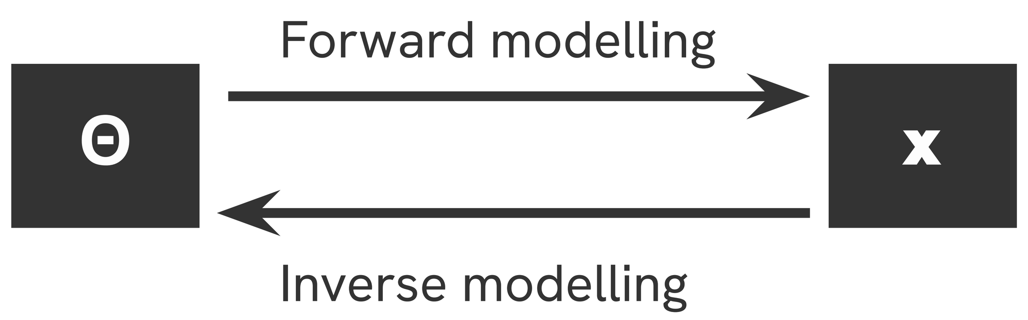
Learning at the Edge: Tailed-Uniform Sampling for Robust Simulation-Based Inference

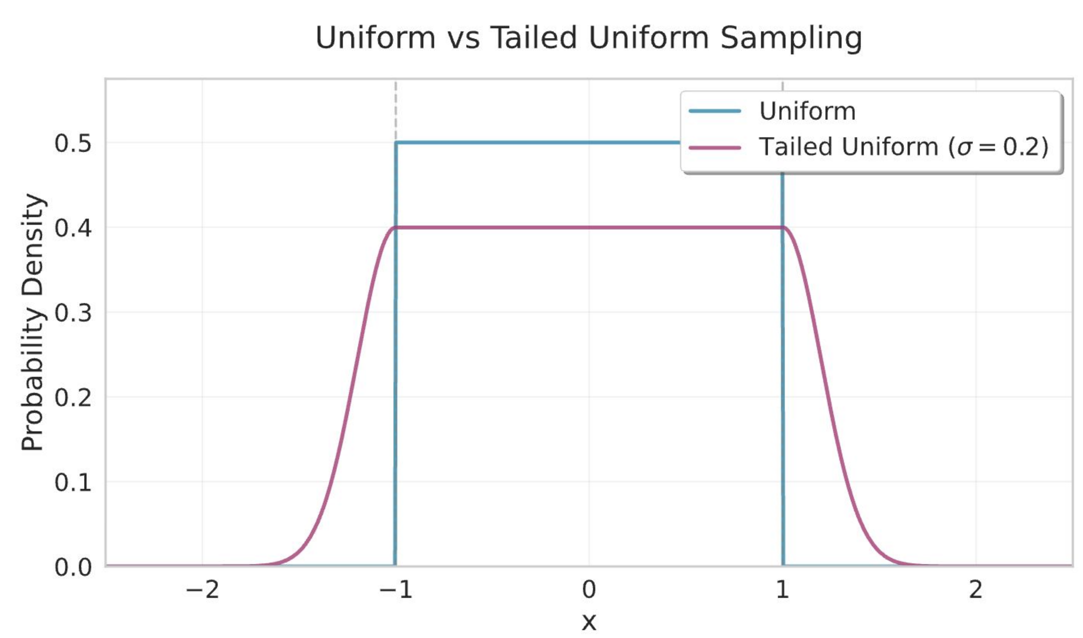
Chaipat Tirapongprasert Matthew Ho
Department of Astronomy, Columbia University, New York, NY, 10027

Introduction



Bayesian inference: $p(\Theta \mid \mathbf{x}_1, \cdots, \mathbf{x}_n) \sim \prod_i p(\mathbf{x}_i \mid \Theta)\, p(\Theta)$

- Simulation-based inference (SBI) leverages neural networks to learn posterior distributions directly from simulation data.
- Simulators for large-scale public campaigns such as CAMELS choose to simulate via a Latin hypercube bounded within a specific region of interest.



Problem: Without gradient information at the boundaries, our flow-based density estimator cannot learn the correct density transitions near the cutoff.

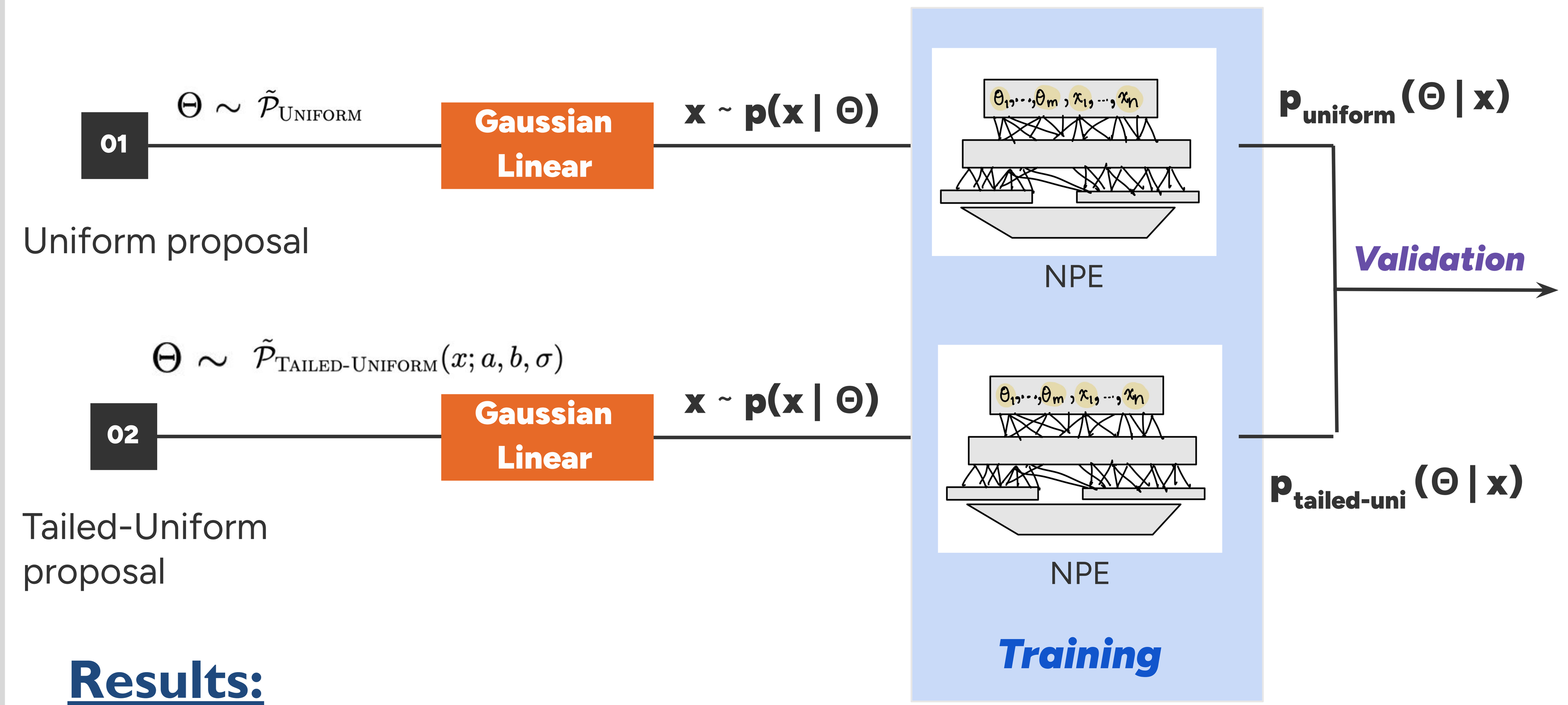
$$\tilde{\mathcal{P}}_{\text{TAILED-UNIFORM}}(x; a, b, \sigma) = \begin{cases} A \cdot \mathcal{N}(a, \sigma^2), & x \leq a \\ B \cdot \mathcal{U}(a, b), & x \in [a, b] \\ A \cdot \mathcal{N}(b, \sigma^2), & x \geq b \end{cases}$$

By feeding the network training samples from outside the uniform region (the extended Gaussian support beyond the original grid)

- Tailed-Uniform captures the posterior density outside the assumed prior bounds, and
- Tailed-Uniform learns a smooth, well-calibrated posterior that extends into boundary-adjacent test points.

Toy Problem

2D Gaussian simulator with an assumed Gaussian prior



Results:

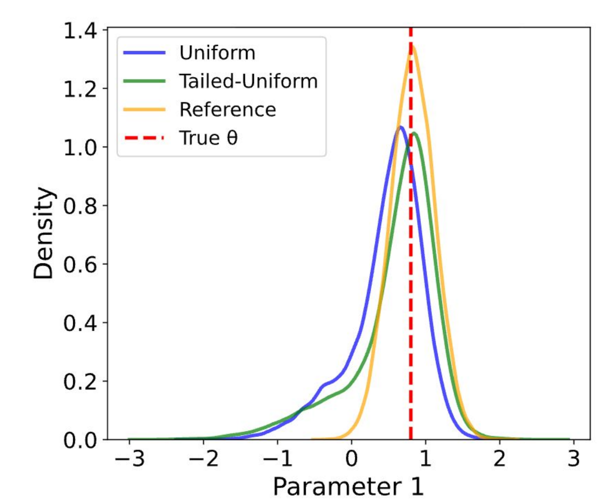
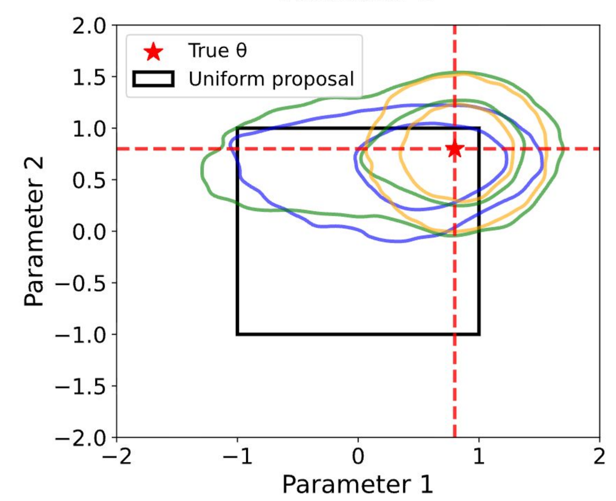
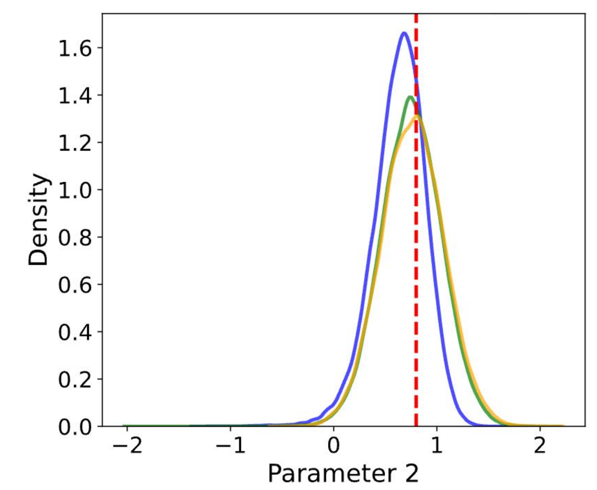
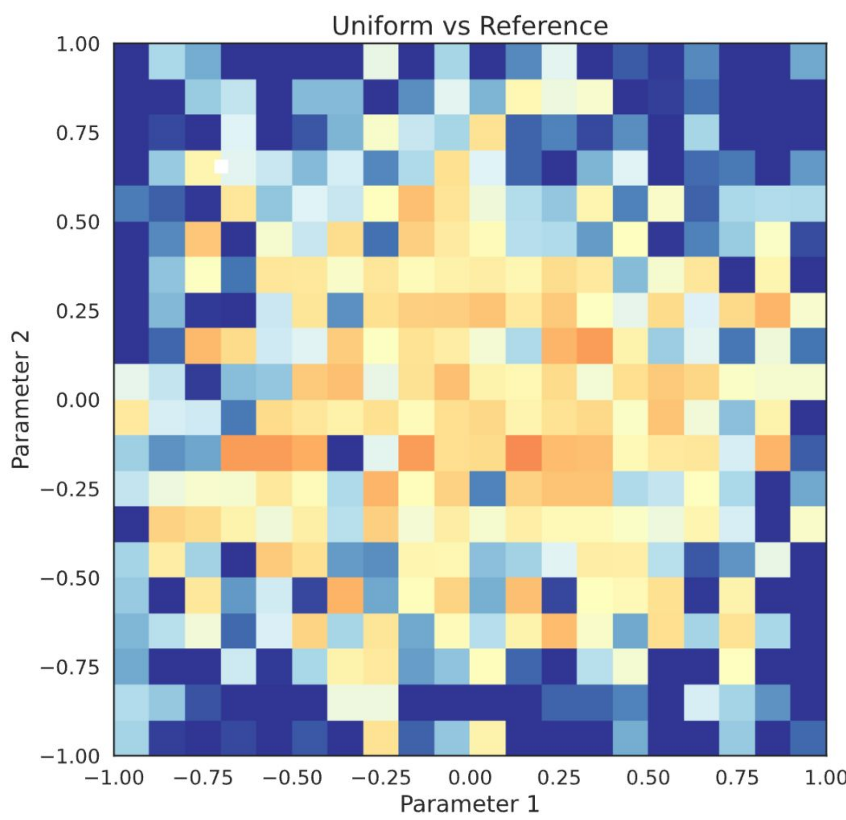


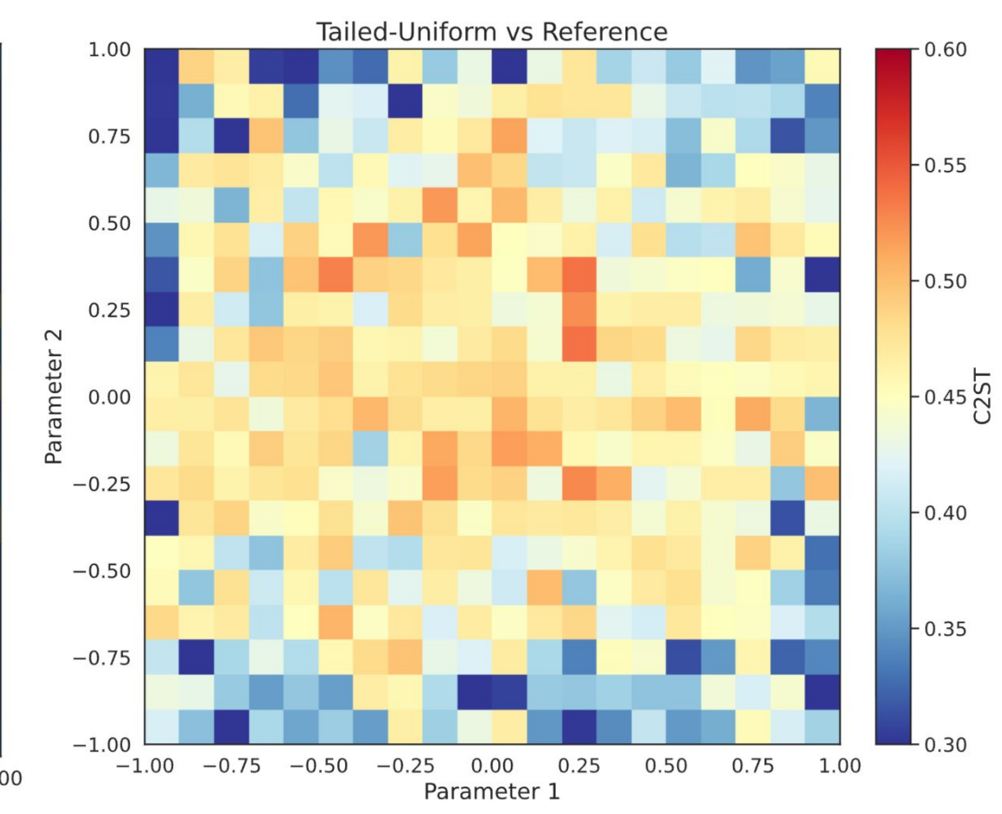
TABLE I. Posterior Performance for $\boldsymbol{\theta}_{\text{true}} = (0.8, 0.8)$.

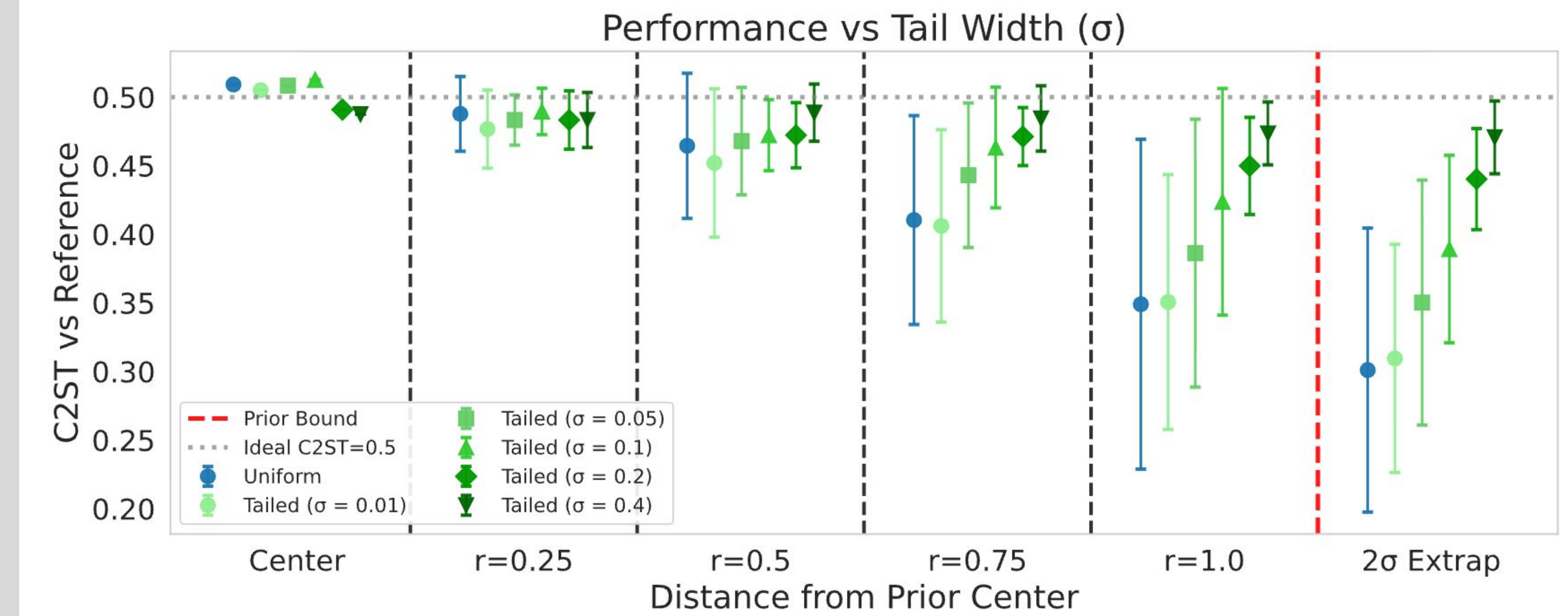
Method	Parameter Estimates θ_1	Parameter Estimates θ_2	C2ST
Analytical Reference	0.825 ± 0.301	0.763 ± 0.299	0.500
TAILED-UNIFORM	0.568 ± 0.590	0.777 ± 0.279	**0.417**
UNIFORM	0.469 ± 0.519	0.644 ± 0.286	0.329
C2ST Improvement			+27%

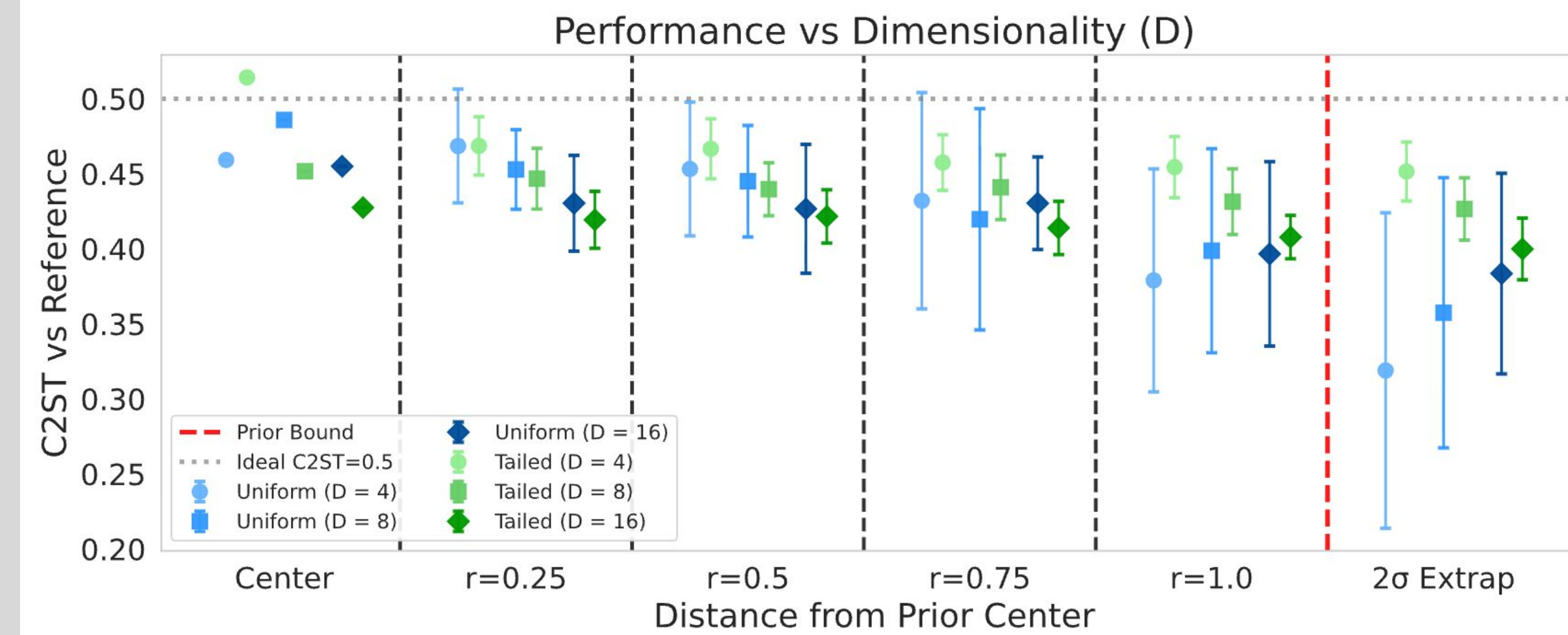






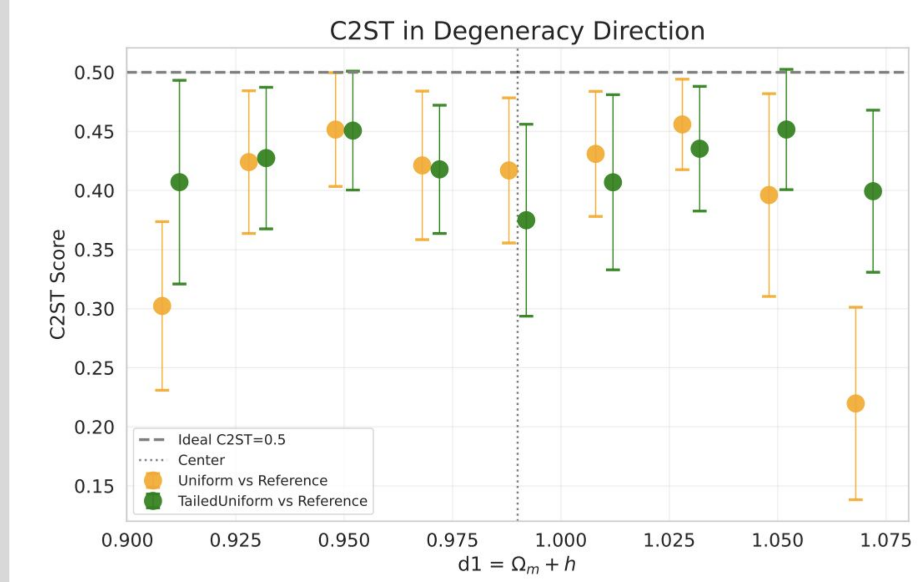


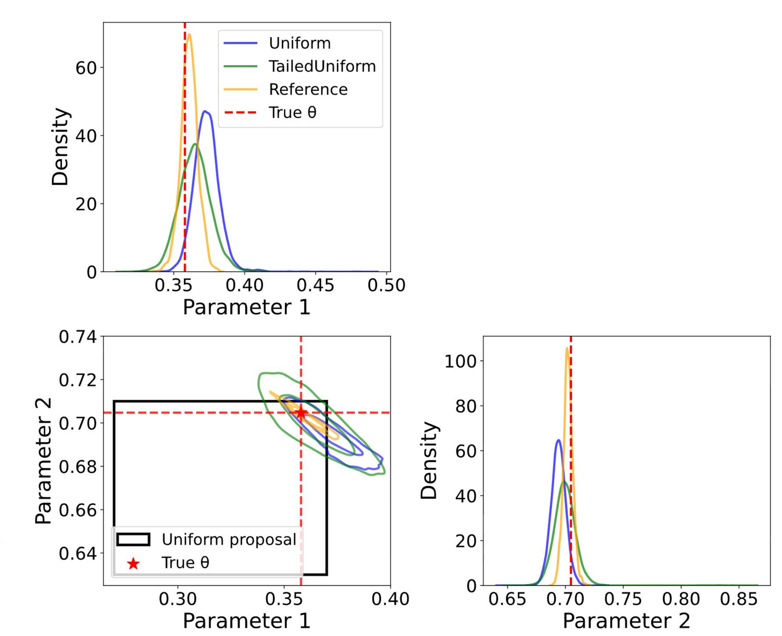


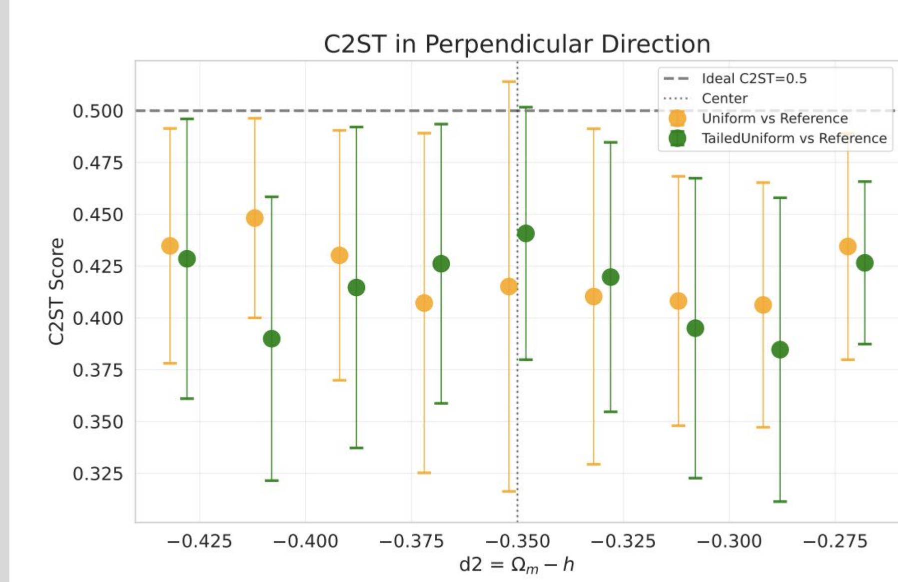


Science Experiments

Goal: Want to infer (Ω_m, h) from observations generated using the syren-new emulator with heteroskedastic cosmic noise.







- Degeneracy direction (d1 = Ωm + h): Uniform degrades at both ends; Tailed-Uniform is stable.
- Perpendicular direction (d2 = Ωm −h): both methods perform about the same