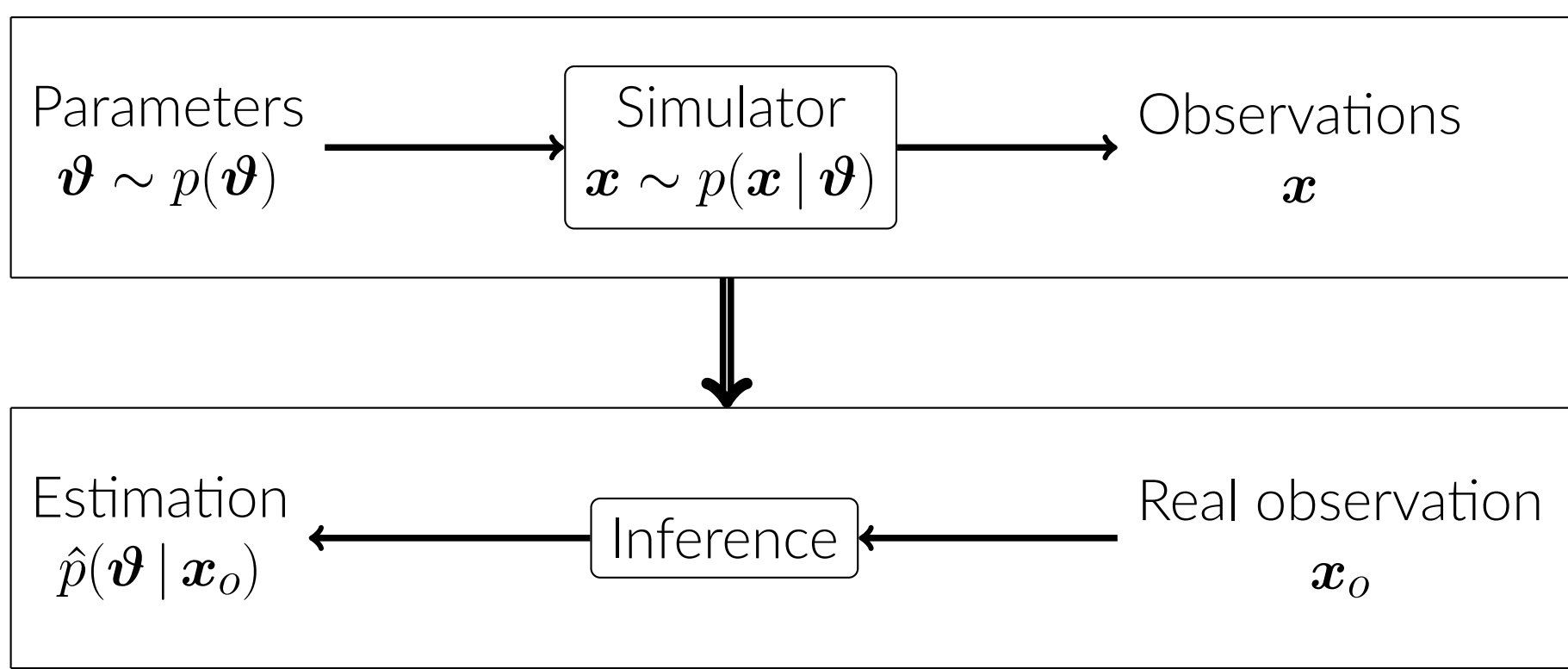
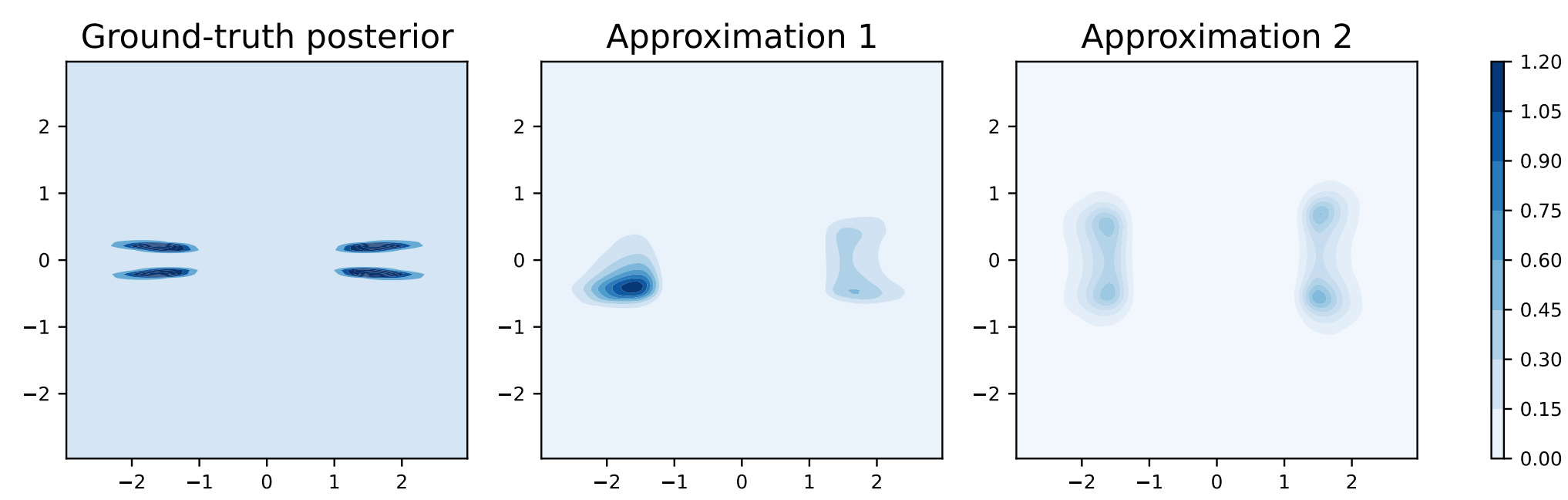


Simulation-based inference



Motivation



The posterior density obtained with the first approximation is sharper than the ground-truth posterior density. Consequently, **it may exclude parameter values that are actually plausible and drive the scientific inquiry in the wrong direction.** A better approximation would be the second one which is flatter and do not reject any plausible value.

Expected coverage

The expected coverage is expressed as

$$\text{expected coverage}(\hat{p}, \alpha) = \mathbb{E}_{p(\vartheta, x)} \left[1 \left[\vartheta \in \Theta_{\hat{p}(\vartheta | x)}(1 - \alpha) \right] \right],$$

where the function $\Theta_{\hat{p}(\vartheta | x)}(1 - \alpha)$ yields the $1 - \alpha$ highest posterior density region of $\hat{p}(\vartheta | x)$.

A **conservative model** is a model such that

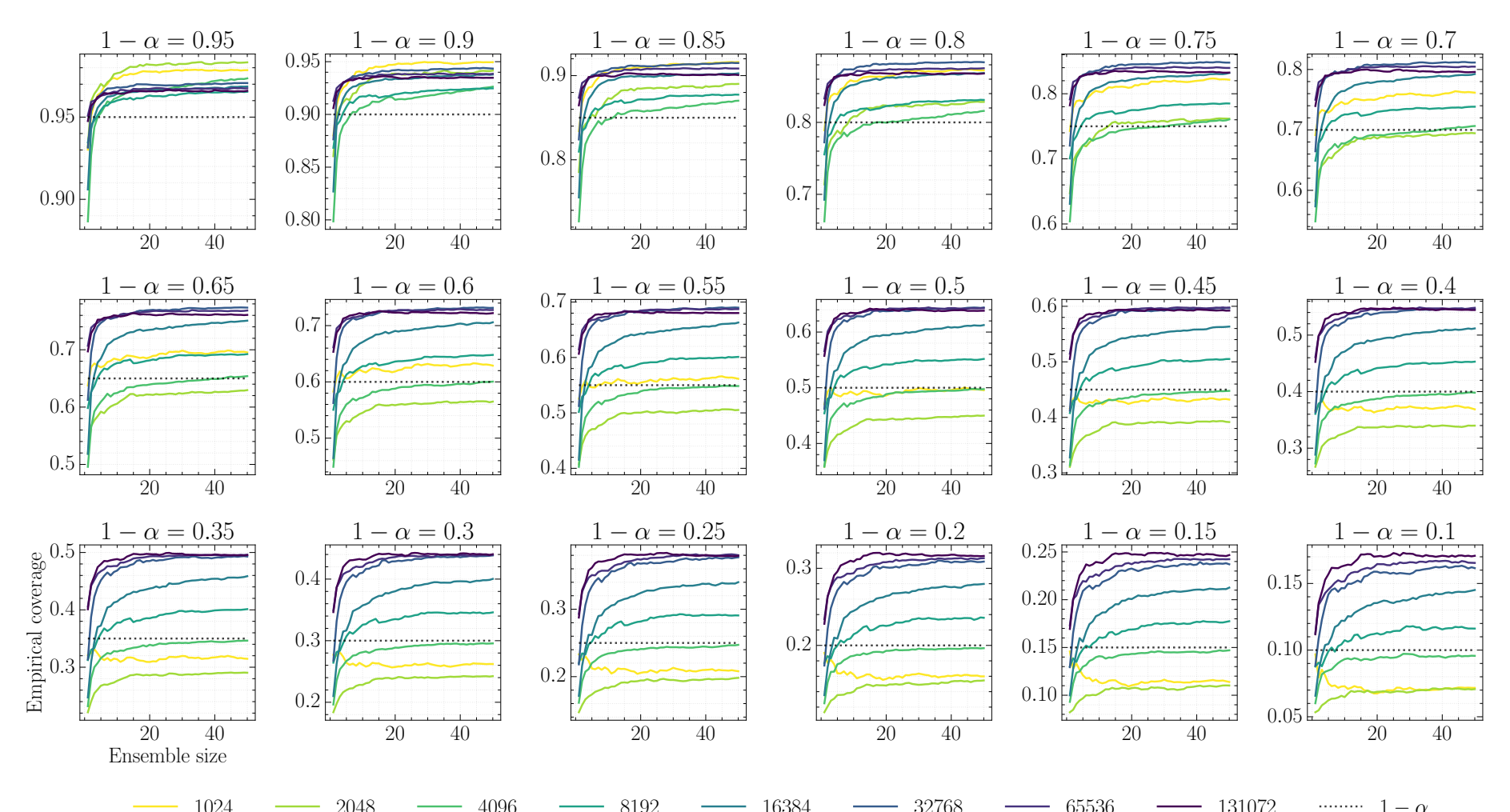
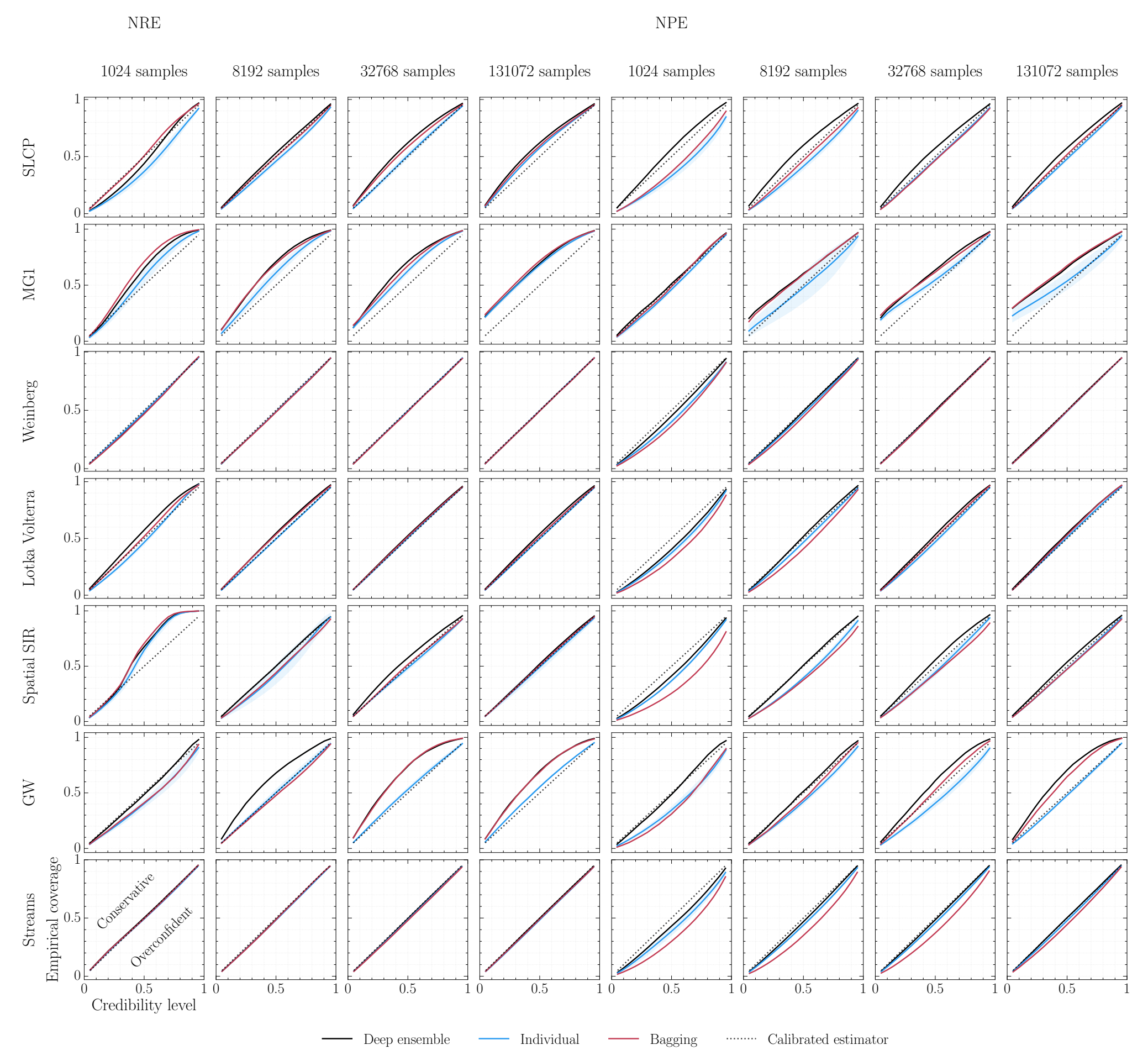
$$\text{expected coverage}(\hat{p}, \alpha) \geq 1 - \alpha, \quad \forall \alpha$$

Amortized vs non-amortized

| Amortized | Non-amortized |
|---|---|
| 1. Build a general model $\hat{p}(\vartheta x)$. | 1. Build a model for $\hat{p}(\vartheta x = x_o)$. |
| 2. Evaluate $\hat{p}(\vartheta x = x_o)$ | 2. Repeat for each observation. |

- Non-amortized algorithms tend to be less conservative than amortized ones.
- Expected coverage computation is expensive for non-amortized approaches since it requires building many posterior models.
- We advise to use amortized algorithms or non-amortized ones that allow local diagnostics.

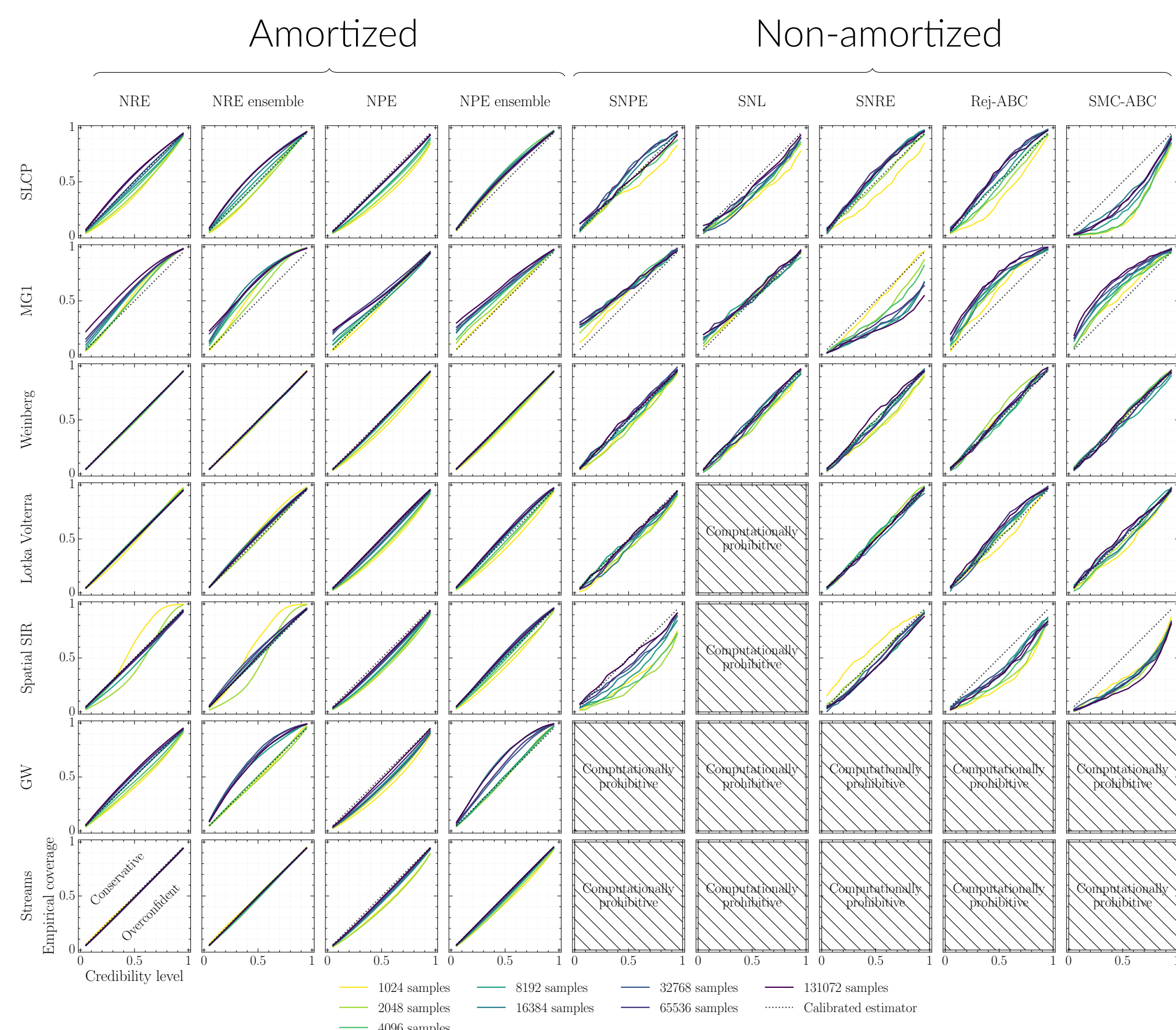
Ensembles



The expected coverage probability of an ensemble model is larger than the average individual model's. The ensemble size positively affects the expected coverage probability.

Take-home messages

- All benchmarked algorithms can produce non-conservative posterior approximations. Consequently, all those algorithms could erroneously exclude plausible parameter values and be potentially misleading.
- Performing diagnostics to identify overconfident posterior approximations is crucial. The use of amortized algorithms or non-amortized ones that allow local diagnostics is then advised.
- Ensembling constitutes an immediately applicable solution to build more conservative approximate posteriors.



All benchmarked algorithms may produce non-conservative posterior approximations.