Efficient Bayesian Methods for Posterior Sampling and Evidence Estimation

He Jia

Peking University

NAOC Workshop, Beijing | Dec 29, 2019

Bayes' Theorem

▷ Given model $\mathcal{M}$, model parameters θ and data D,

$$\underbrace{P(\theta|D,\mathcal{M})}_{\text{Posterior}} = \frac{\overbrace{P(D|\theta,\mathcal{M})}^{\text{Likelihood}}\overbrace{P(\theta|\mathcal{M})}^{\text{Prior}}}{\underbrace{P(D|\mathcal{M})}_{\text{Evidence}}}$$

How to get information from the posterior?

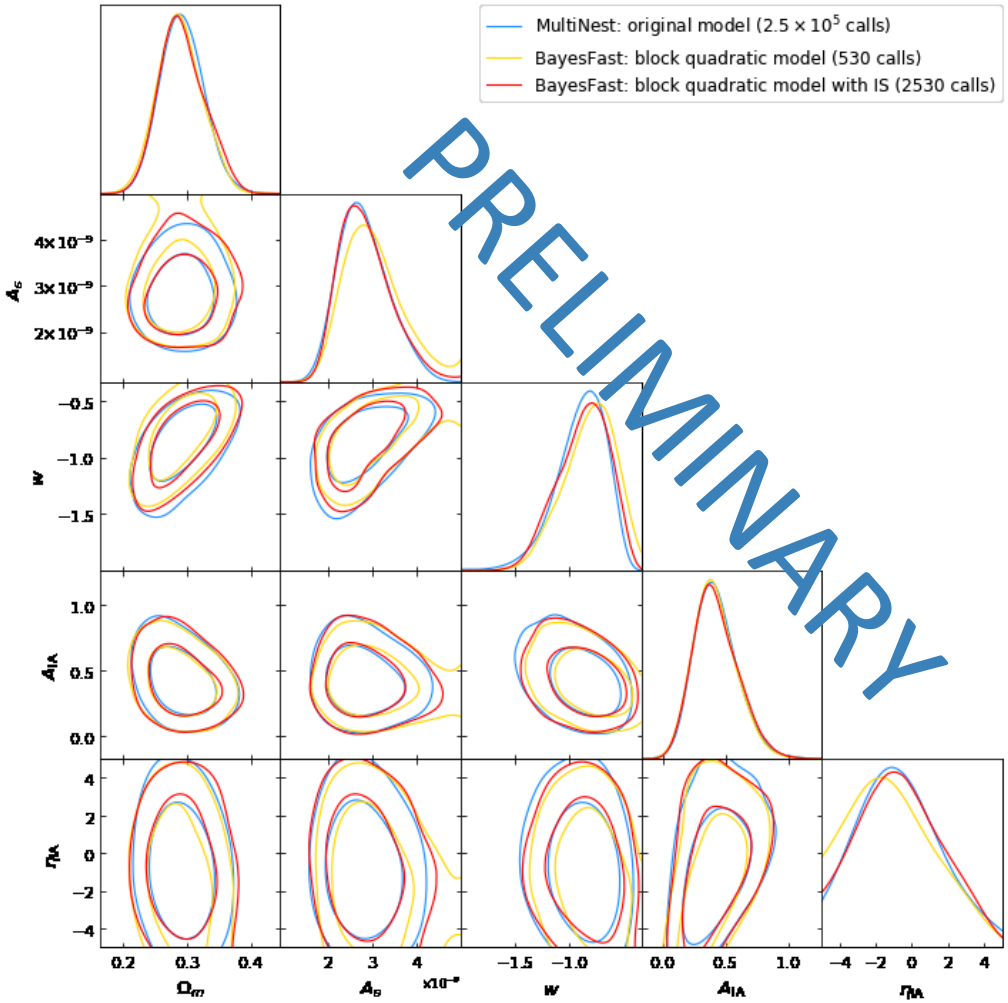
- Say we want to know $\langle f(\theta) \rangle_p$ given the posterior $p(\theta)$
- Naïve integration: $\mathcal{O}(e^D)$
- In practice, we need to first get samples from $p(\theta)$
- MCMC: typically 10^5 evaluations for 10-d, 10^6 for 20-d

Proposed Method

- Usually, we can write the likelihood as $\mathcal{L}(m(\theta))$
- Only $m(\theta)$ is expensive, so we approximate it with appropriate surrogate functions
- We found quadratic / cubic polynomials sufficient for most cosmological examples
- We can add importance weights $p(\theta)/q(\theta)$ to retrieve asymptotic consistency

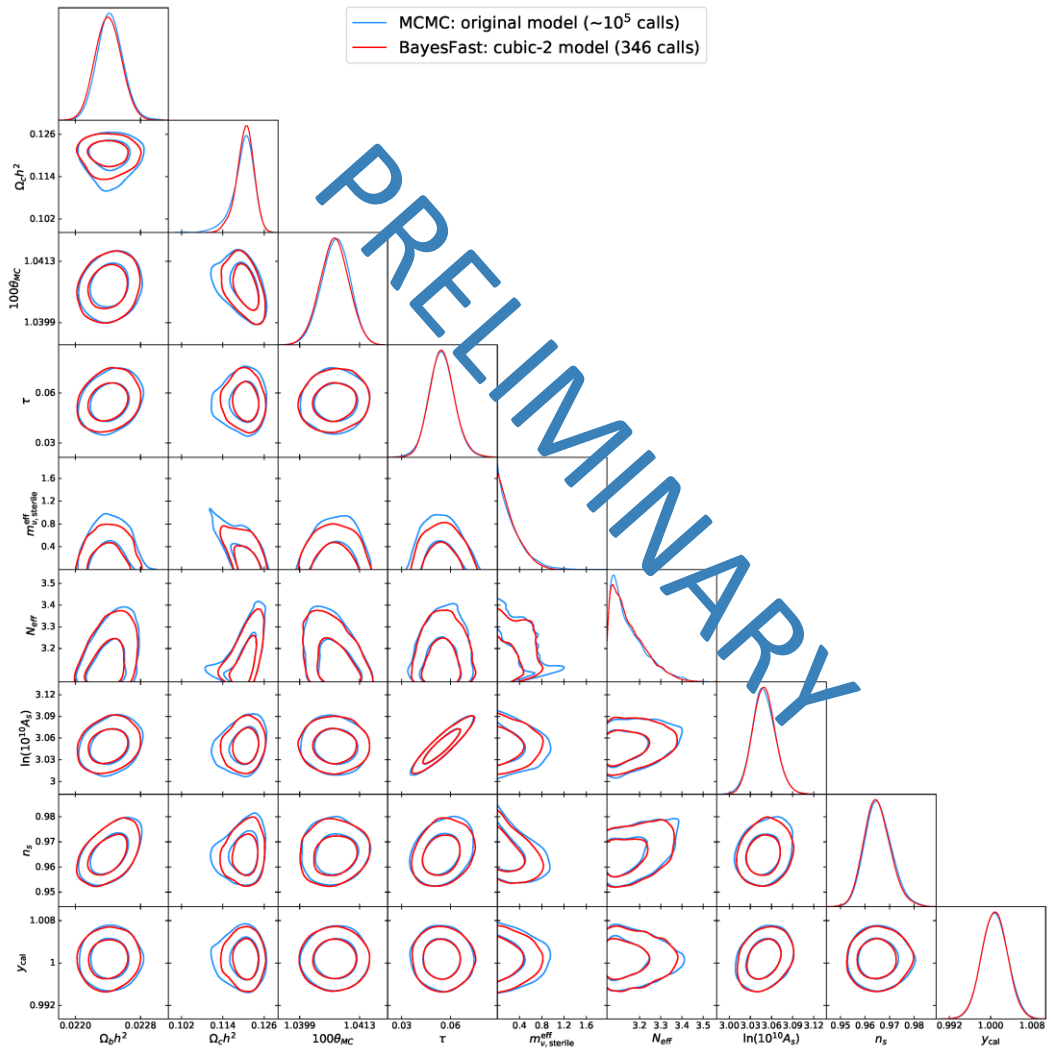
He Jia and Uros Seljak, *BayesFast: A Fast and Scalable Method for Cosmological Bayesian Inference*, in prep.

Results: DES Year 1



He Jia and Uros Seljak, *BayesFast: A Fast and Scalable Method for Cosmological Bayesian Inference*, in prep.

Results: Planck 2018



He Jia and Uros Seljak, *BayesFast: A Fast and Scalable Method for Cosmological Bayesian Inference*, in prep.

Bayesian Evidence

- Definition:

$$p(d|\mathcal{M}) \equiv \int_{\Omega_{\mathcal{M}}} p(d|\theta, \mathcal{M}) p(\theta|\mathcal{M}) \mathrm{d}\theta \quad \text{(Bayesian evidence)}.$$

- Empirical rule for model selection:

$\lvert \ln B_{01} \rvert$	Odds	Probability	Strength of evidence
< 1.0	$\lesssim 3:1$	< 0.750	Inconclusive
1.0	$\sim 3:1$	0.750	Weak evidence
2.5	$\sim 12:1$	0.923	Moderate evidence
5.0	$\sim 150:1$	0.993	Strong evidence

Proposed Method

- Bridge Sampling:

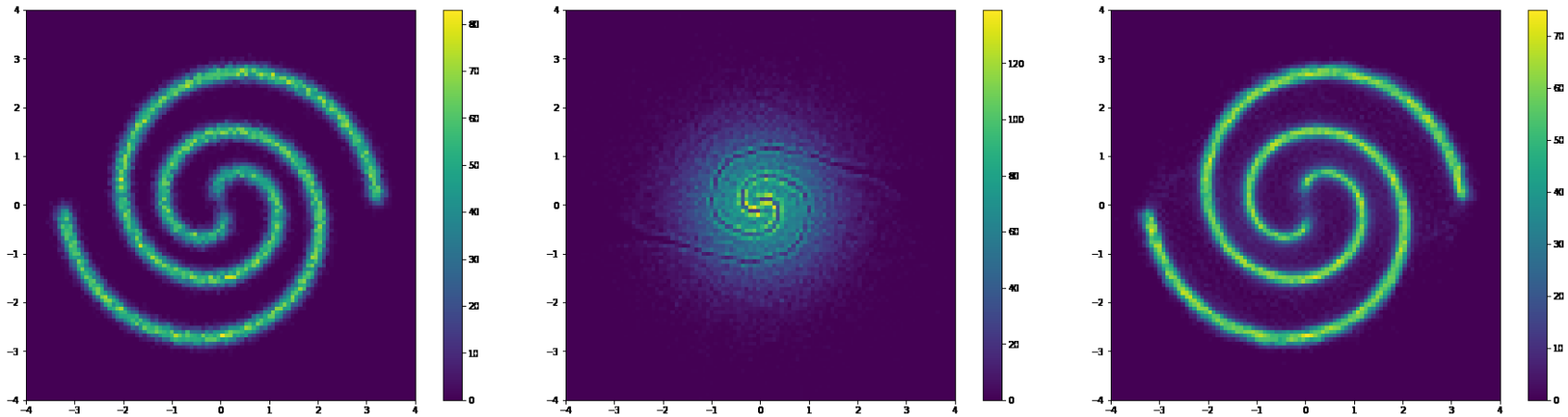
$$\int_{\Omega} \alpha(\boldsymbol{x}) p(\boldsymbol{x}) q(\boldsymbol{x}) d\boldsymbol{x} = \mathcal{Z}_p \left\langle \alpha(\boldsymbol{x}) q(\boldsymbol{x}) \right\rangle_p = \mathcal{Z}_q \left\langle \alpha(\boldsymbol{x}) p(\boldsymbol{x}) \right\rangle_q ,$$

- Choice of $\alpha(x)$: optimal bridge function can be constructed
- Choice of $q(x)$: approximated from the posterior samples

He Jia and Uros Seljak, *Normalizing Constant Estimation with Gaussianized Bridge Sampling*, arXiv:1912.06073.

Proposed Method

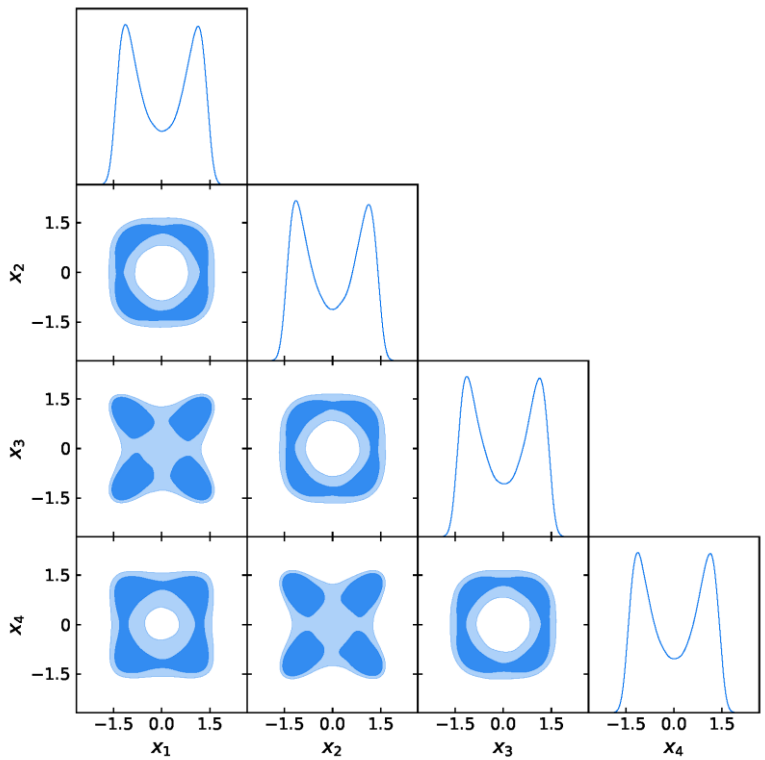
▷ Iterative Neural Transform (INT):



Original Gaussianized Generated

The details will be presented in another paper.

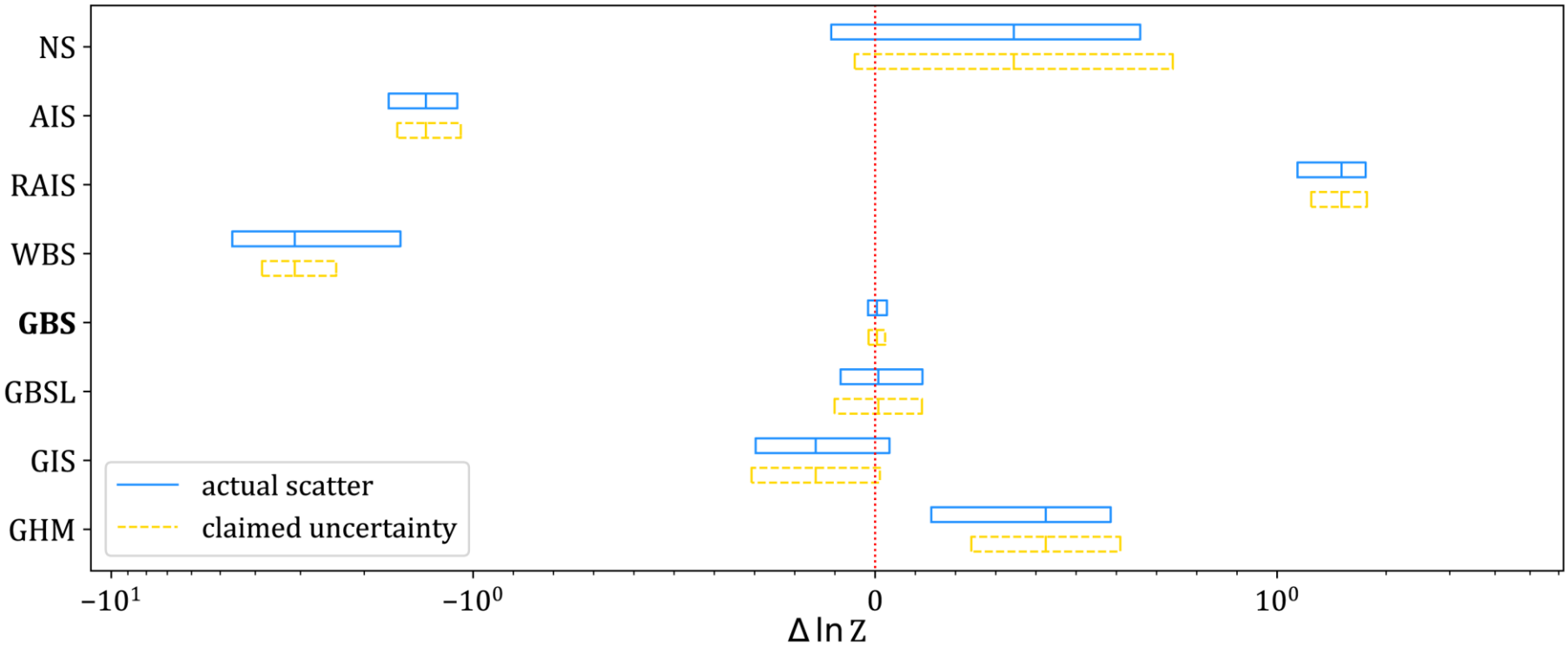
Ring Example: Corner Plot



The example is in 64 dim. Here we only show the first 4 dim.

He Jia and Uros Seljak, *Normalizing Constant Estimation with Gaussianized Bridge Sampling*, arXiv:1912.06073.

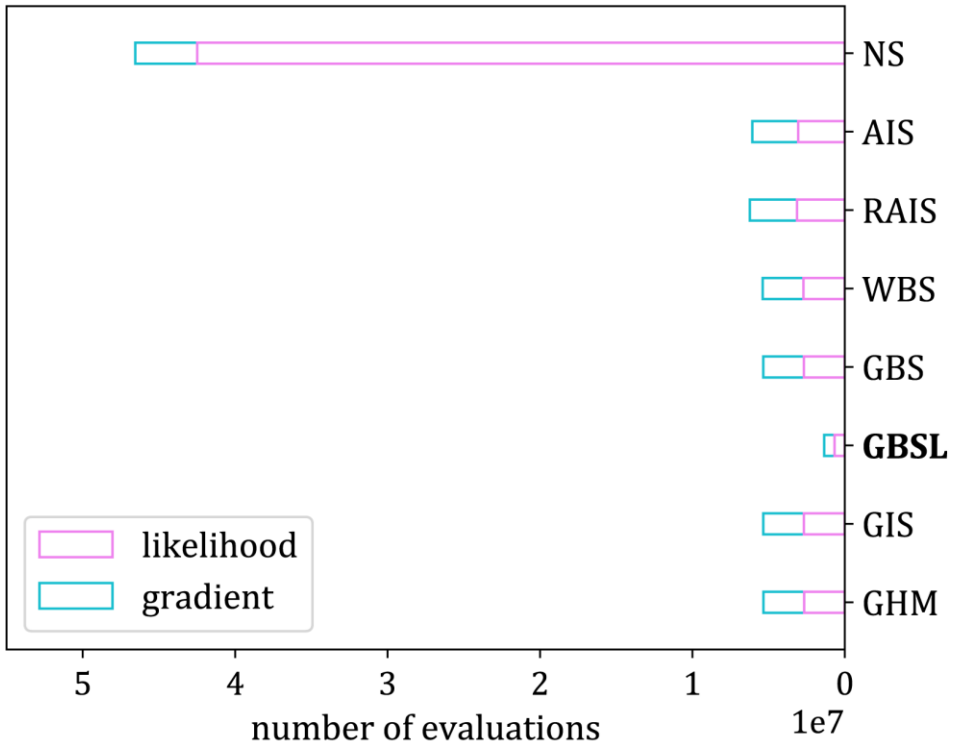
Ring Example: Accuracy



Abbreviation keys: NS, Nested Sampling; AIS, Annealed Importance Sampling; RAIS, Reversed Annealed Importance Sampling; WBS, Warp Bridge Sampling; GBS, Gaussianized Bridge Sampling; GBSL, Gaussianized Bridge Sampling Lite; GIS, Gaussianized Importance Sampling; GHM, Gaussianized Harmonic Mean.

He Jia and Uros Seljak, *Normalizing Constant Estimation with Gaussianized Bridge Sampling*, arXiv:1912.06073.

Ring Example: Computation Cost



Abbreviation keys: NS, Nested Sampling; AIS, Annealed Importance Sampling; RAIS, Reversed Annealed Importance Sampling; WBS, Warp Bridge Sampling; GBS, Gaussianized Bridge Sampling; GBSL, Gaussianized Bridge Sampling Lite; GIS, Gaussianized Importance Sampling; GHM, Gaussianized Harmonic Mean.

He Jia and Uros Seljak, *Normalizing Constant Estimation with Gaussianized Bridge Sampling*, arXiv:1912.06073.

Software Development

- BayesFast:
  - https://github.com/HerculesJack/bayesfast
  - Fast posterior sampling with polynomial surrogate models and No-U-Turn Sampler.
  - Fast evidence evaluation with GBS algorithm.
- CosmoFast:
  - https://github.com/HerculesJack/cosmofast
  - Cosmology add-on for BayesFast.

Thanks!

Any questions?