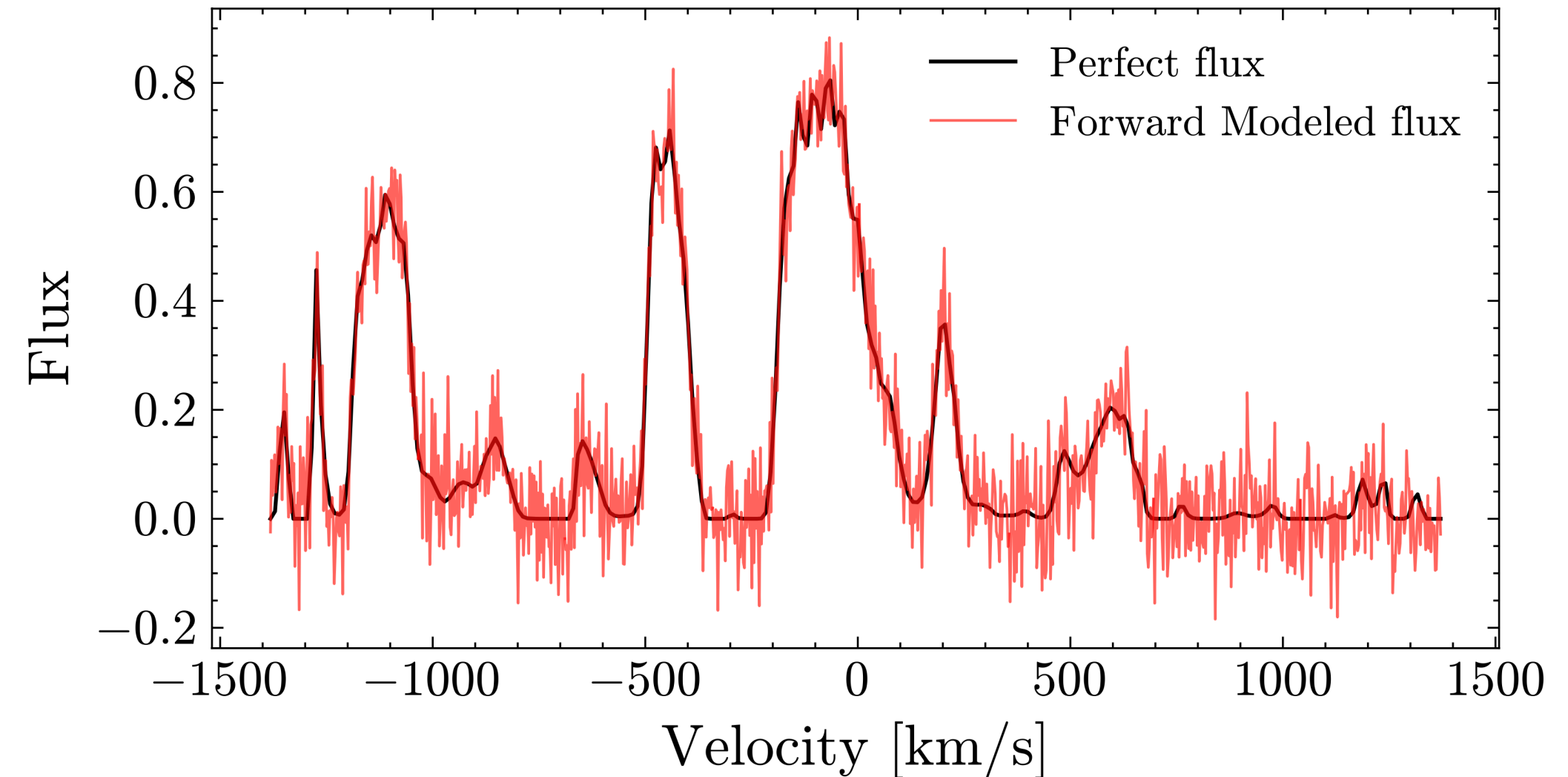
- Bayesian framework with a summary statistic (namely, the 1D power spectrum).
- Gaussian likelihood.
- Forward-modeled flux skewers are used to compute mean power spectra and covariance matrices.



Parameter Estimation

The Mean Power Spectrum

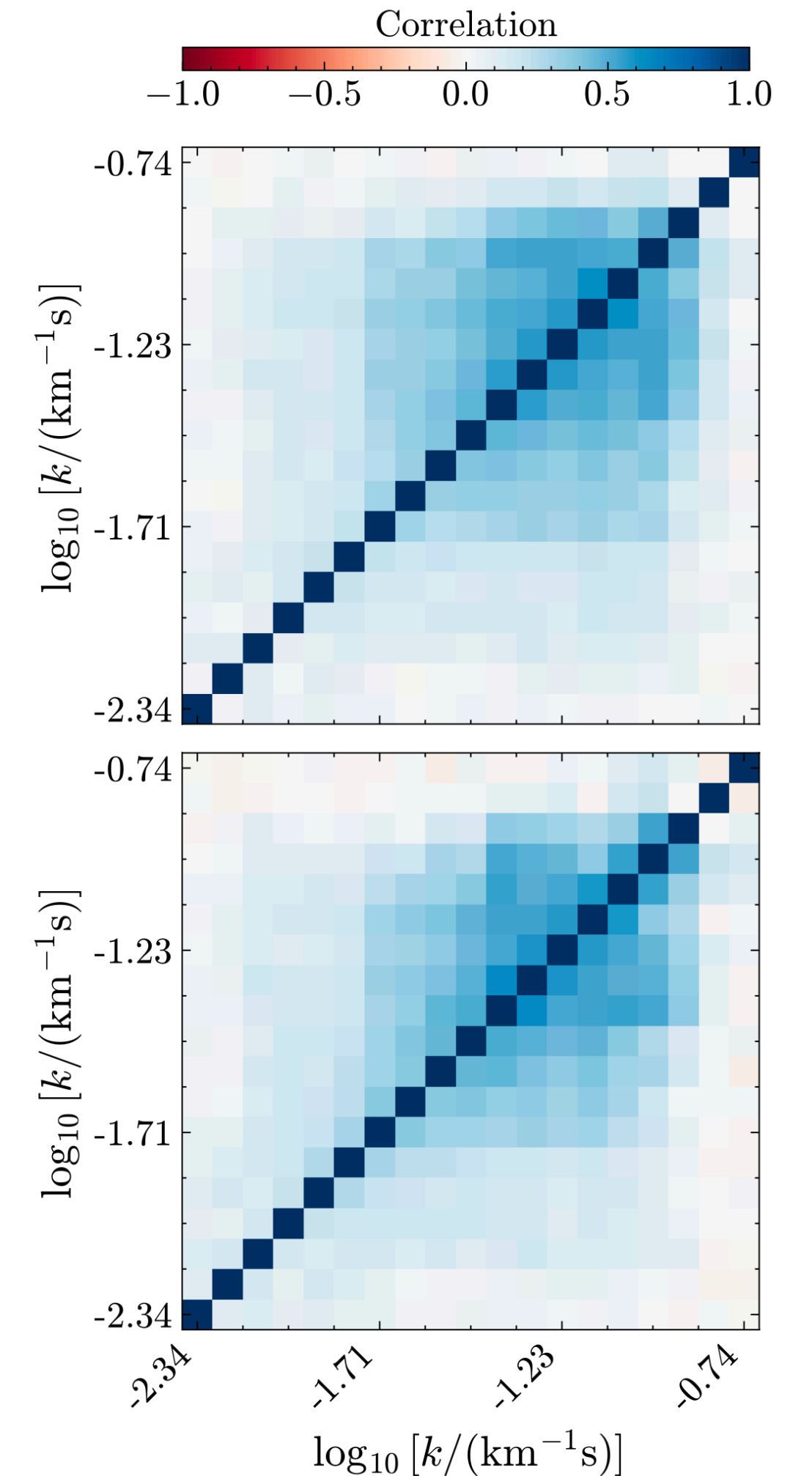
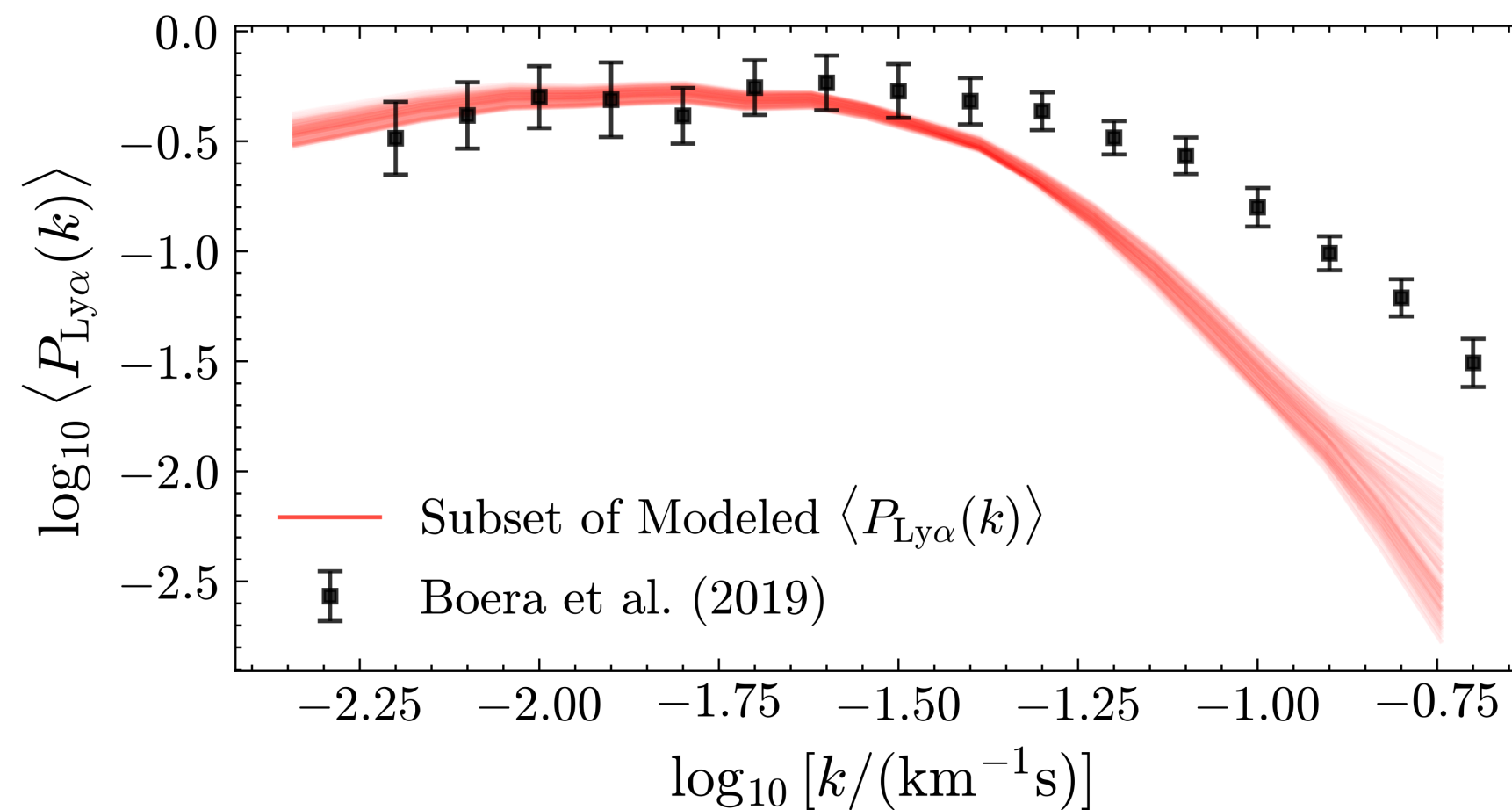
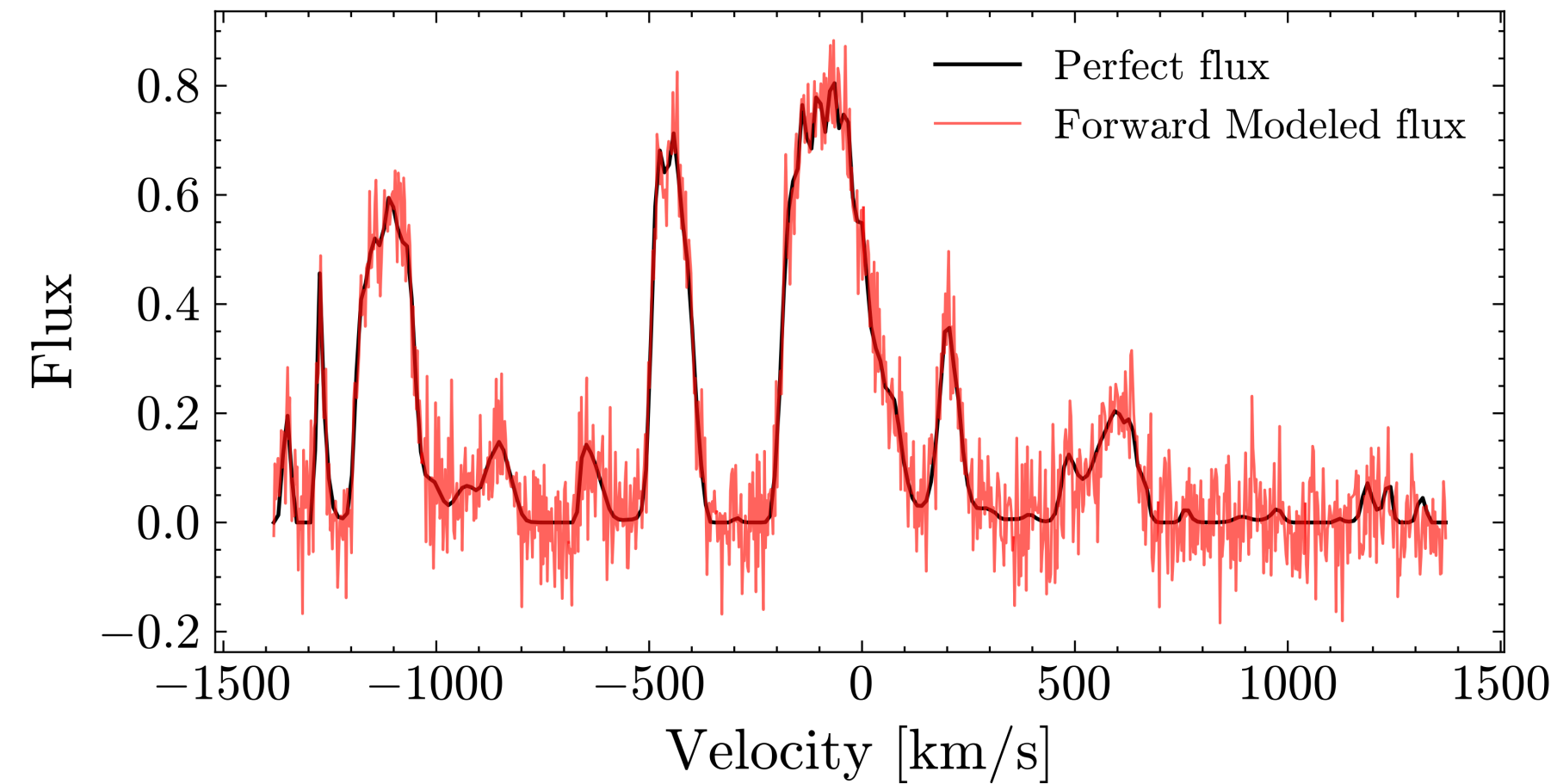
- Bayesian framework with a summary statistic (namely, the 1D power spectrum).
- Gaussian likelihood.
- Forward-modeled flux skewers are used to compute mean power spectra and covariance matrices.



Parameter Estimation

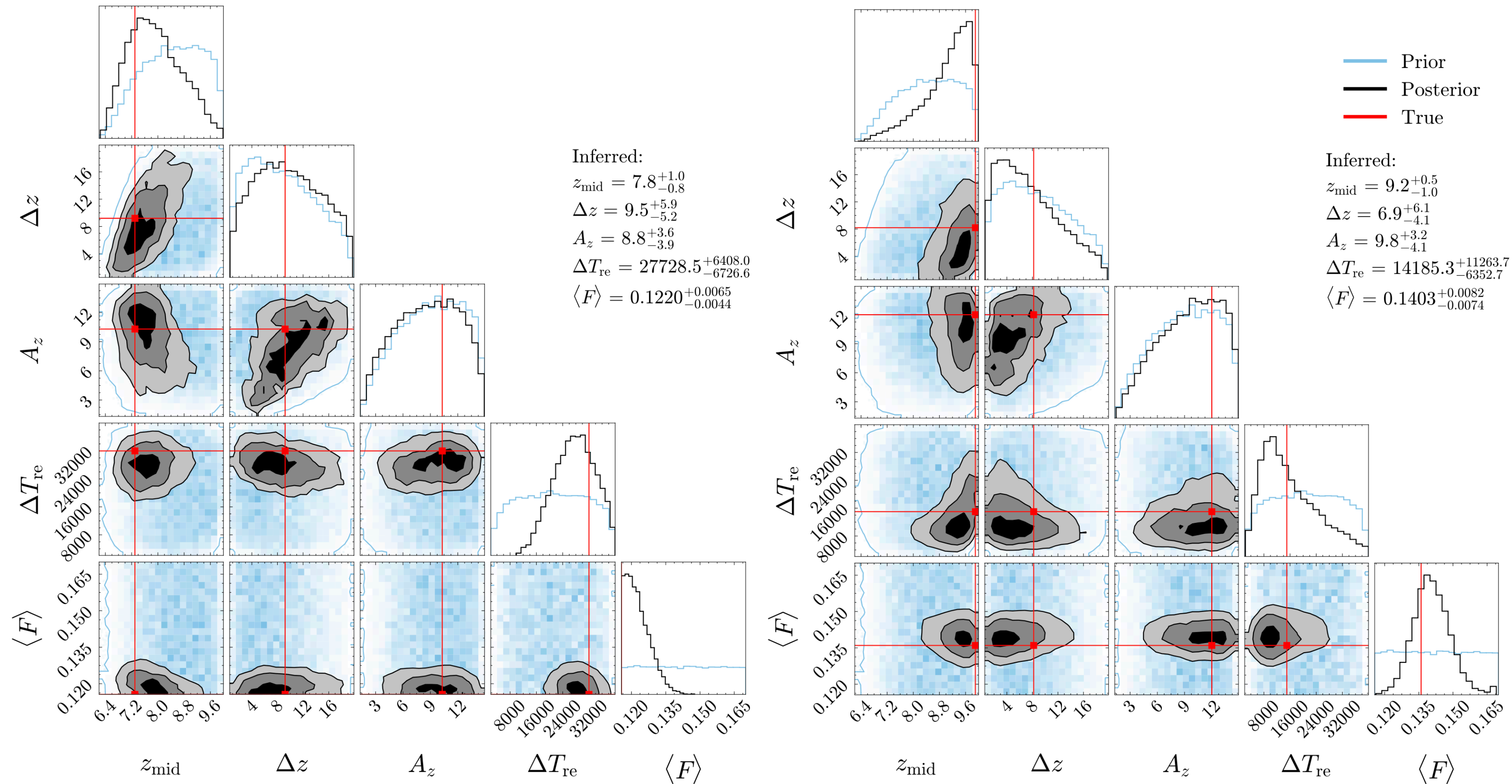
The Covariance Matrix

- Bayesian framework with a summary statistic (namely, the 1D power spectrum).
- Gaussian likelihood.
- Forward-modeled flux skewers are used to compute mean power spectra and covariance matrices.



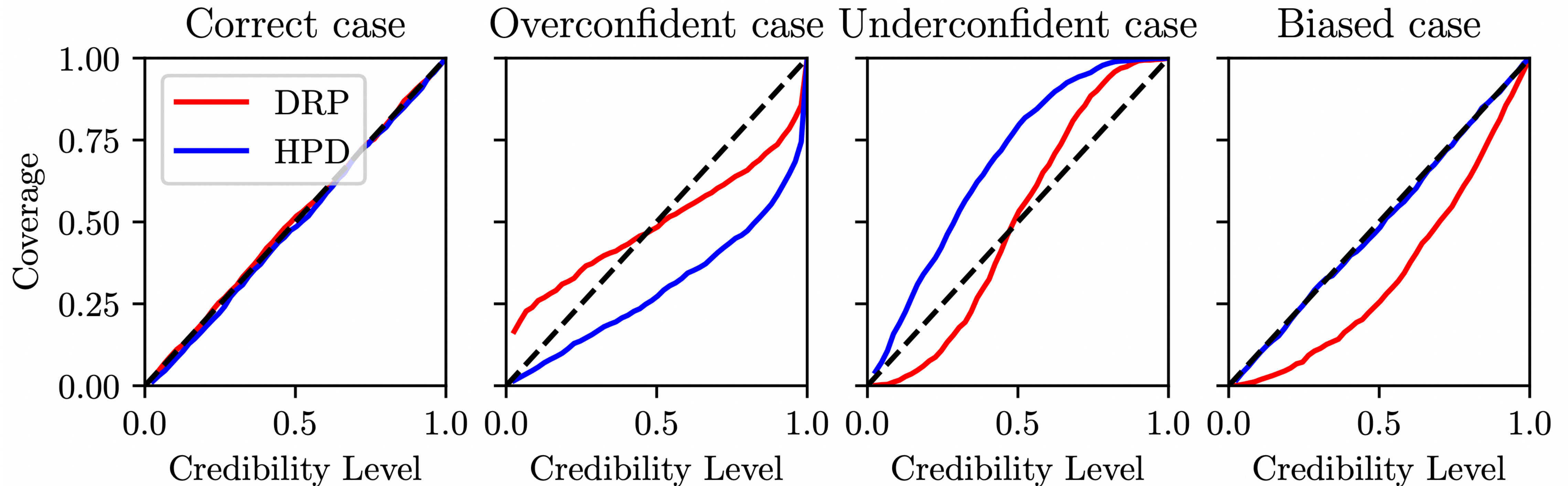
Inference on Synthetic Observations

With the Gaussian Likelihood Assumption (González-Hernández et al. 2025a)



Coverage Tests

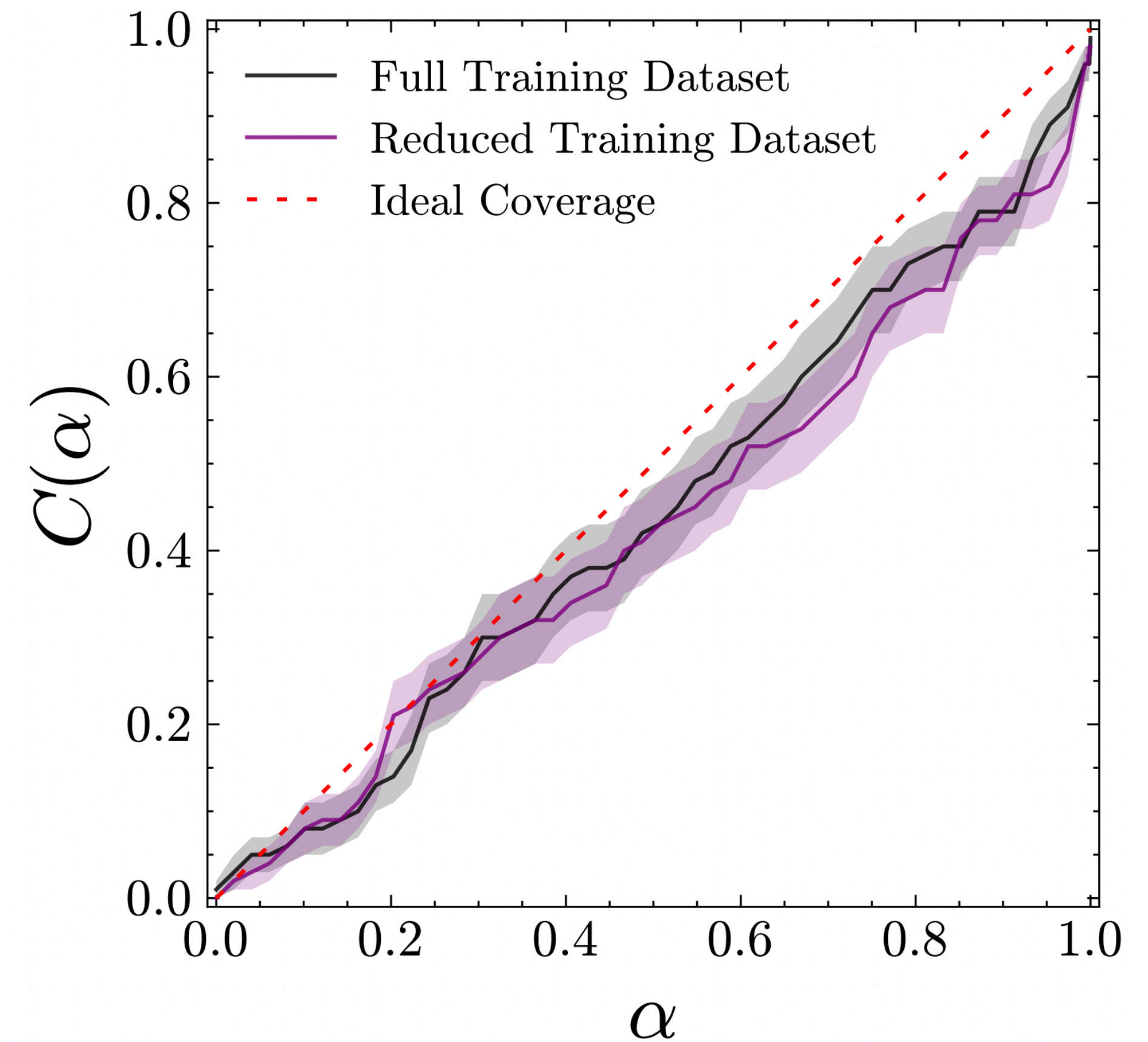
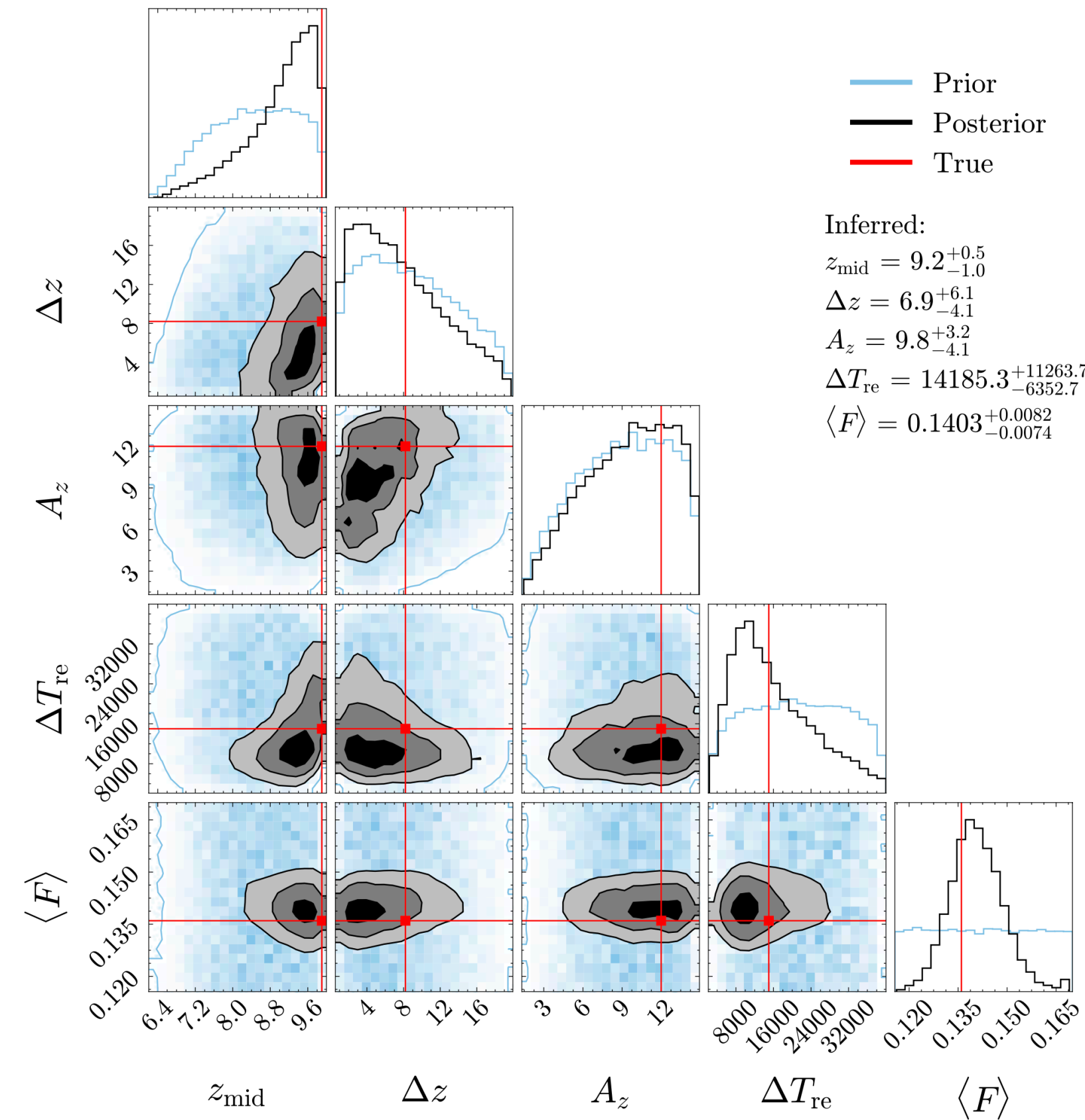
To Test the Posterior Calibration



* Adapted from **Lemos, P., Coogan, A., Hezaveh, Y., & Perreault-Levasseur, L.** Sampling-Based Accuracy Testing of Posterior Estimators for General Inference. *ICML, 2023*.

Uncalibrated Posterior Estimates

Using a Coverage Test (González-Hernández et al. 2025a)



Using Simulation-Based Inference (SBI)

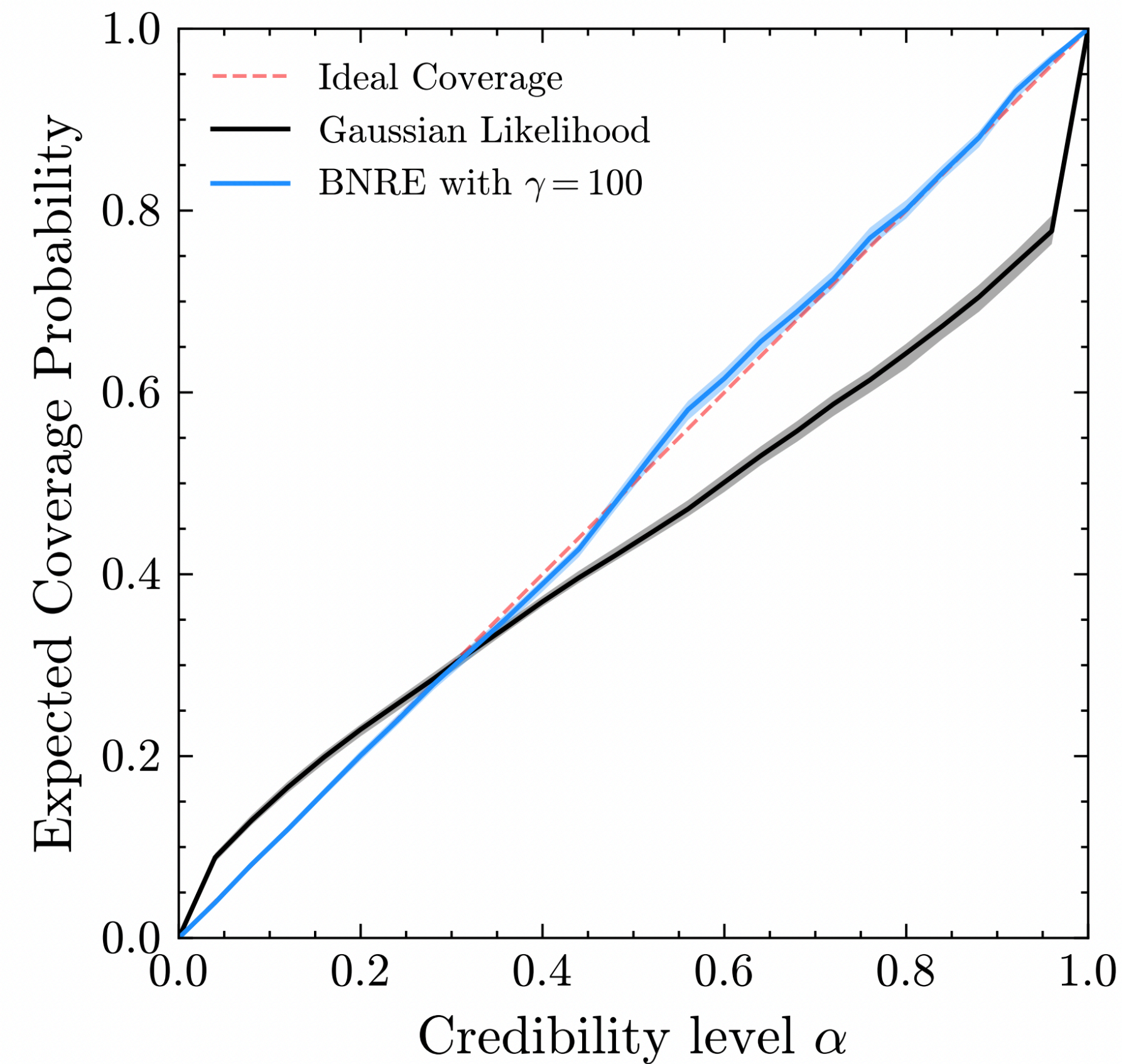
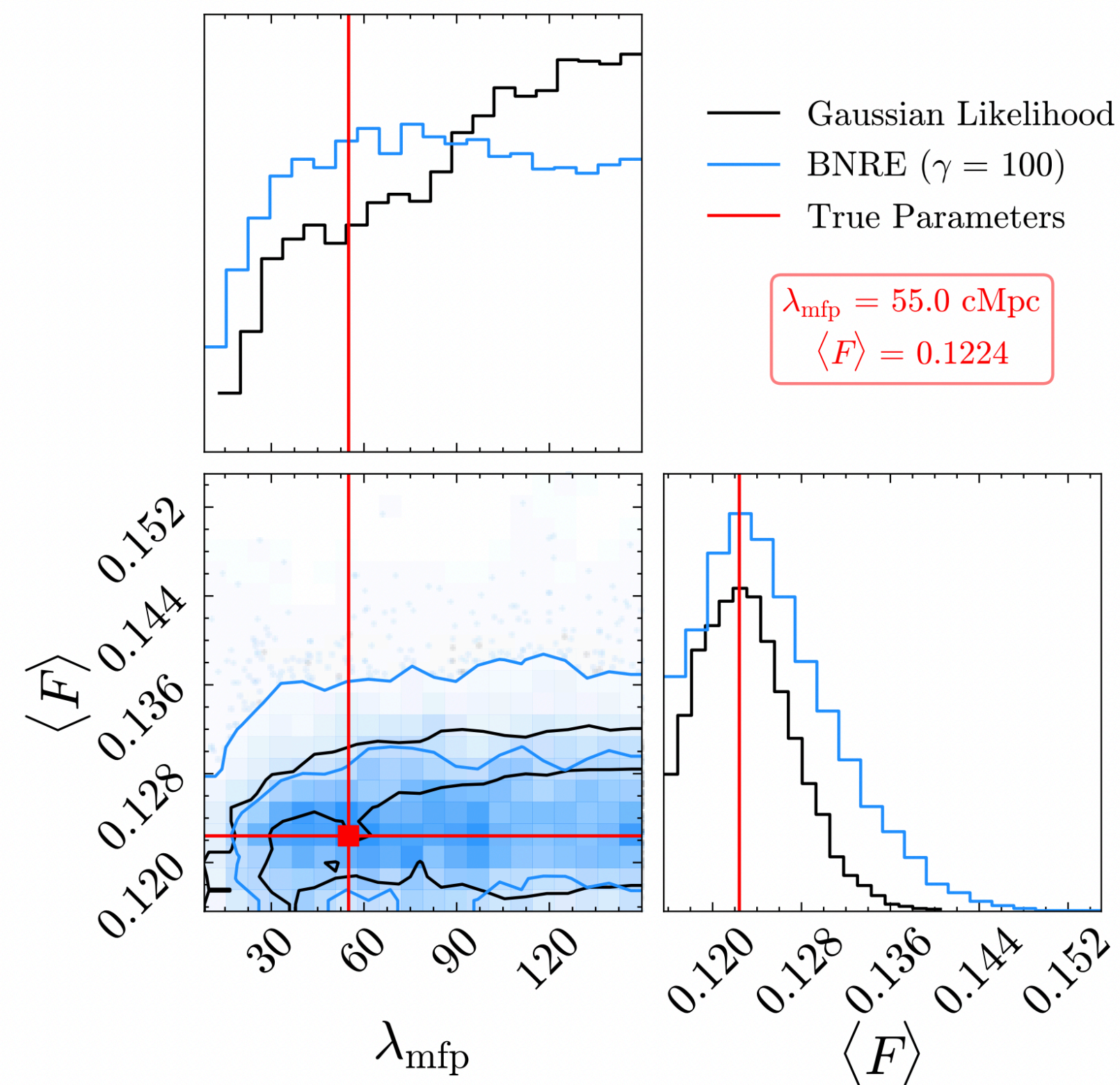
To address this challenge

- Balanced Neural Ratio Estimation (BNRE; Delaunoy et al. 2022) is cheap, and it specifically allows you to tune the calibration, allowing for rapid iteration.

Neural Posterior Estimation (NPE)	- Estimates the posterior directly - Does not require a sampling algorithm	- Density Estimator
Neural Likelihood Estimation (NLE)	- Estimates the likelihood - Requires a sampling algorithm	- Density Estimator
Neural Ratio Estimation (NRE)	- Estimates the likelihood ratio (likelihood over evidence) - Requires a sampling algorithm	- Classifier

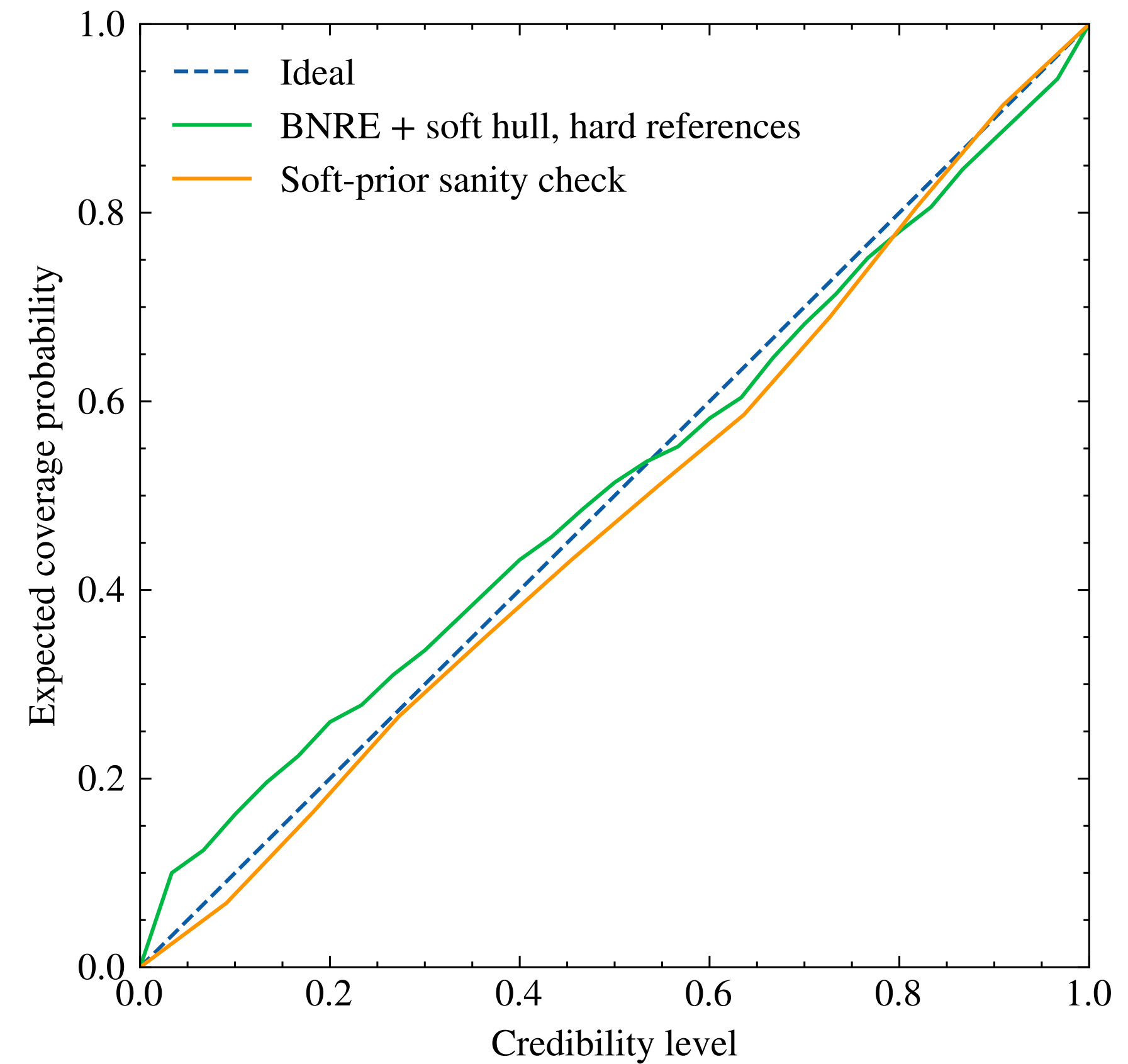
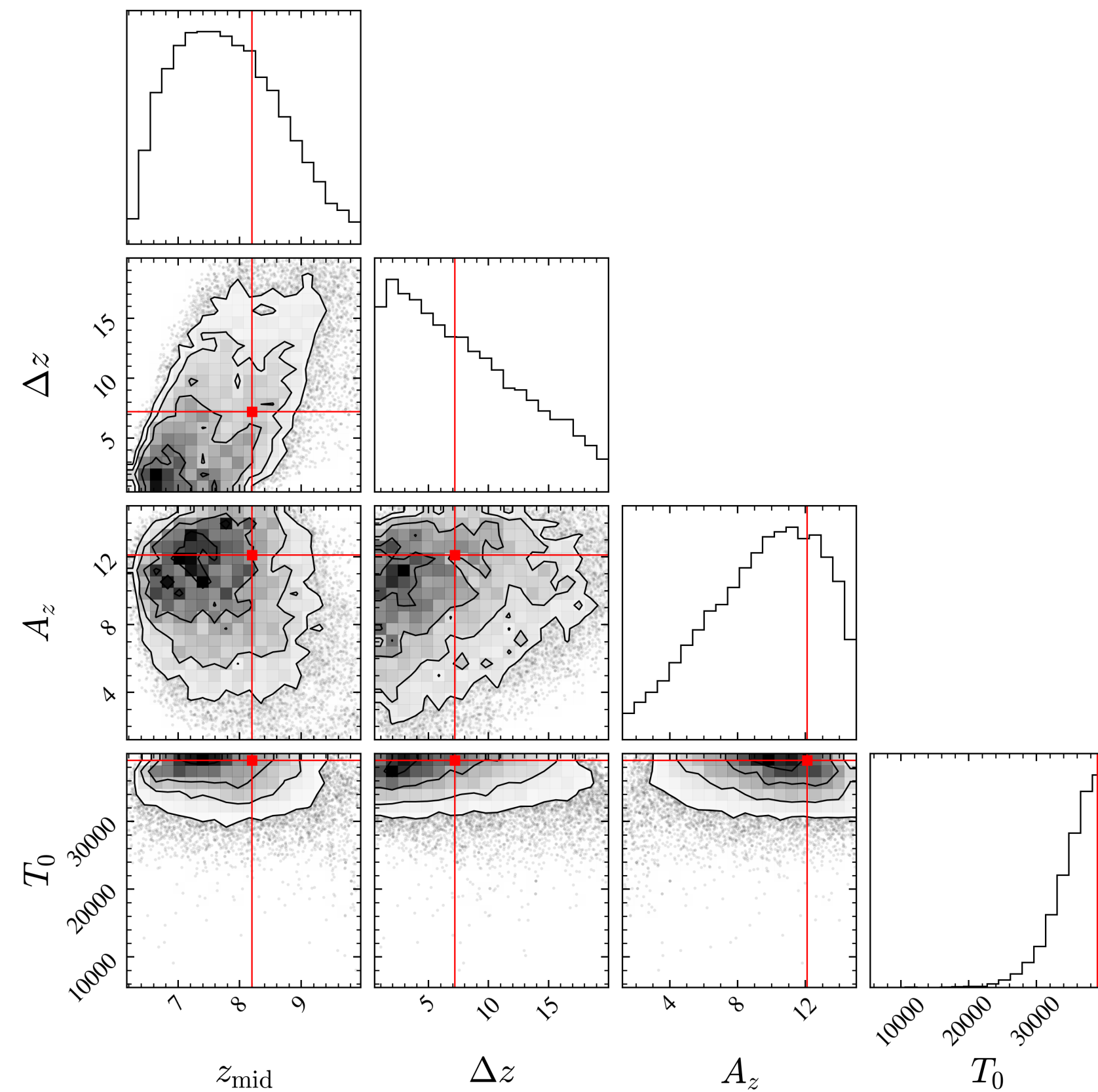
Using Balanced Neural Ratio Estimation

Solves the problem (González-Hernández et al. 2025b)



Using Balanced Neural Ratio Estimation

Solves the problem (González-Hernández et al. In Prep.)

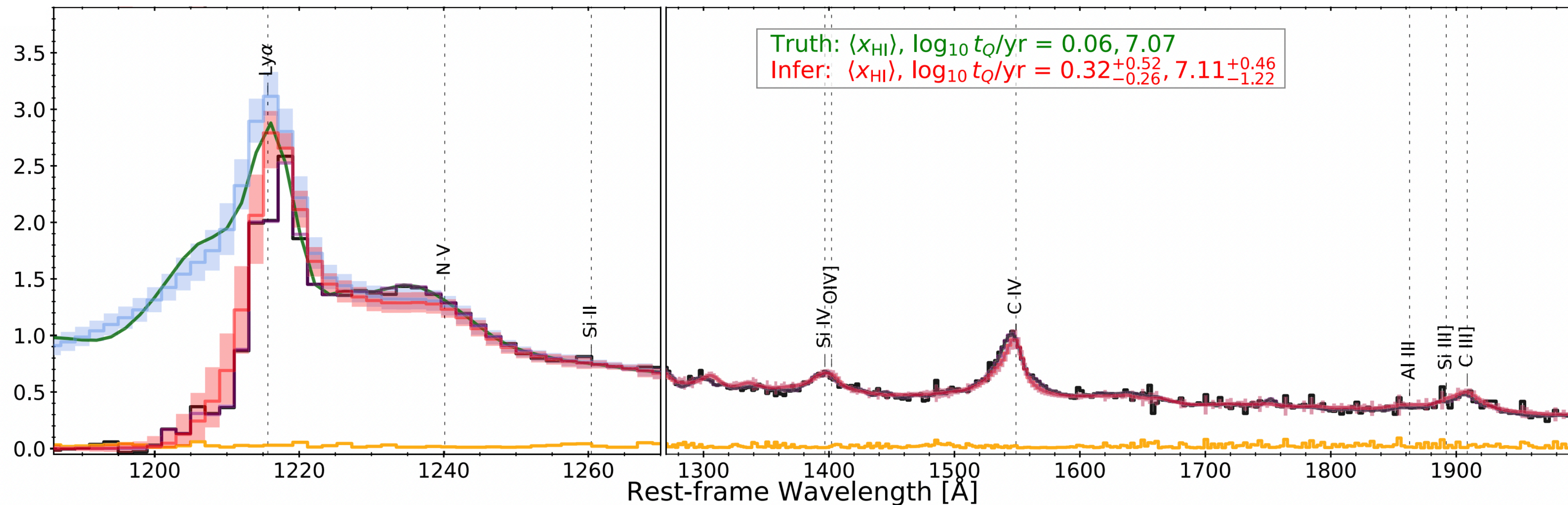


Parameter Estimation on Observations

Beyond Summary Statistics

- Hennawi et al. (2025) derived an analytical likelihood for Ly α Damping Wings observed in quasar spectra.

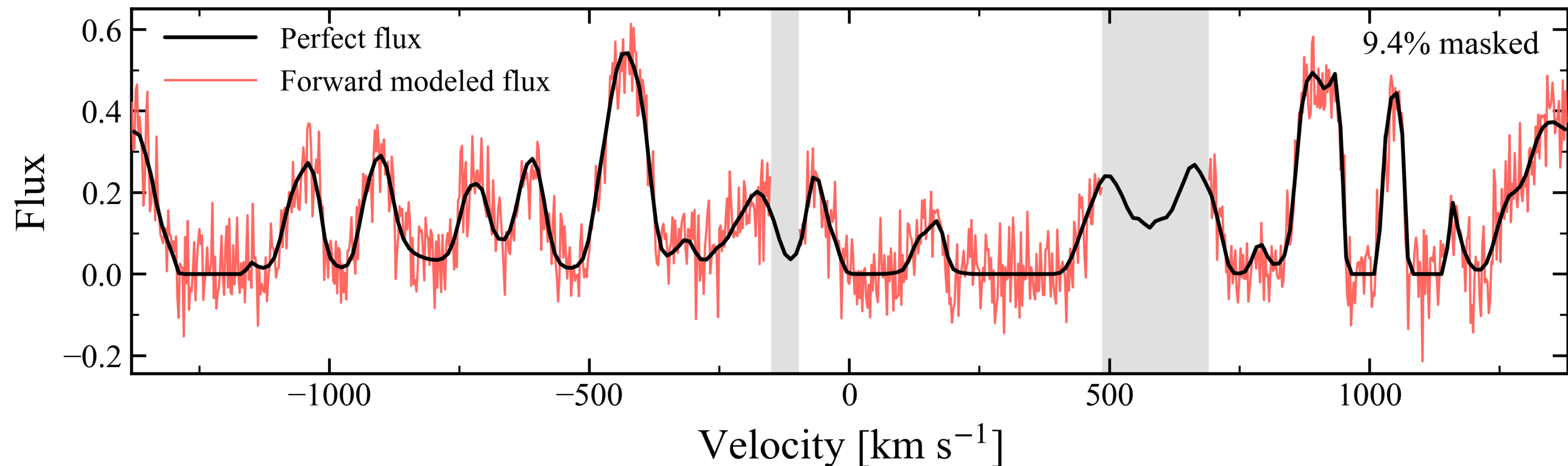
$$\begin{aligned} L(\mathbf{f}|\boldsymbol{\sigma}, \boldsymbol{\theta}, \boldsymbol{\eta}) &= \int \mathcal{N}(\mathbf{f}; \langle \mathbf{t} \rangle \circ \mathbf{s}, \boldsymbol{\Sigma} + \langle \mathbf{S} \rangle \mathbf{C}_t \langle \mathbf{S} \rangle) \\ &\quad \times \mathcal{N}(\langle \mathbf{t} \rangle \circ \mathbf{s}; \langle \mathbf{t} \rangle \circ \langle \mathbf{s} \rangle, \langle \mathbf{T} \rangle \mathbf{C}_s \langle \mathbf{T} \rangle) d(\langle \mathbf{t} \rangle \circ \mathbf{s}). \\ &= \mathcal{N}(\mathbf{f}; \langle \mathbf{t} \rangle \circ \langle \mathbf{s} \rangle, \boldsymbol{\Sigma} + \langle \mathbf{S} \rangle \mathbf{C}_t \langle \mathbf{S} \rangle + \langle \mathbf{T} \rangle \mathbf{C}_s \langle \mathbf{T} \rangle) \end{aligned}$$



The First Step

Towards Parameter Estimation on Flux Observations

- Consider forward-modeled Ly α flux skewers, with no quasar continuum, and missing data chosen via a missing completely at random (MCAR) mechanism with up to 10% of missingness.



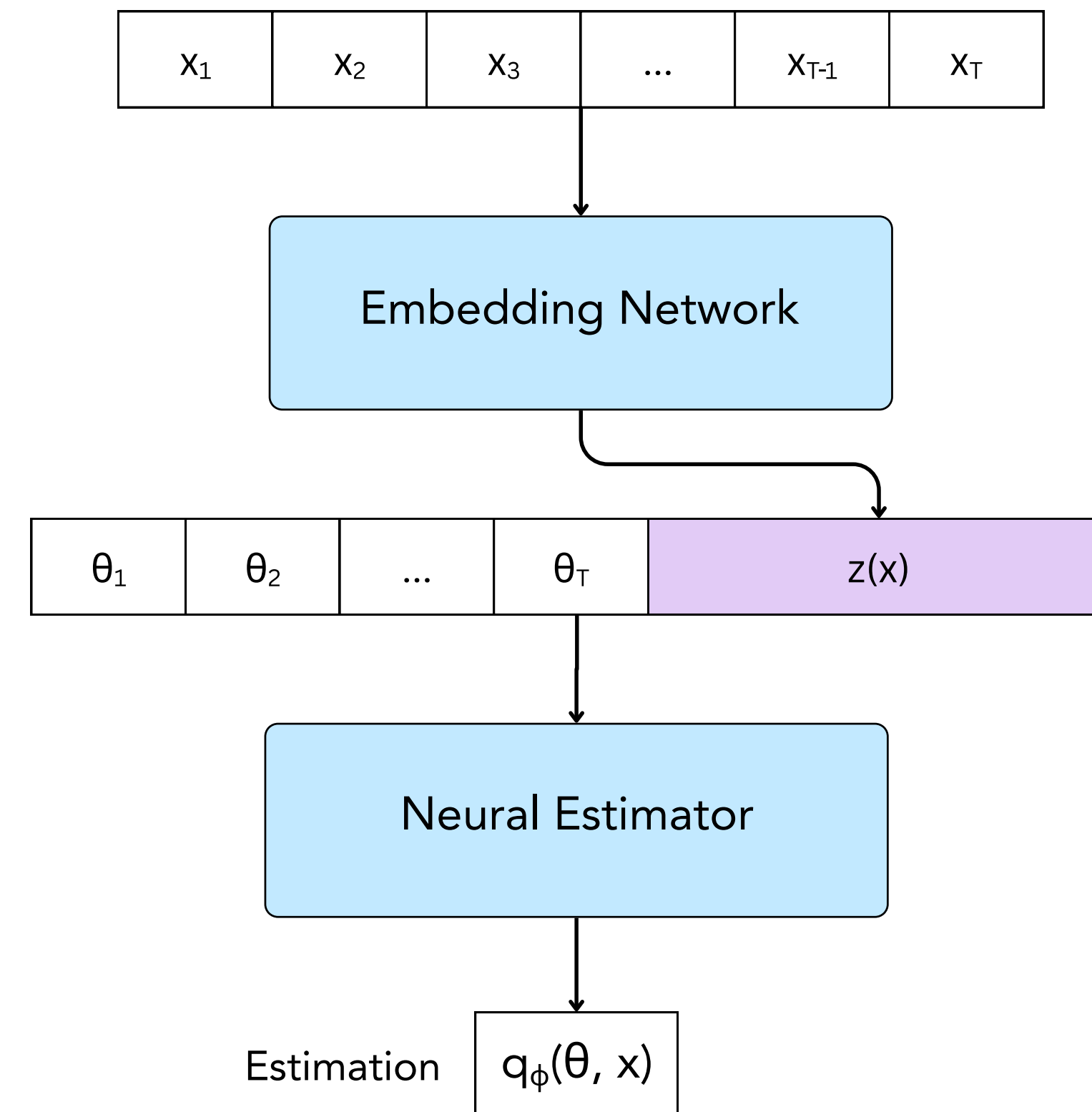
Embedding Networks

To process raw observations

- Embedding Networks are used to process observations and learn a representation of them. This representation is then used by a given estimator.
- Lueckmann et al. (2017) uses MLPs, Gloeckler et al. (2024) uses transformers and score diffusion, Verma et al. (2025) uses neural processes, etc.

Embedding Network

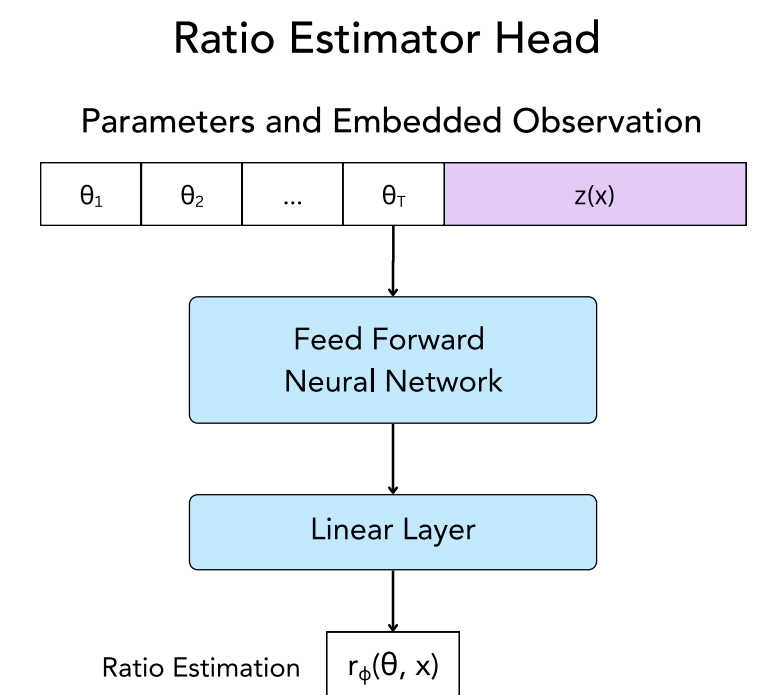
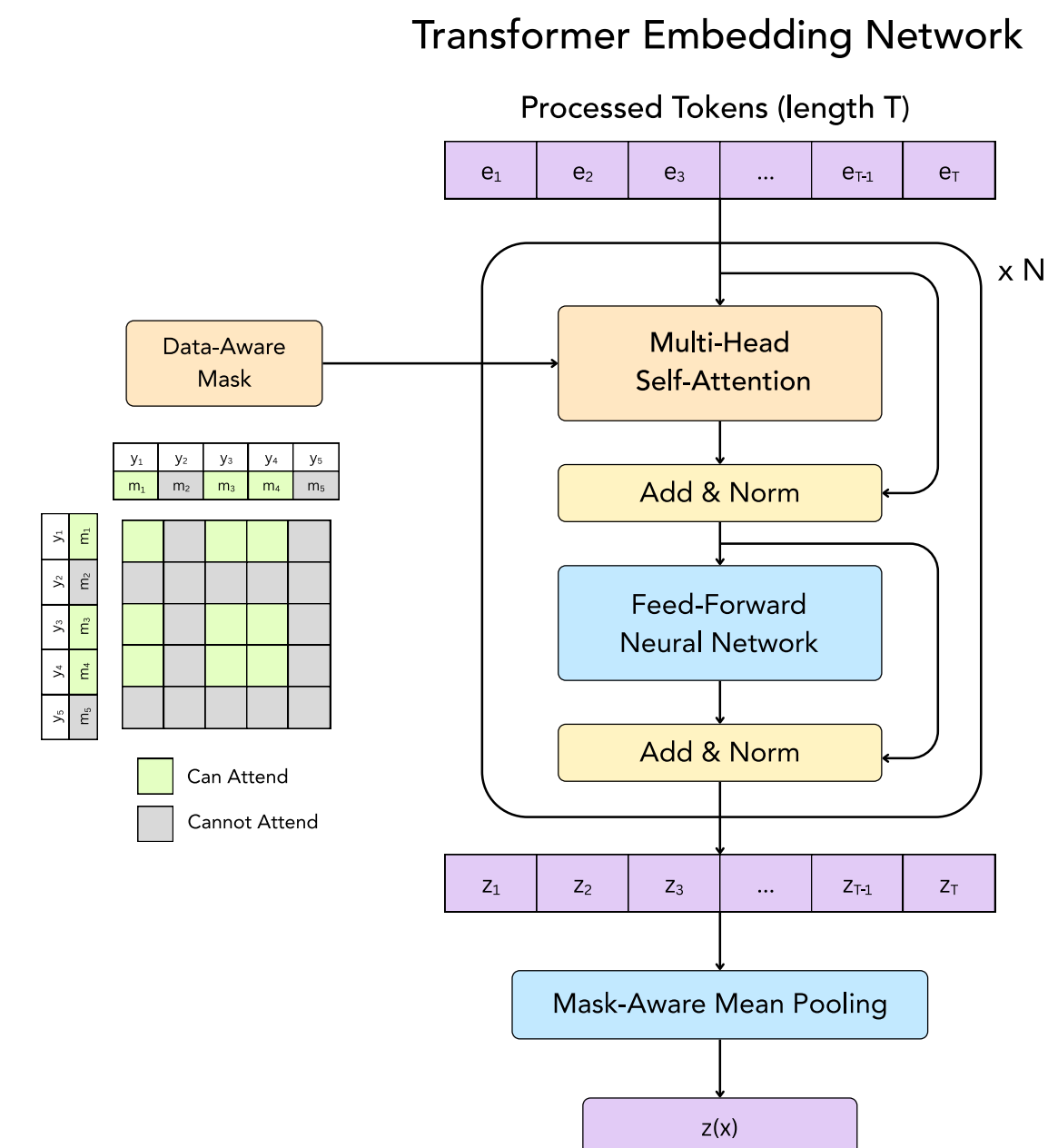
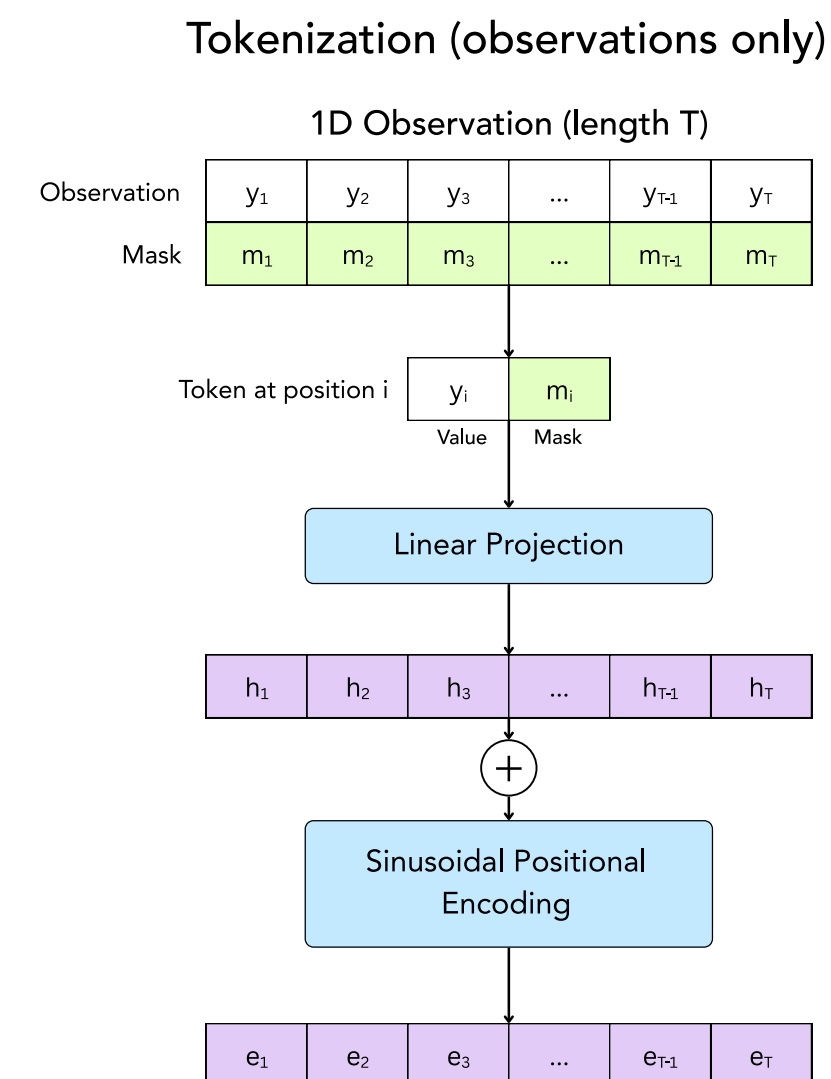
Only applied to observations



Using a Transformer

As an embedding network

- Attention mechanism.
- A simple tokenization for this problem.
- Joint training, as is usual for embedding networks.

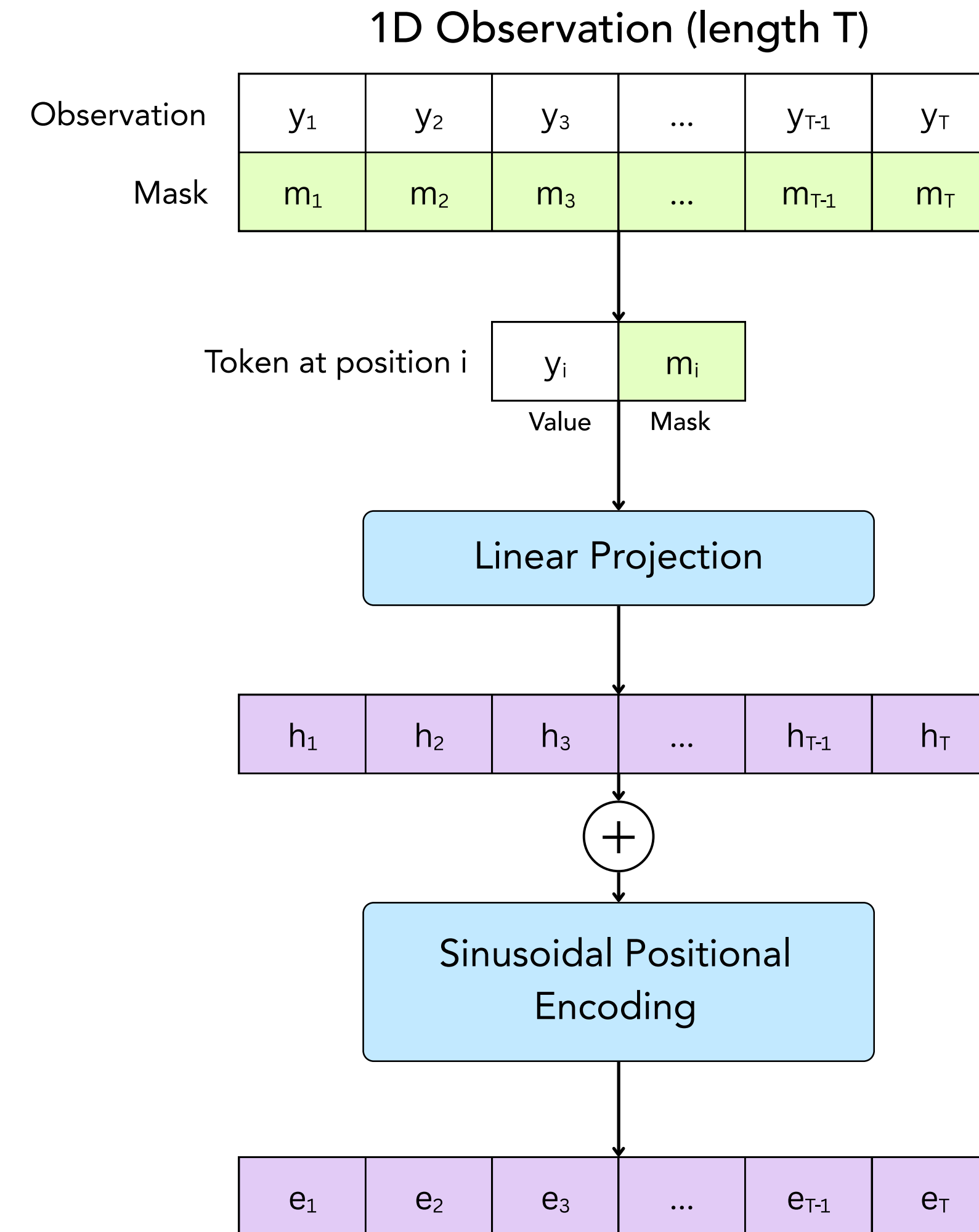


Using a Transformer

A Simple Tokenization

- Only observations are tokenized.
- Raw input tokens are observations and masks.
- Learnable linear projection with sinusoidal positional encoding.

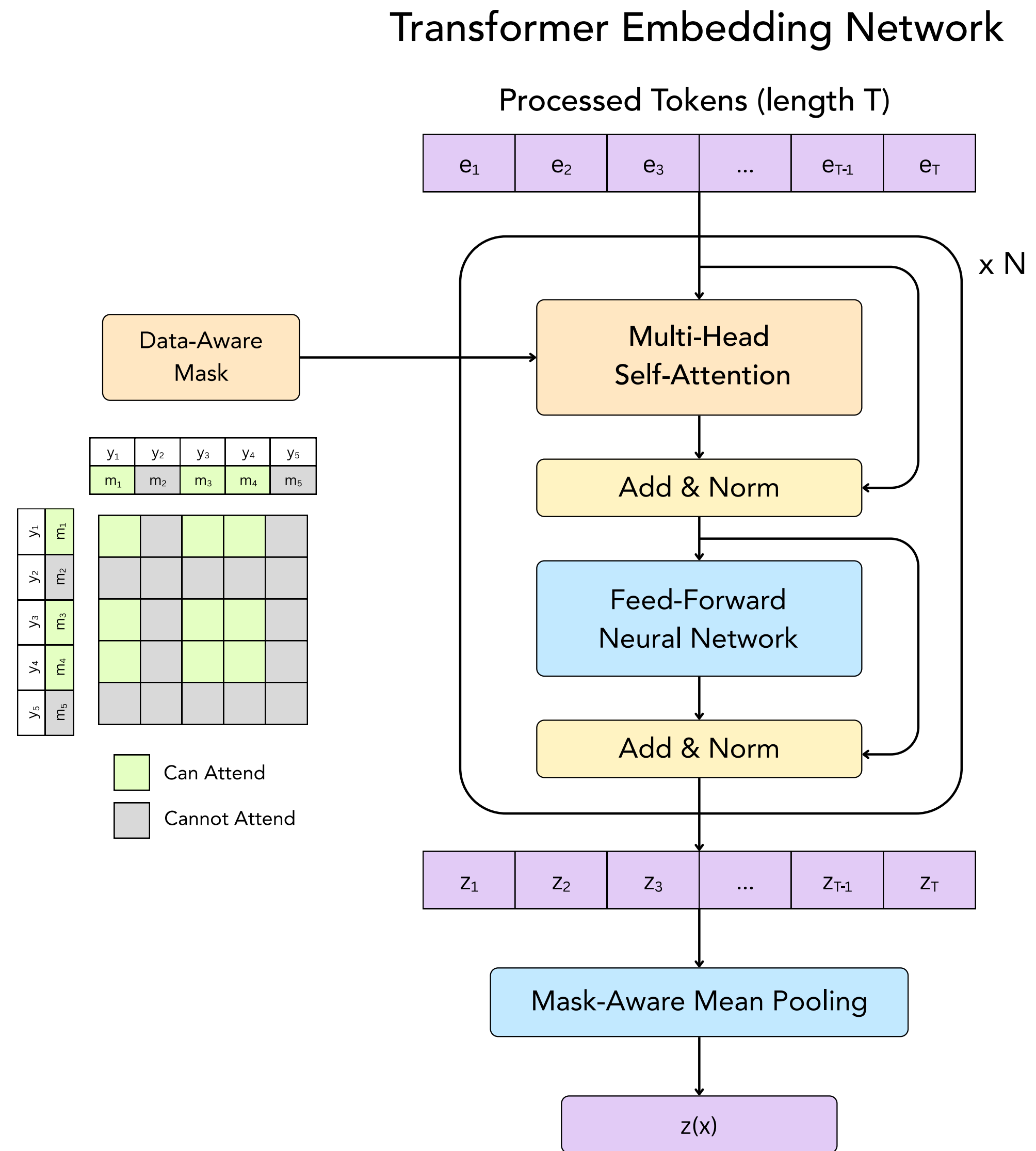
Tokenization (observations only)



Using a Transformer

Attention and Transformer Blocks

- Processed tokens as input.
- Missing values dictate the attention mask.
- Mean pooling only uses observed values.
- Learned representation.



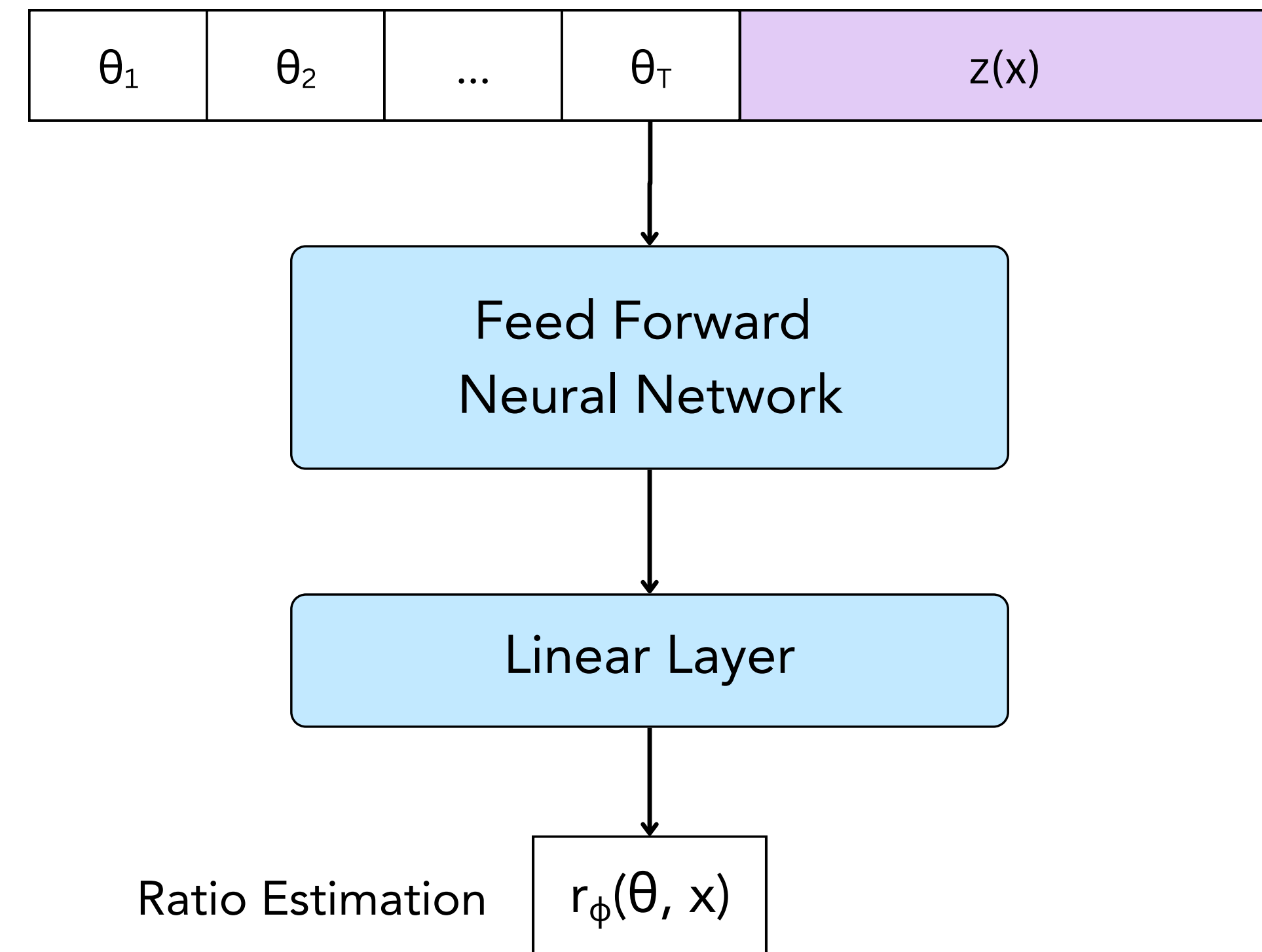
Using a Transformer

The Ratio Estimator

- Learned representation of observation and parameters are concatenated.
- A simple feed-forward neural network predicts the likelihood ratio.

Ratio Estimator Head

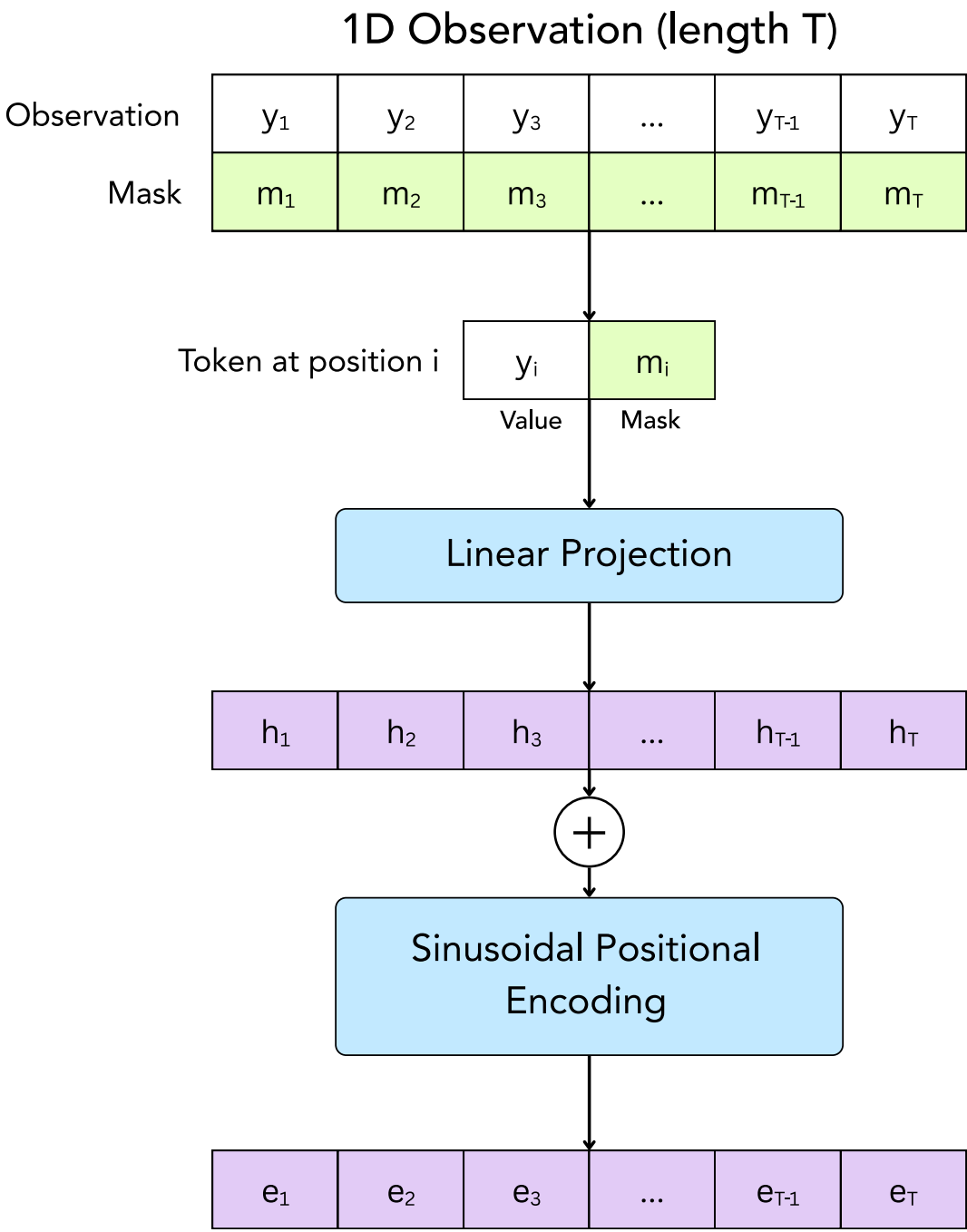
Parameters and Embedded Observation



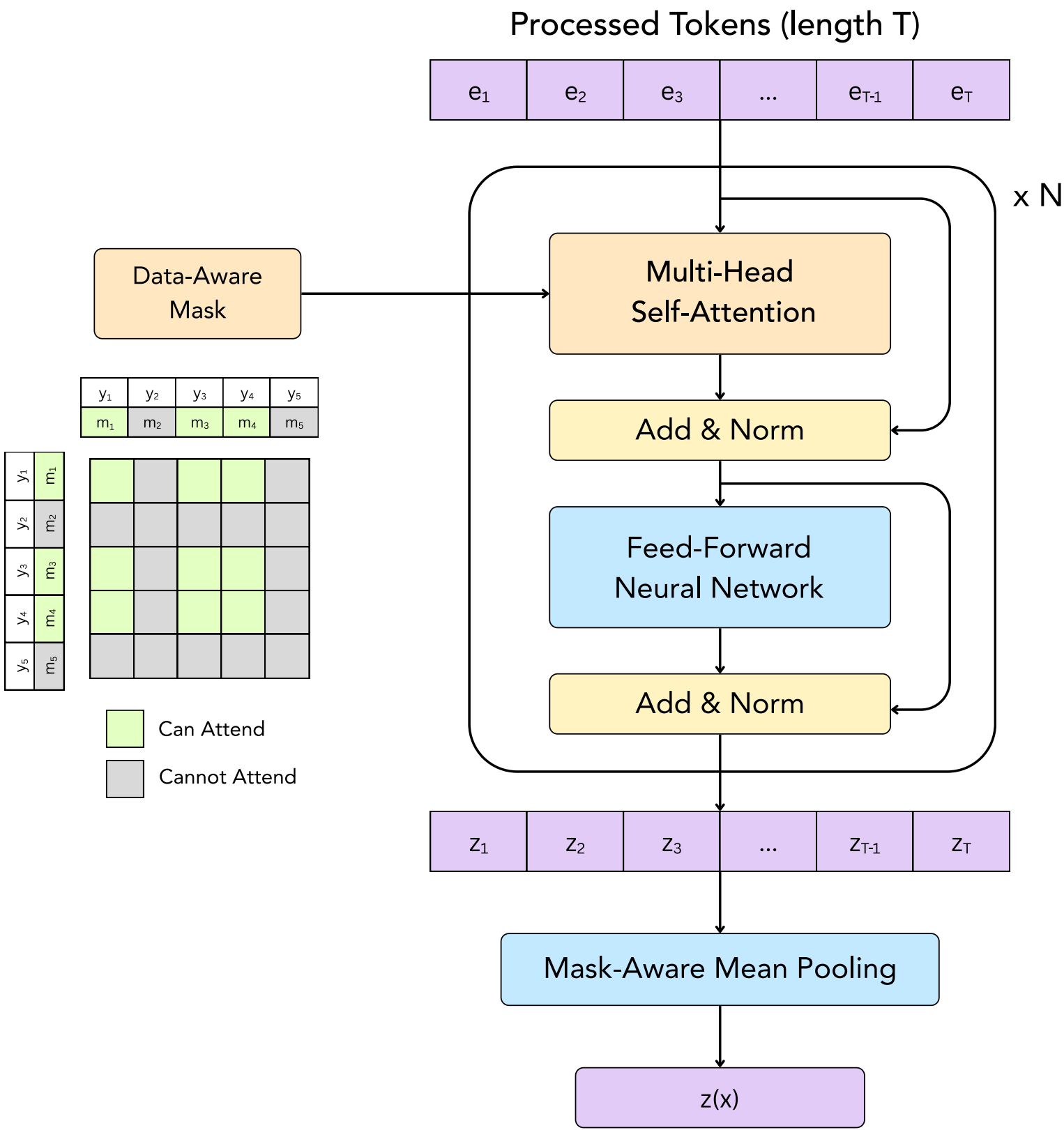
Using a Transformer

Overview

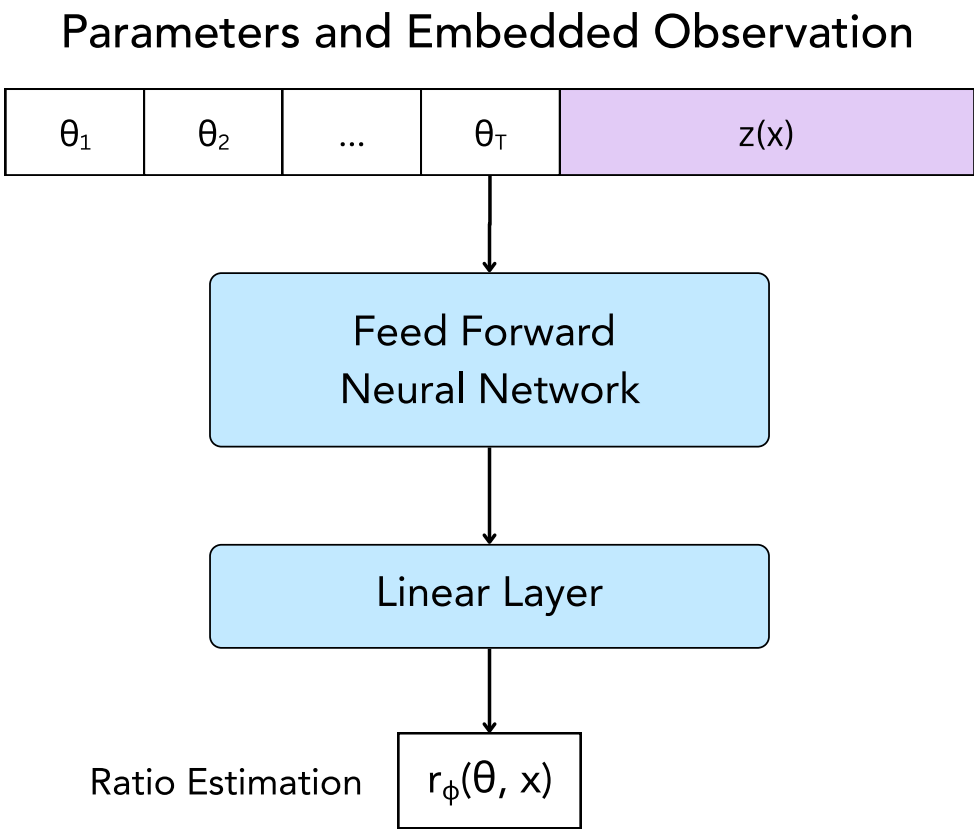
Tokenization (observations only)



Transformer Embedding Network



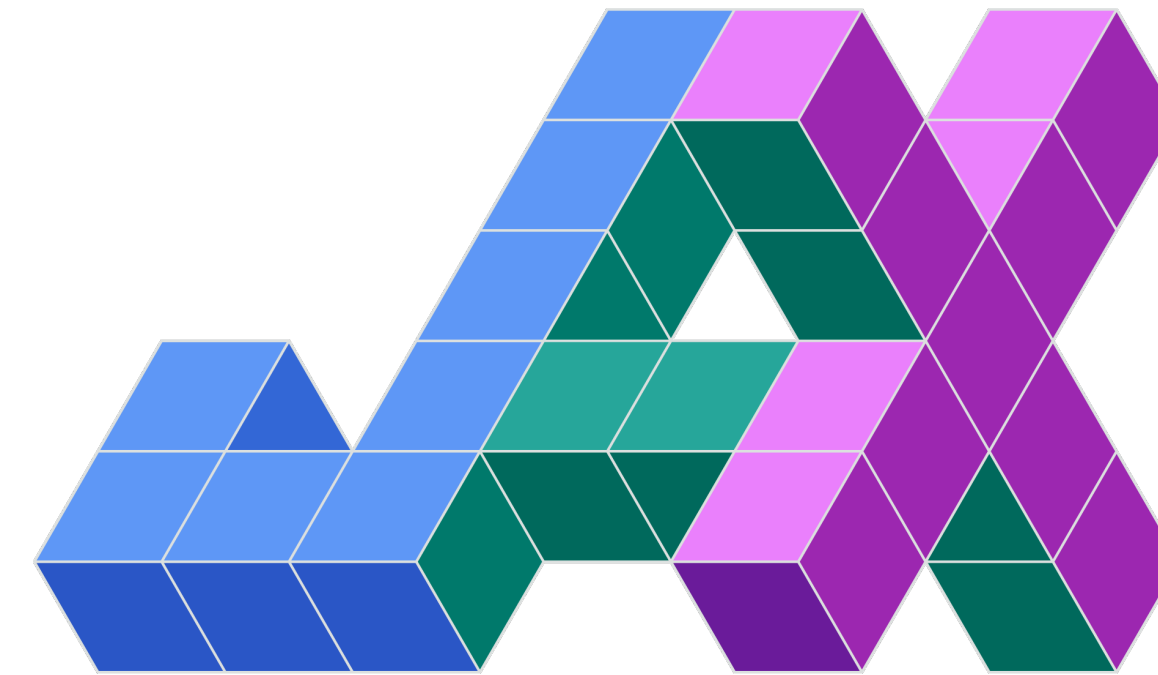
Ratio Estimator Head



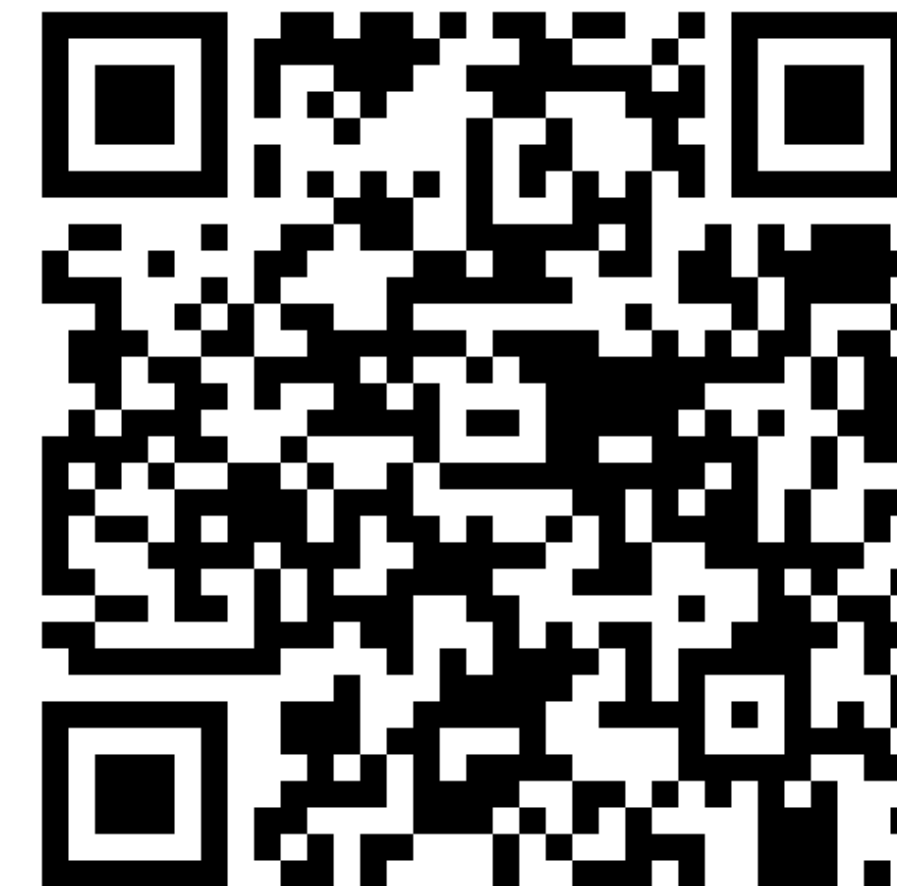
The Implementation

Of this method

- JAX backend for fast Hamiltonian Monte Carlo.
- Simple MLP for the classifier and a simple transformer for the embedding network.
- A standard dataset contract that allows for a coding agent to run a full experiment (as a lab assistant).

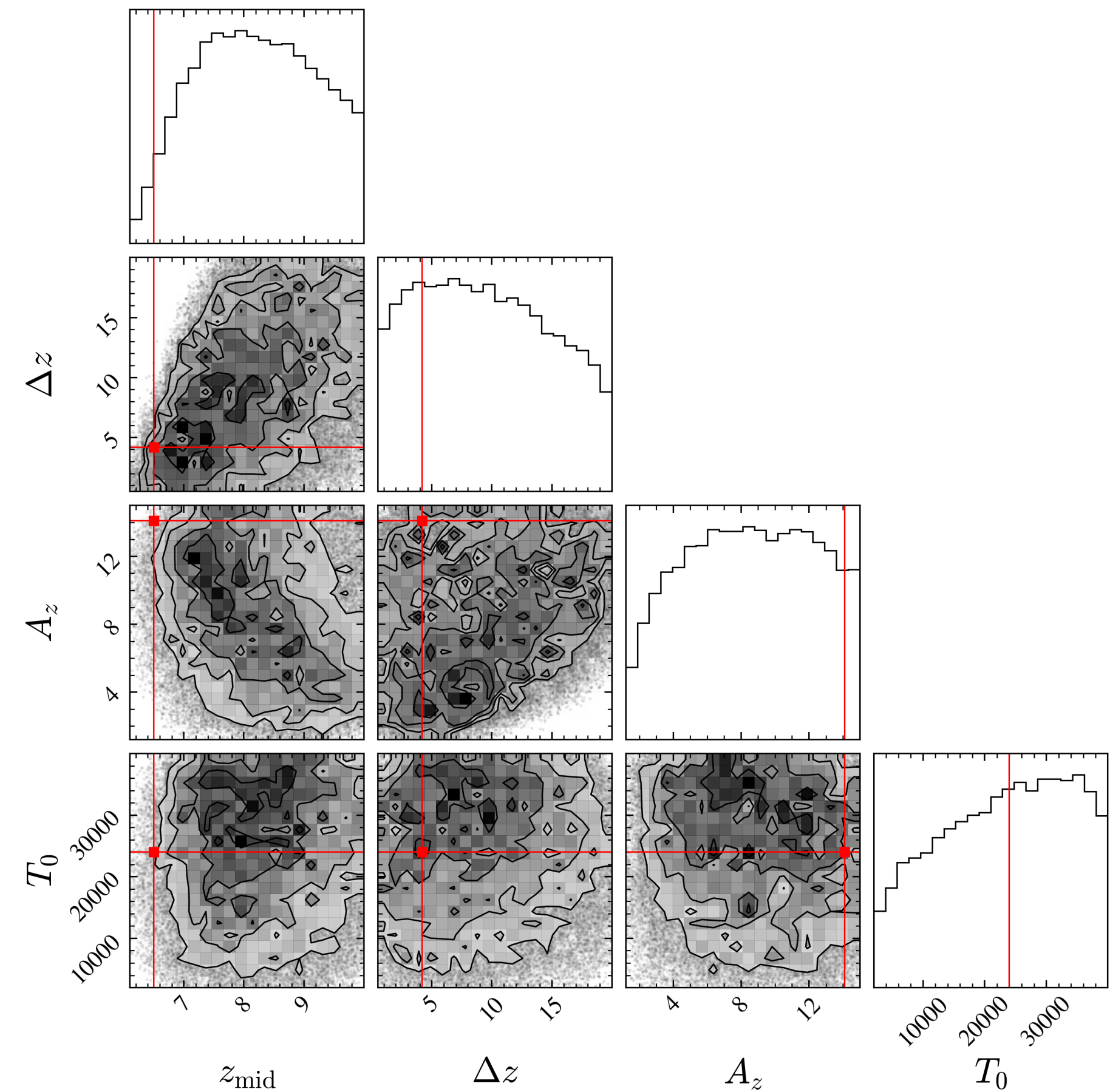
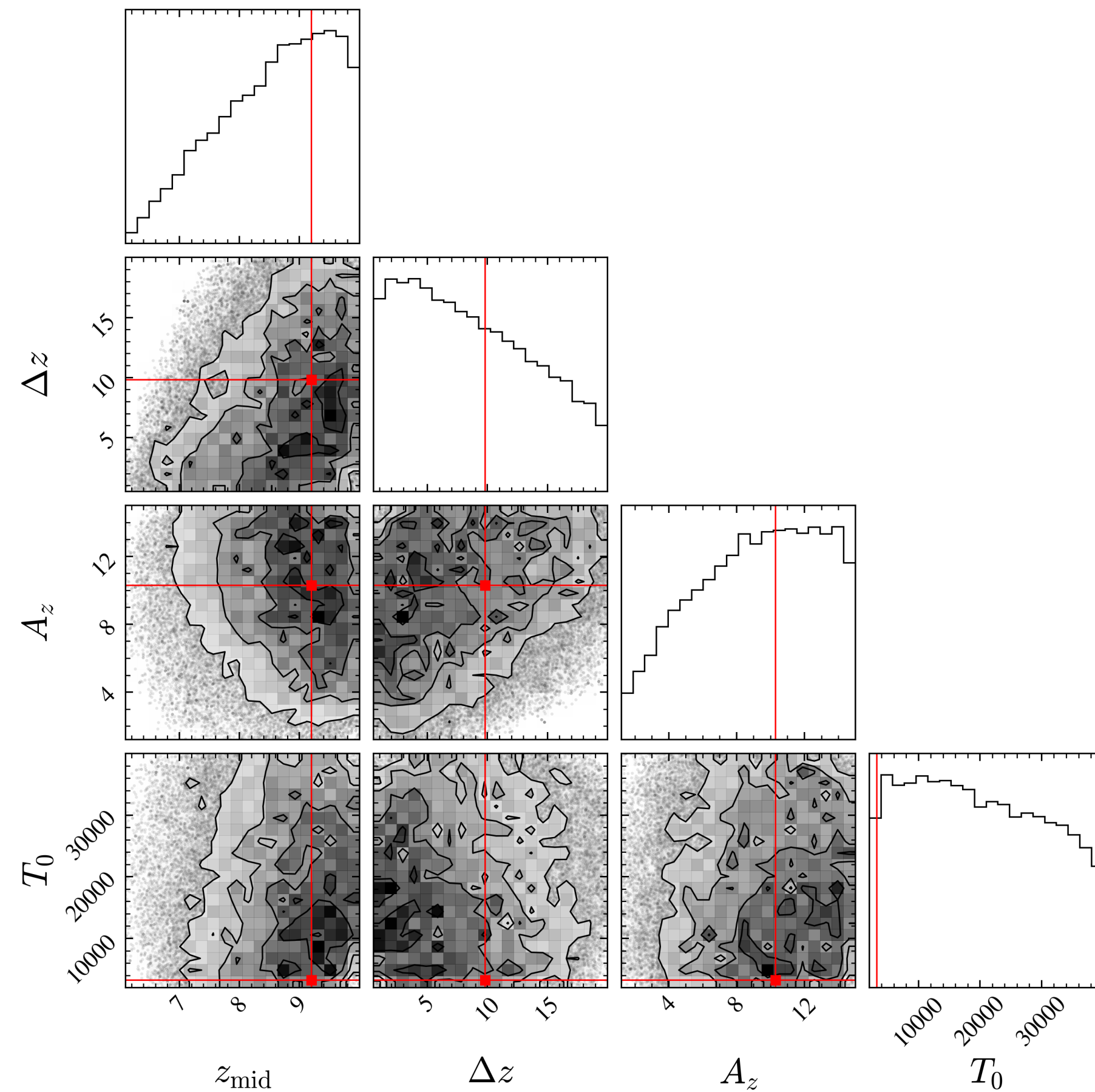


Code:



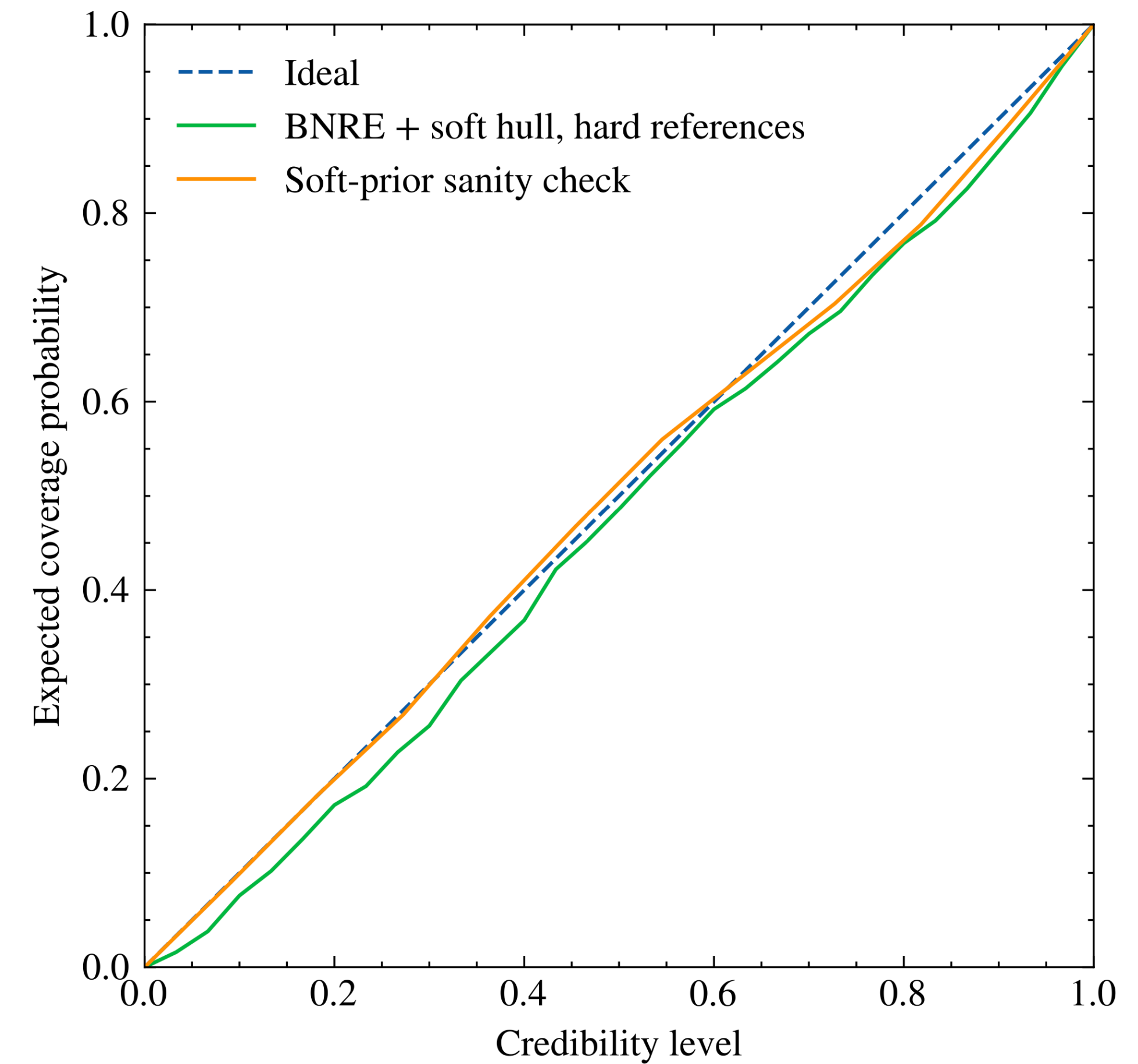
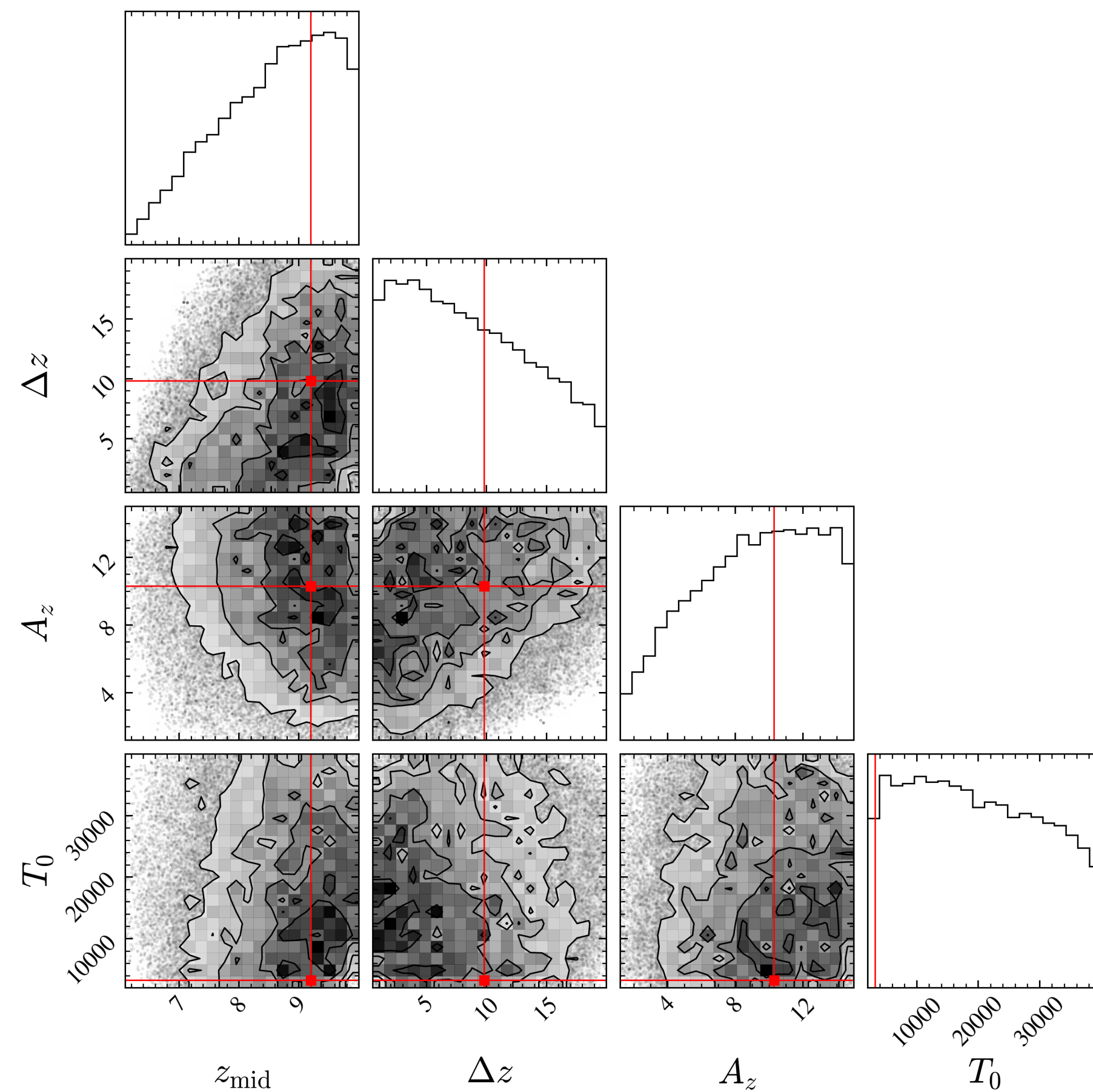
Results: Single Observation

Parameter Estimation



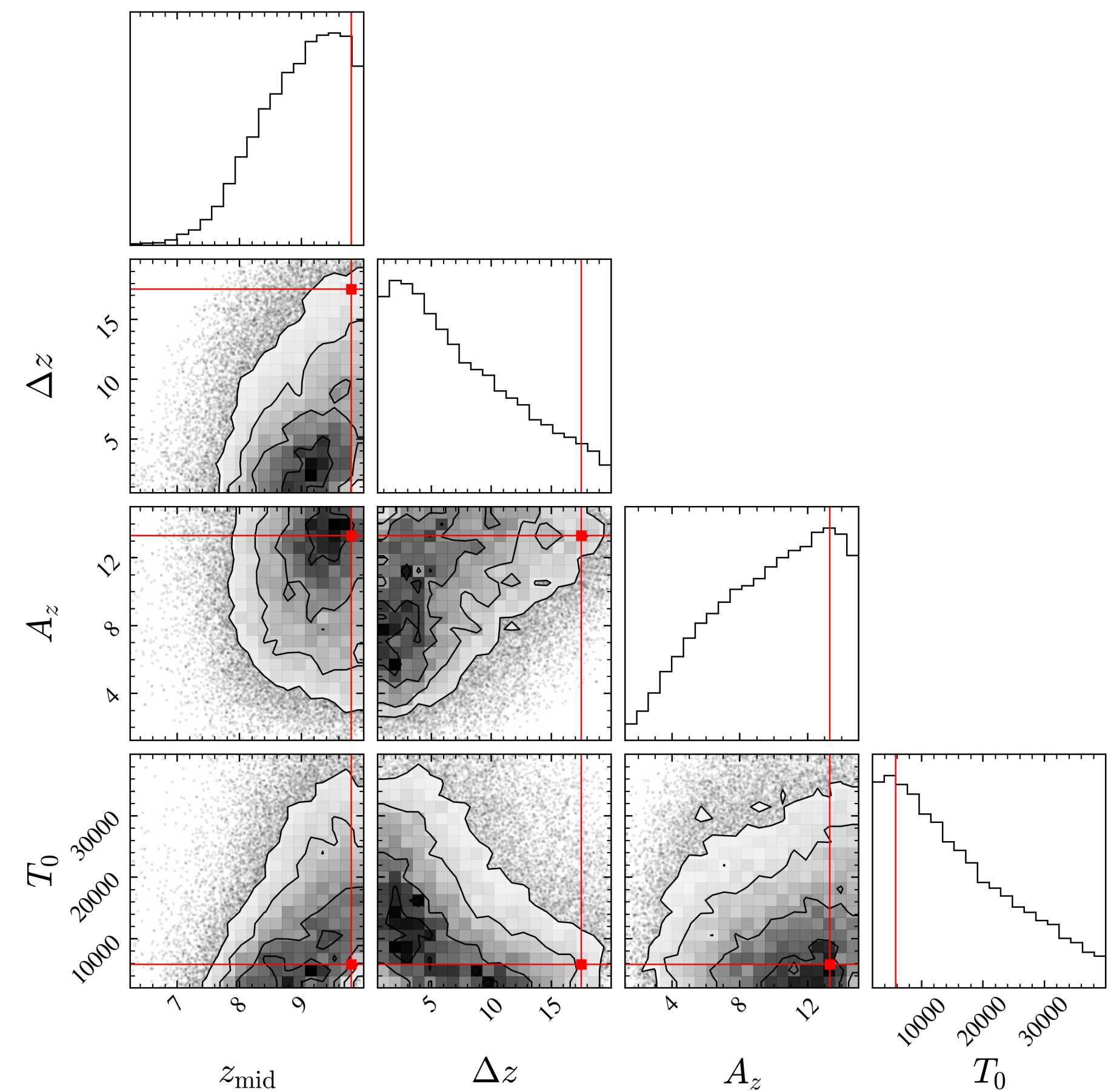
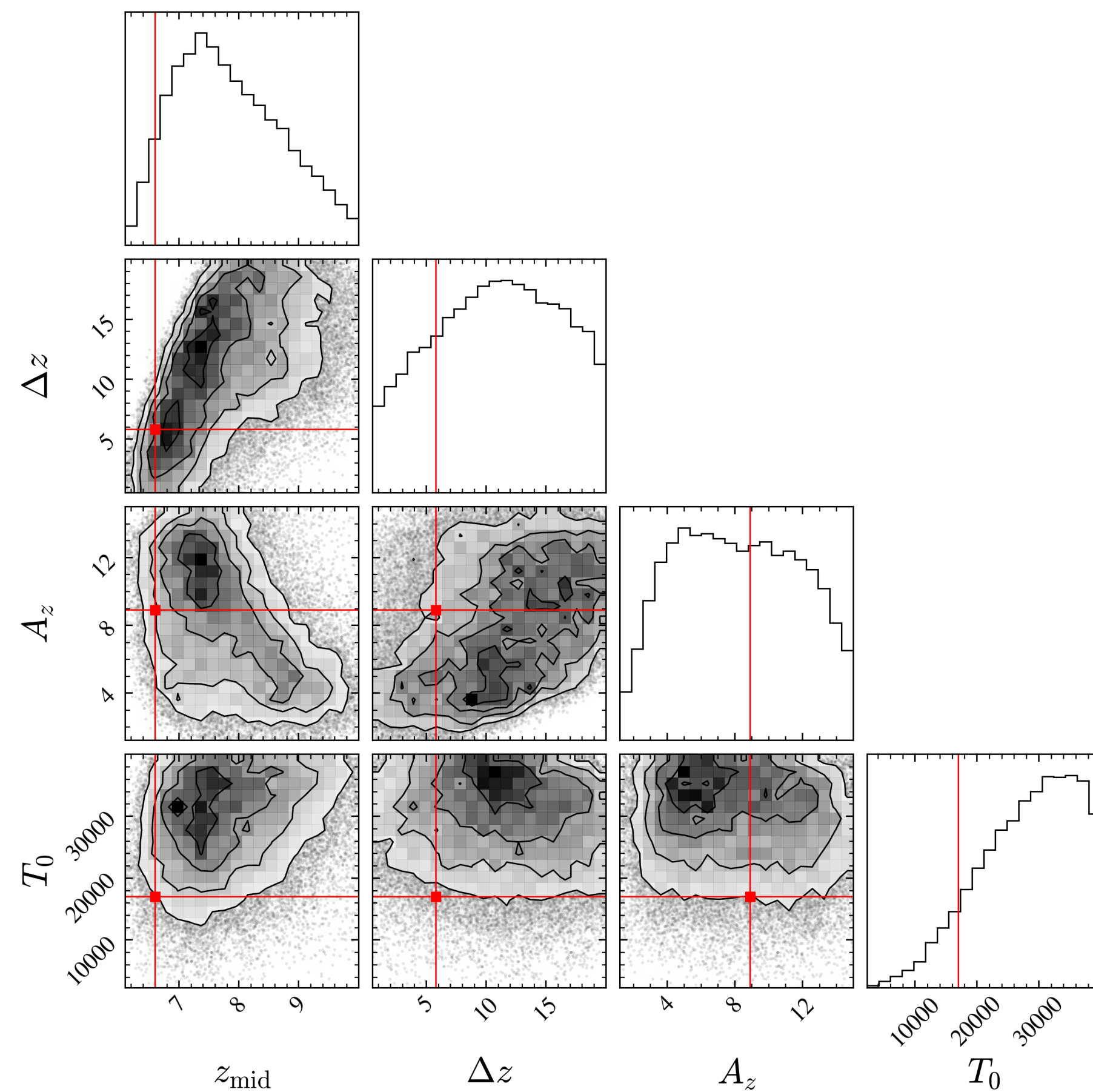
Results: Single Observation

Calibrated Posteriors



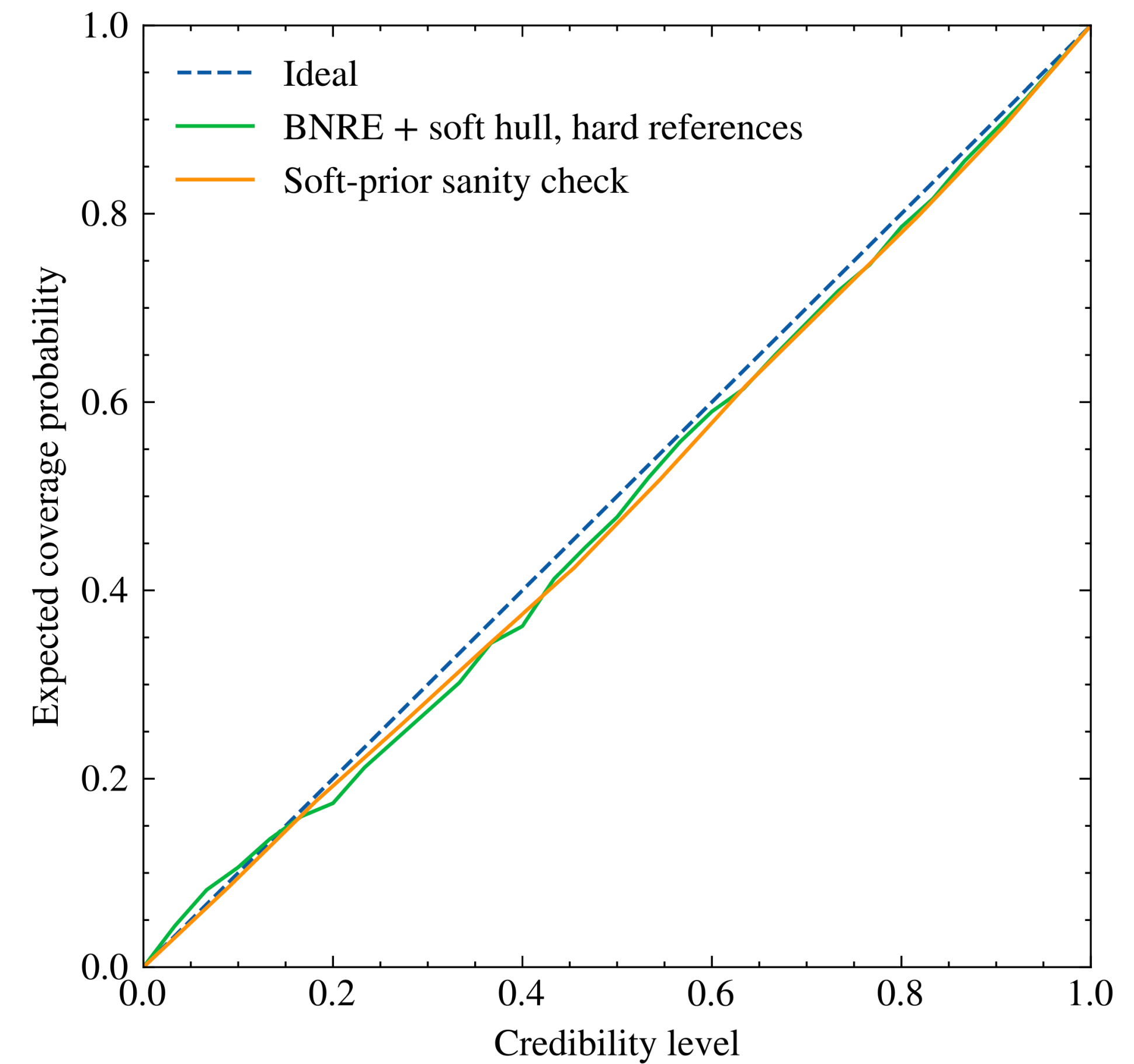
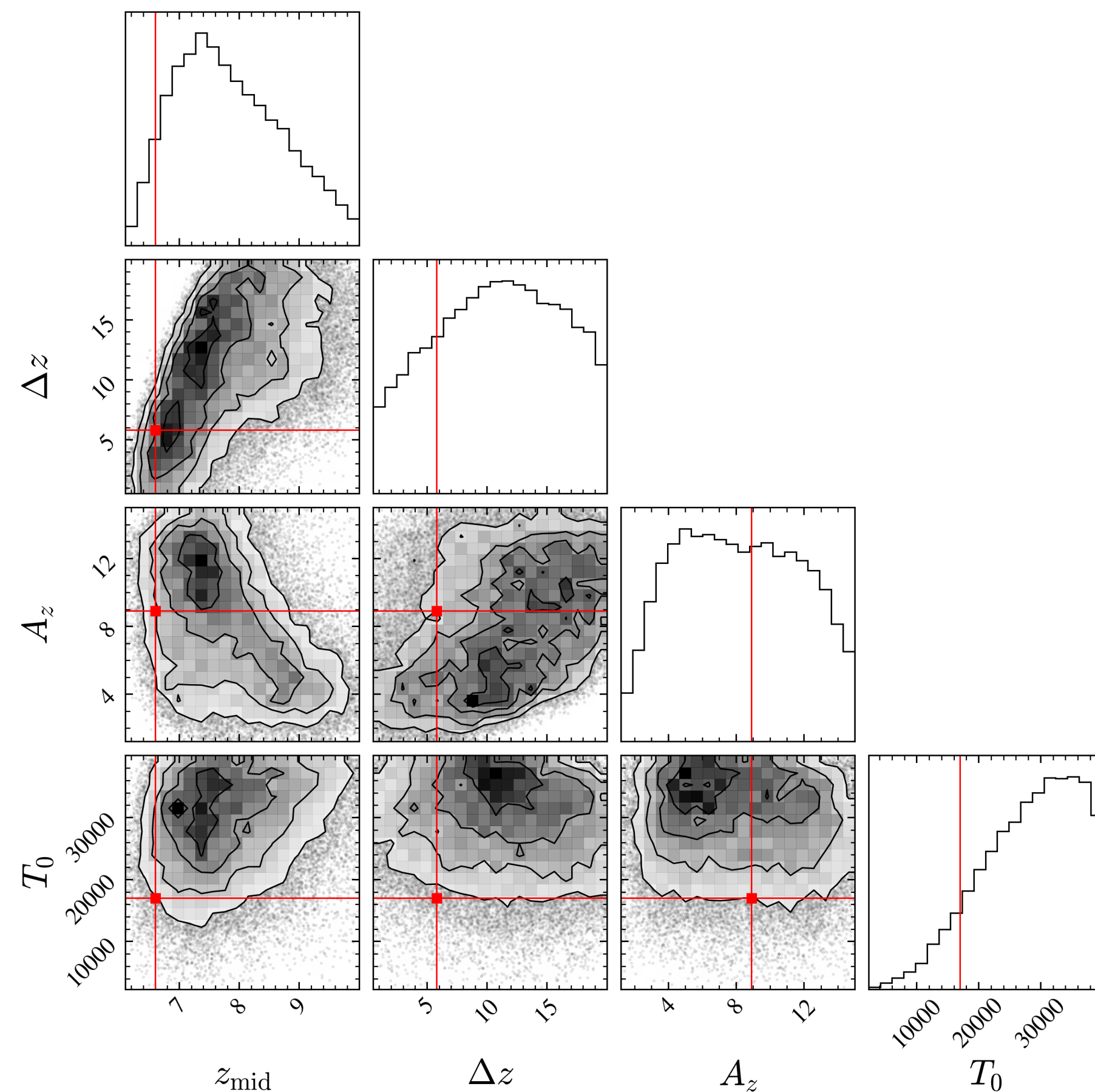
Results: Multiple Observations

Parameter Estimation with 10 “i.i.d” Observations



Results: Multiple Observations

Calibrated Posteriors

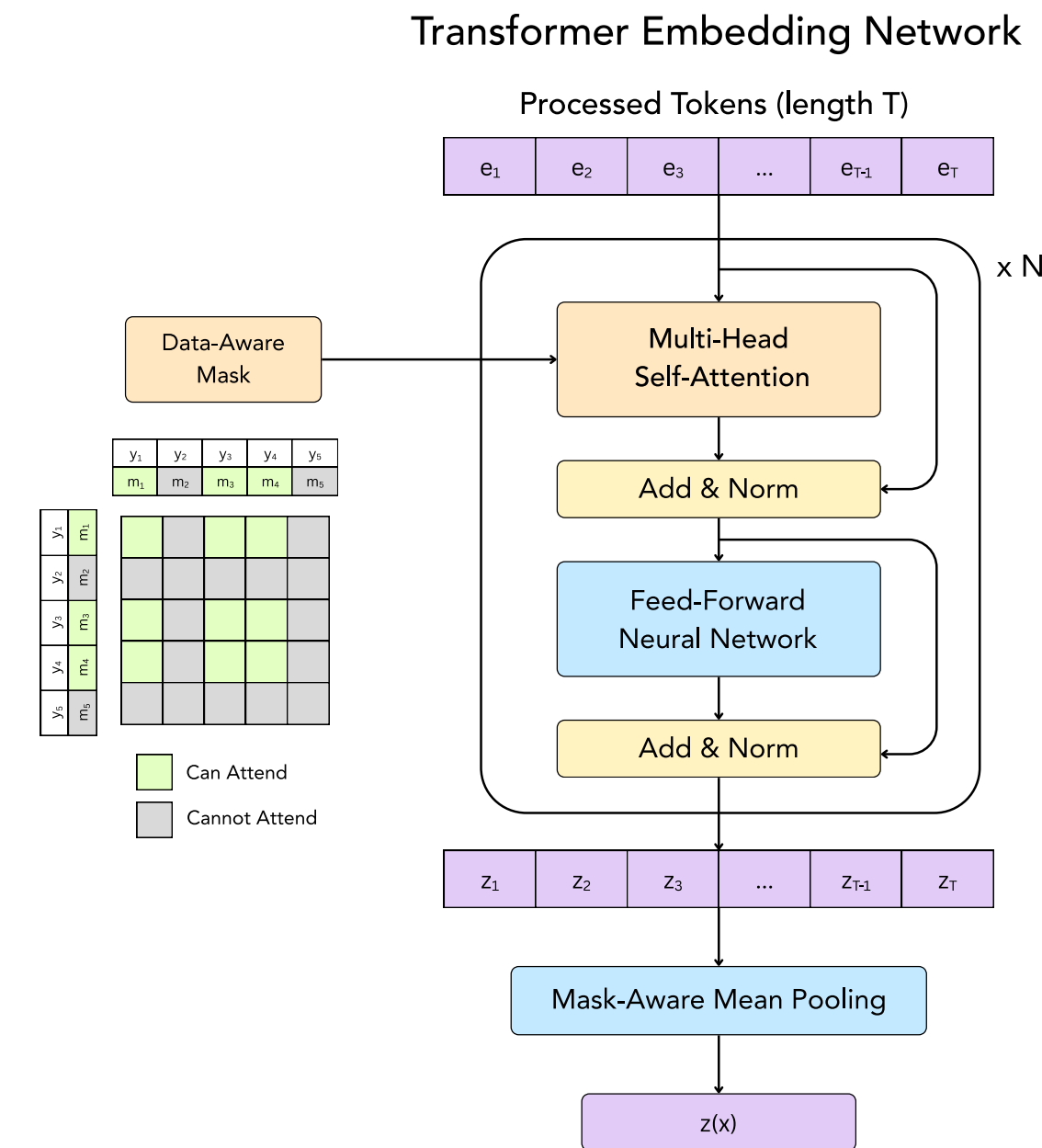


What's next?

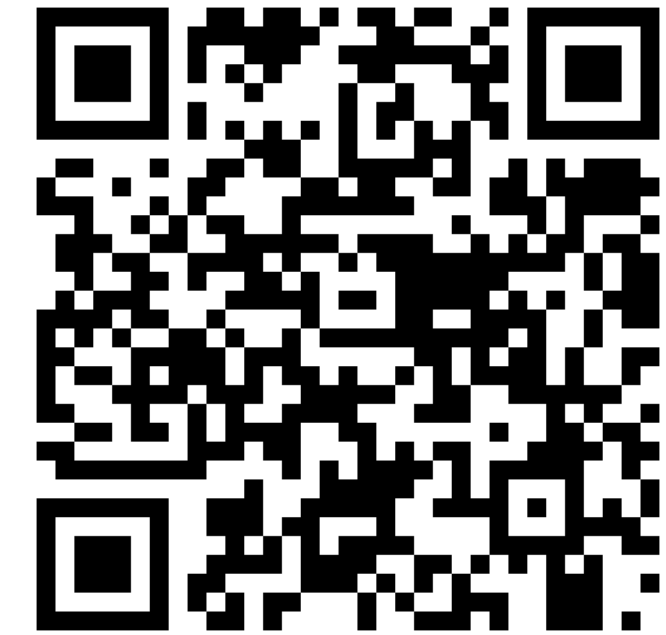
- Include QSO continuum reconstruction
- Improvements to multiple observations
- Use a hi-res simulation dataset
- Extend to other types of estimators

References

- Trac, H., Chen, N., Holst, I., Alvarez, M. A., & Cen, R. AMBER: A Semi-numerical Abundance Matching Box for the Epoch of Reionization. *ApJ*, 2022.
- Almgren, A. S., Bell, J. B., Lijewski, M. J., Lukić, Z., & Van Andel, E. Nyx: A Massively Parallel AMR Code for Computational Cosmology. *ApJ*, 2013.
- Lukić, Z., Stark, C. W., Nugent, P., White, M., Meiksin, A. A., & Almgren, A. The Lyman α Forest in Optically Thin Hydrodynamical Simulations. *MNRAS*, 2015.
- González-Hernández, D., Doughty, C., Wolfson, M., Hennawi, J. F., & Jin, Z. Using Neural Emulators and Hamiltonian Monte Carlo to constrain the Epoch of Reionization's History with the Ly α Forest Power Spectrum. *arXiv:2509.13498*, 2025a.
- González-Hernández, D., Wolfson, M., & Hennawi, J. F. Reliable Parameter Inference for the Epoch of Reionization using Balanced Neural Ratio Estimation. *arXiv:2511.02808*, 2025b.
- Hennawi, J. F., Kist, T., Davies, F. B., & Tamasas, J. Precisely measuring the cosmic reionization history from IGM damping wings towards quasars. *MNRAS*, 2025.
- Lueckmann, J.-M., Goncalves, P. J., Bassetto, G., Öcal, K., Nonnenmacher, M., & Macke, J. H. Flexible statistical inference for mechanistic models of neural dynamics. *arXiv:1711.01861*, 2017.
- Gloeckler, M., Deistler, M., Weilbach, C., Wood, F., & Macke, J. H. All-in-one simulation-based inference. *arXiv:2404.09636*, 2024.
- Verma, Y., Bharti, A., & Garg, V. Robust Simulation-Based Inference under Missing Data via Neural Processes. *ICLR*, 2025.

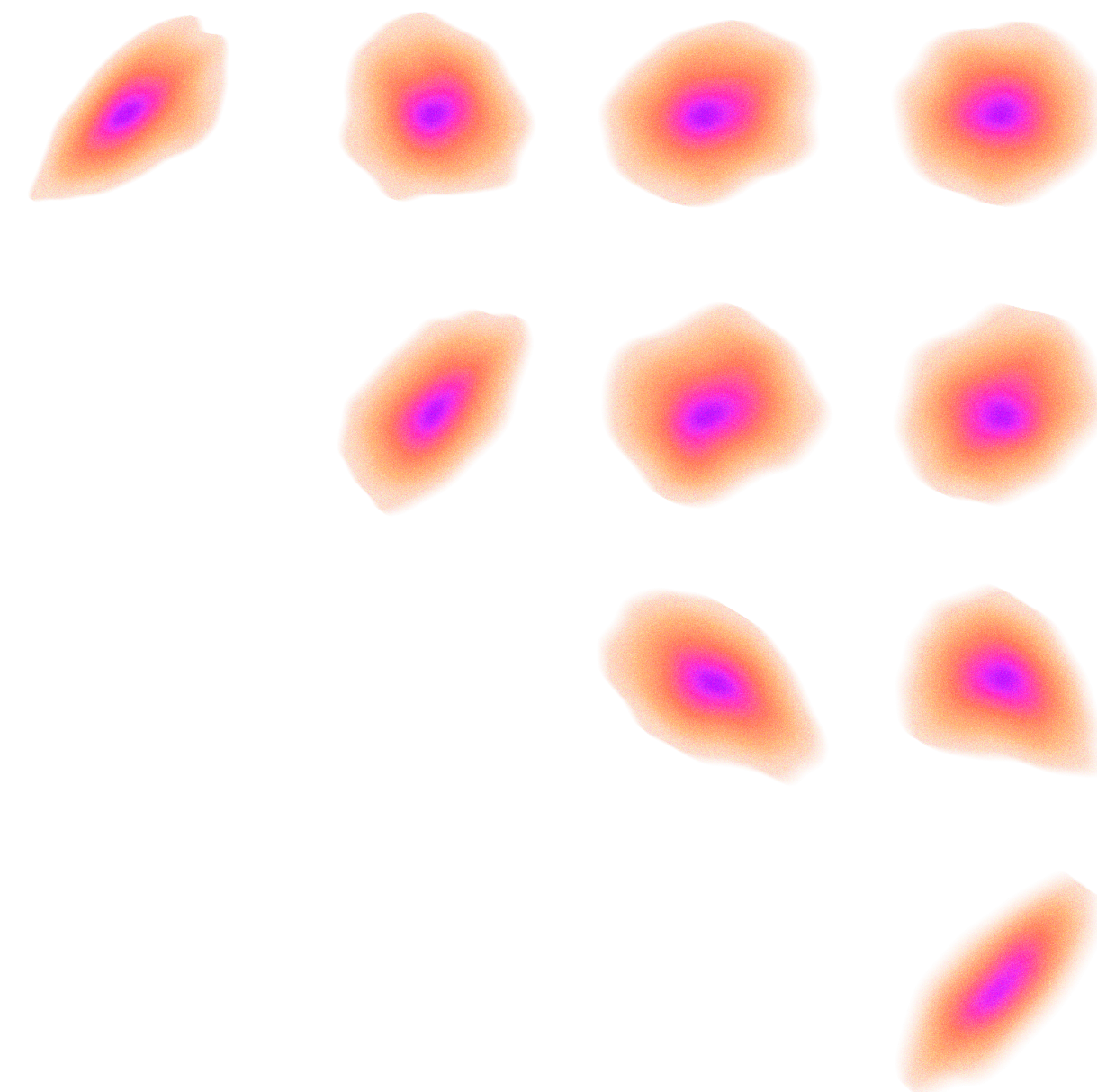
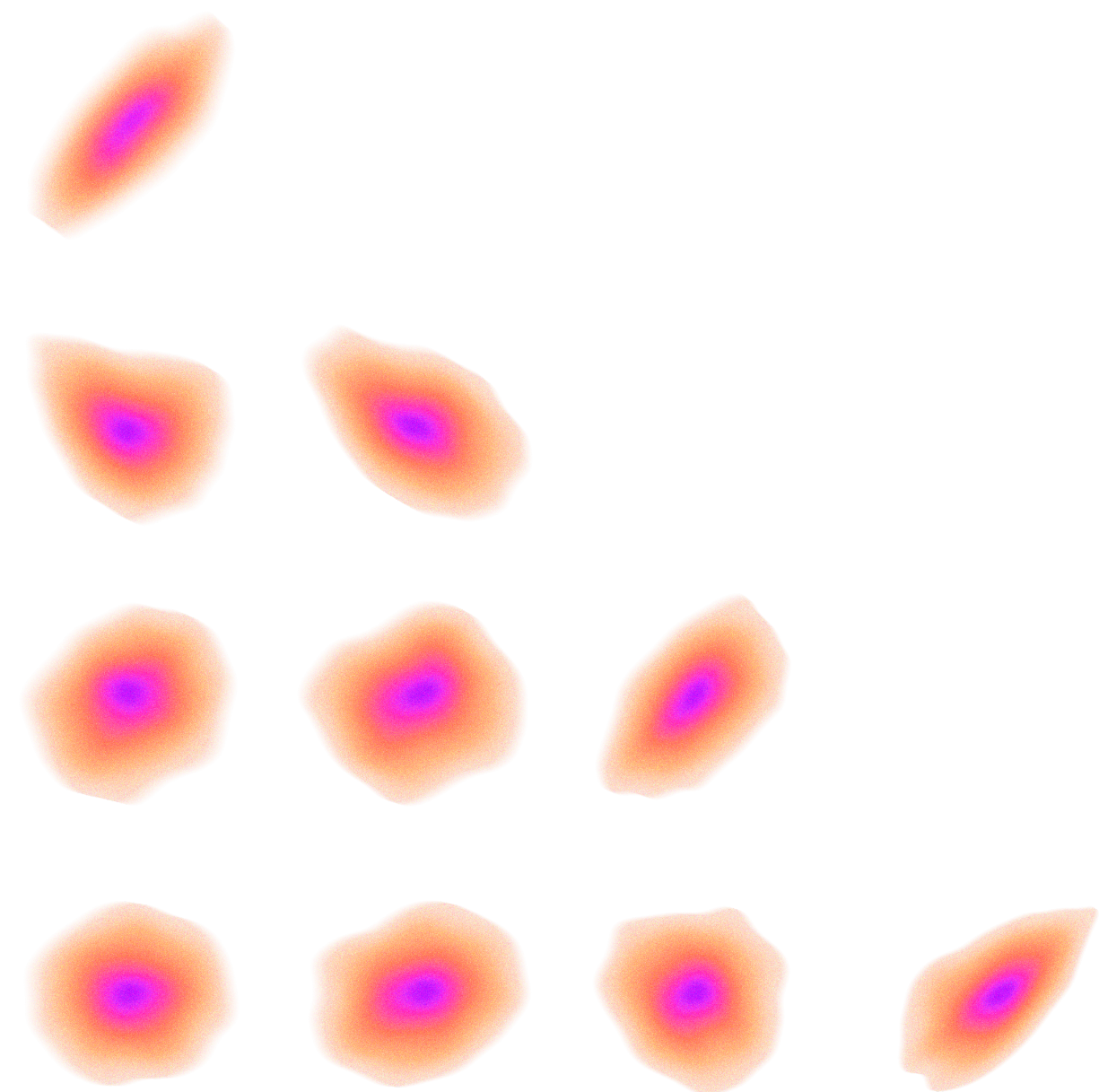


Code:

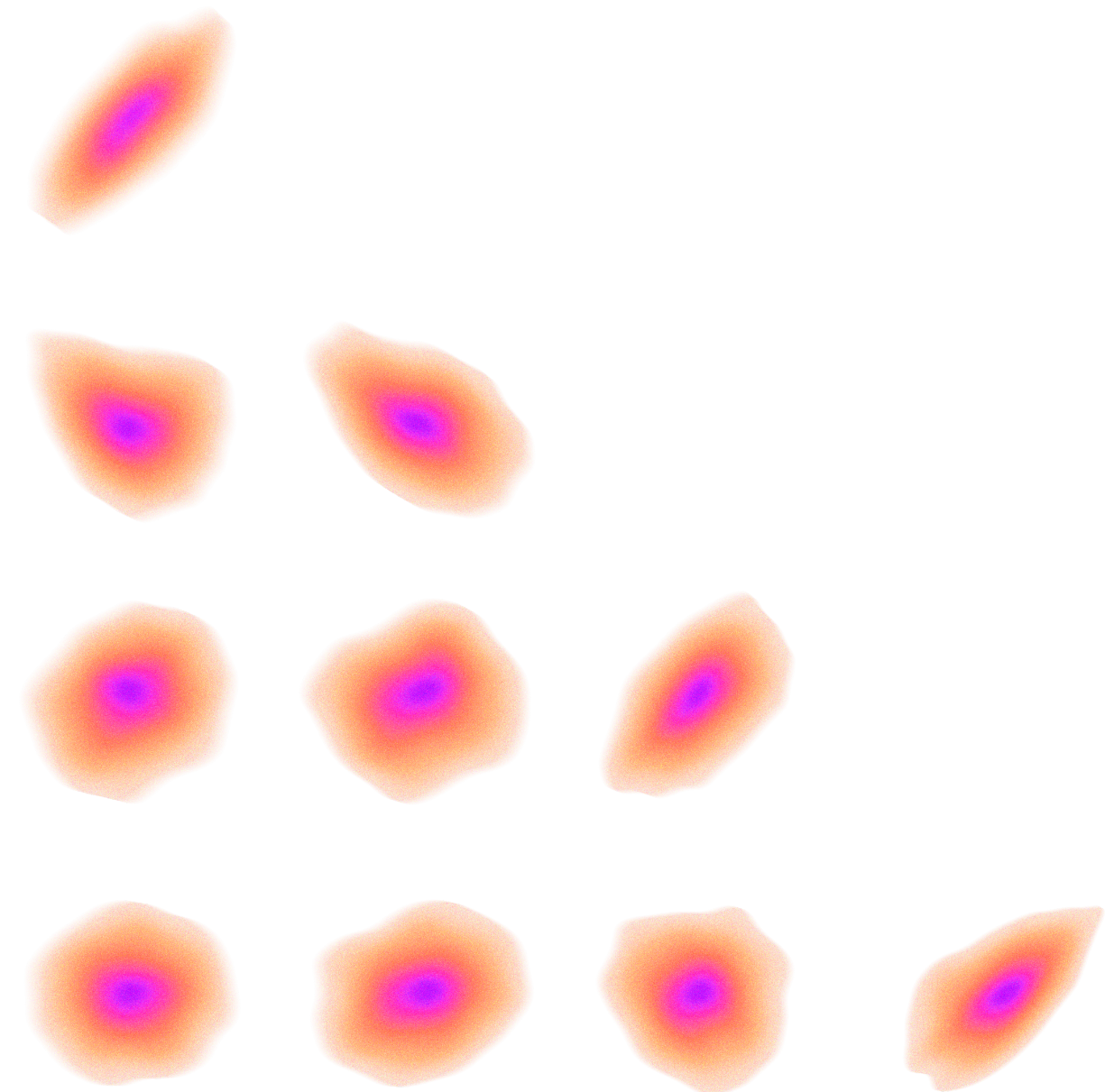


Thank you!

Questions?



Backup Slides



Filtering out Reionization Histories

