Fermilab | U.S. DEPARTMENT OF ENERGY | Office of Science

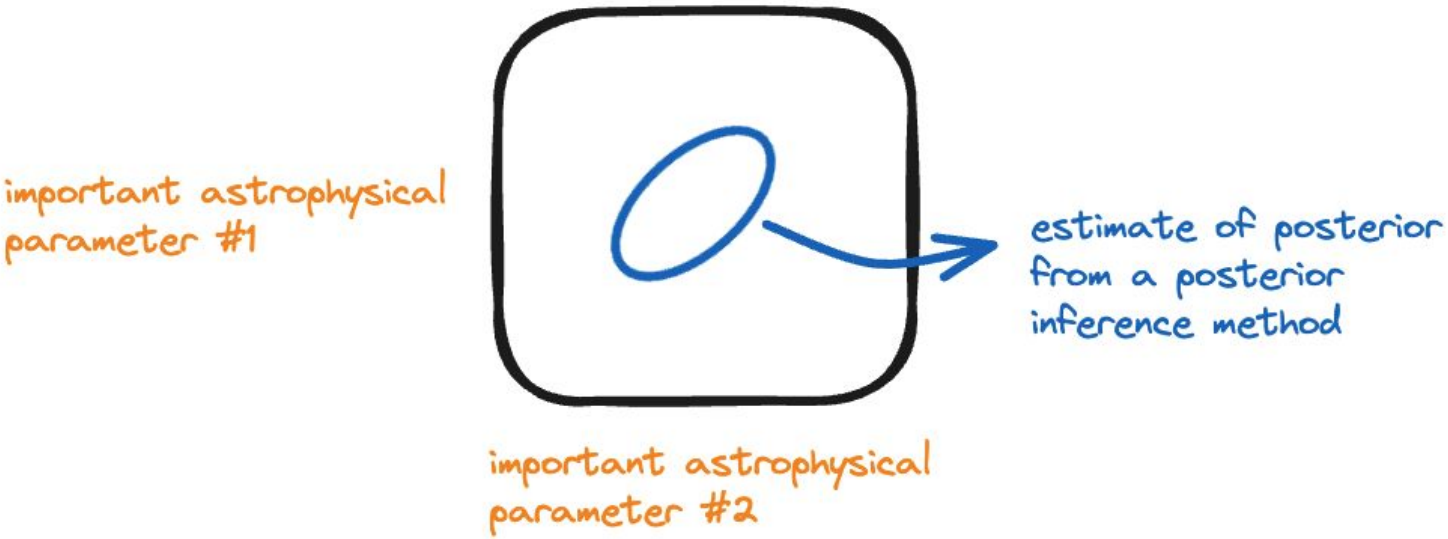
Testing the fidelity of posterior inference methods for astrophysics

Becky Nevin and the DeepSkies Lab

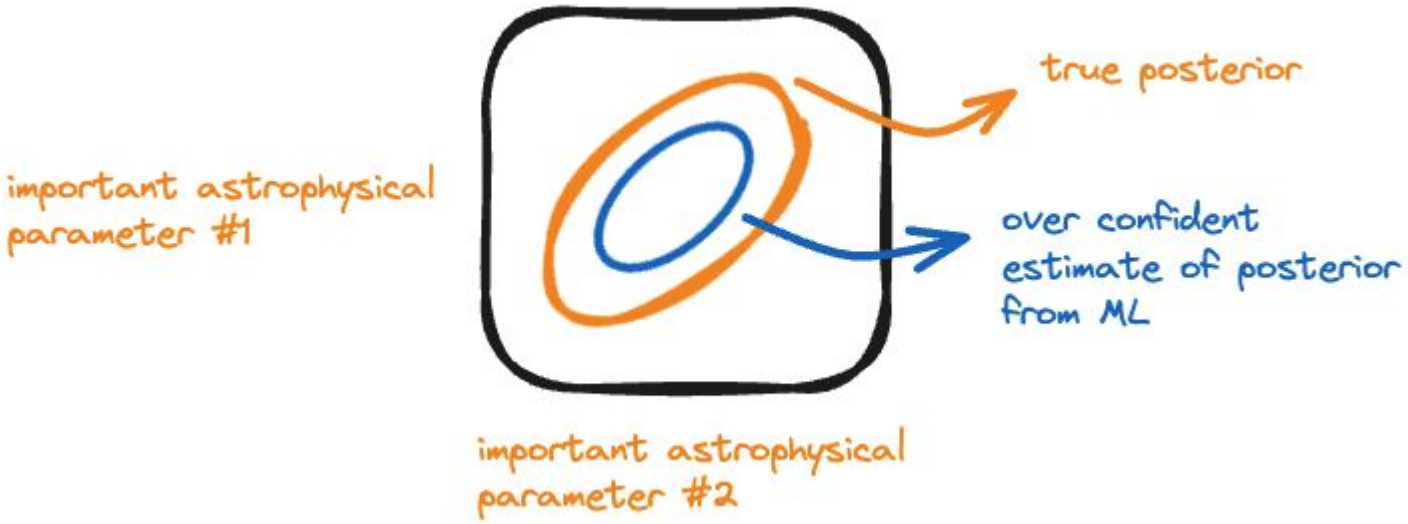
Nov 27, 2023

ML-IAP/CCA Debating the potential of ML in astronomical surveys

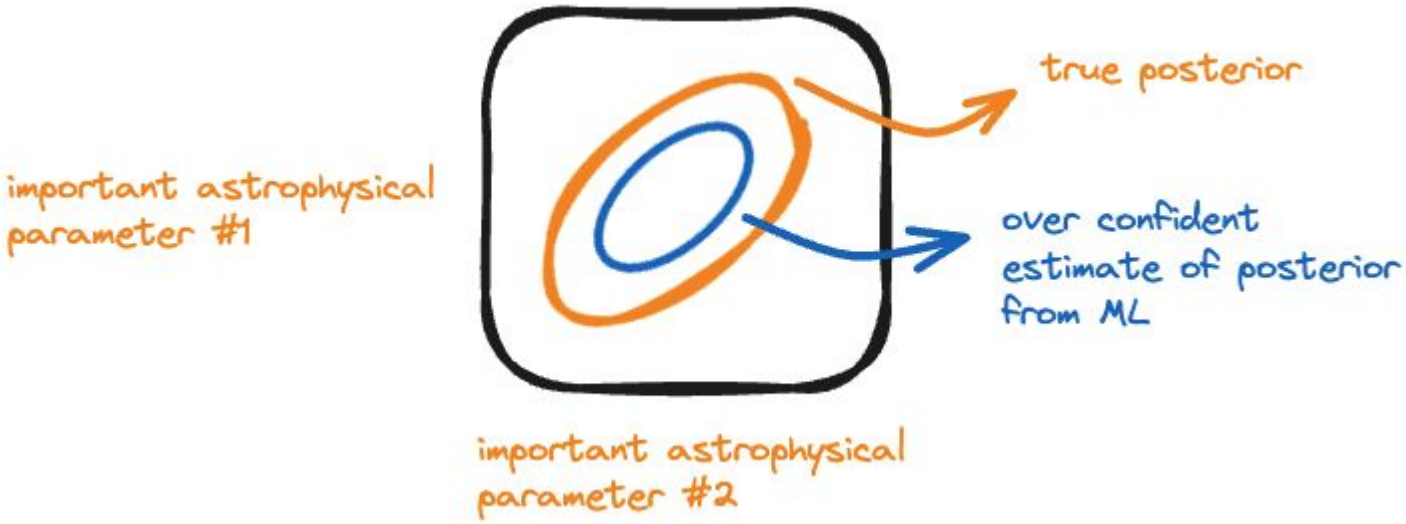
Why test the fidelity of posterior inference methods?



Why test the fidelity of posterior inference methods?



Why test the fidelity of posterior inference methods?

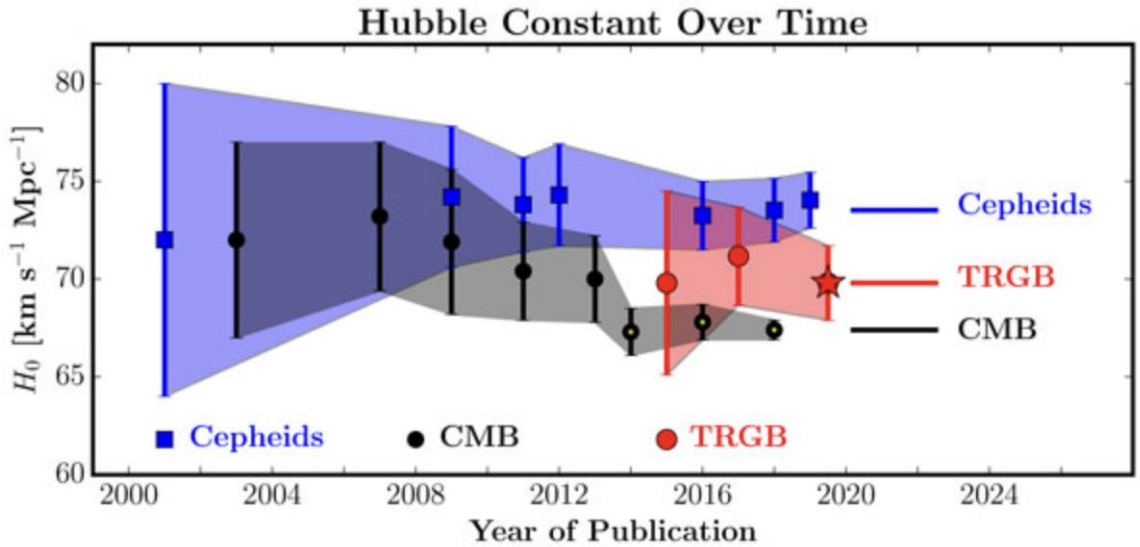


Many ML posterior inference methods are over-confident

([A crisis in simulation-based inference])

What does it mean to **test the fidelity** of these methods?

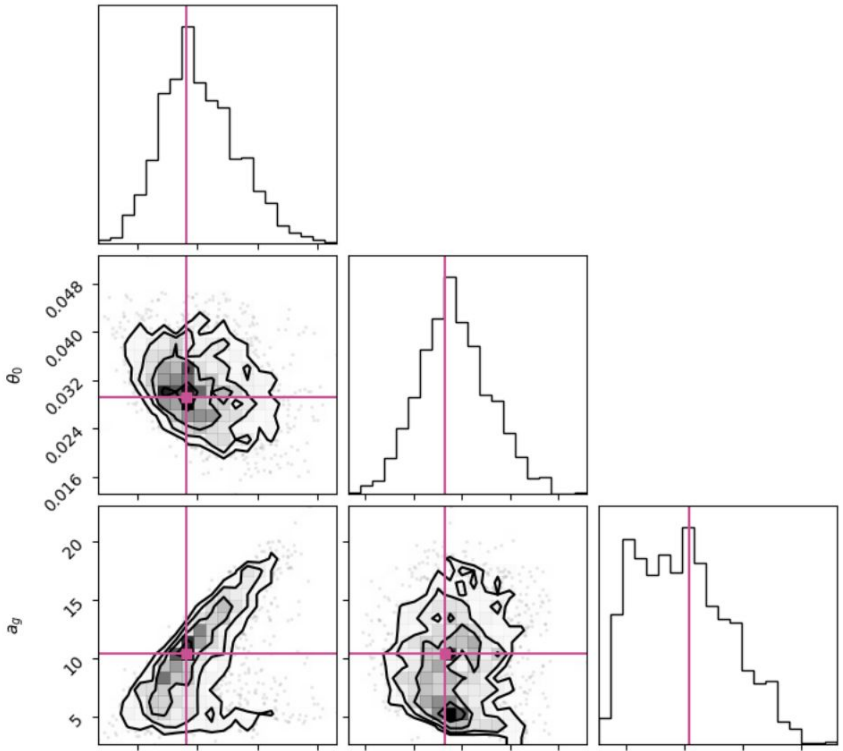
Error is important in astrophysics, people's entire careers depend upon / define these error bars:



Testing the fidelity of **posterior inference methods** for astrophysics

Posterior inference method = determining possible parameter values that make sense for your data, p(θ|X)

distribution of possible parameter values, θs



Fermilab

Posterior inference methods

Bayesian Inference
Analytic likelihood and HMC sampling

Simulation-based inference
No likelihood needed, uses a density estimator

Deep Ensembles
An amalgamation of neural networks

Posterior inference methods

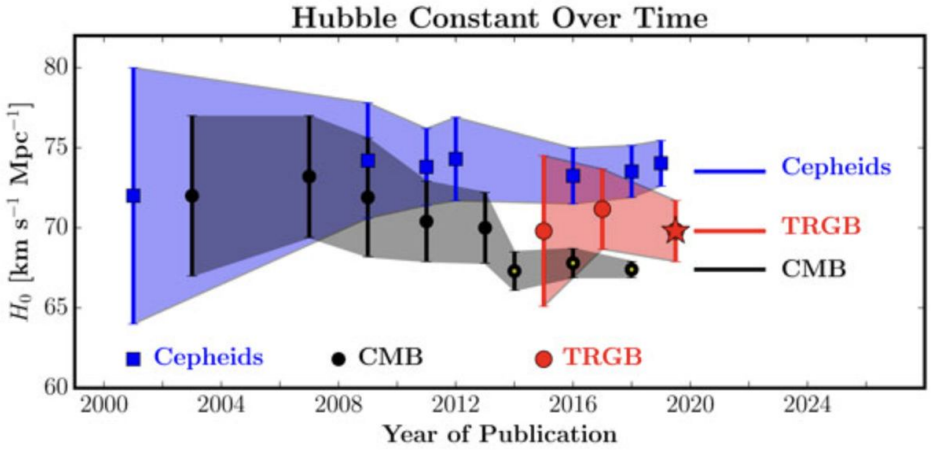
Bayesian Inference + Hierarchical
Analytic likelihood and HMC sampling

Simulation-based inference + Hierarchical
No likelihood needed, uses a density estimator

Deep Ensembles
An amalgamation of neural networks

What does it mean to **test the fidelity** of these methods?

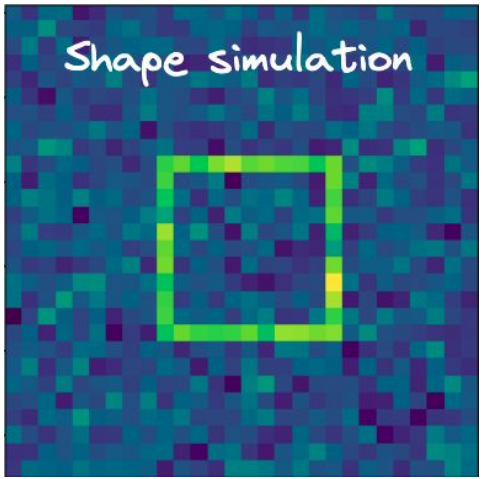
Can we trust the uncertainty prediction of the model?

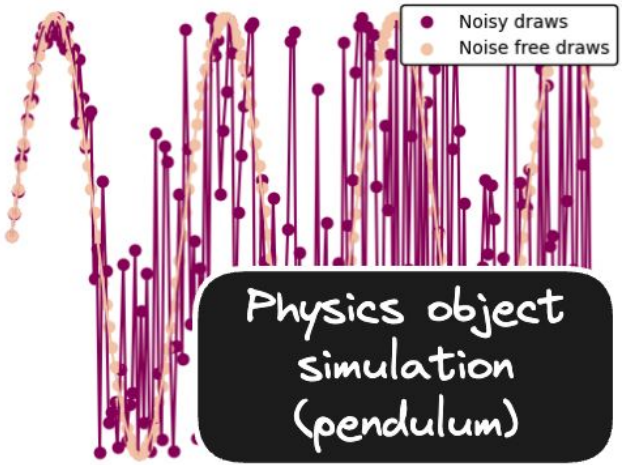


- ^Uncertainty is clearly really important in astrophysics^
- Can we trust the confidence of these models?
- How do these different models perform under different uncertainty conditions? Different types of uncertainty?

We require fine control over uncertainty properties → we use Deepskies' **DeepBench** software

https://github.com/deepskies/DeepBench





Testing the fidelity of posterior inference methods for astrophysics

- Comparing posterior inference methods (Bayesian inference, simulation-based inference, and deep ensembles)
- Developing software tools to assess uncertainty prediction from these methods