SAND2022-7253C

Sandia National Laboratories

Data-Driven Model-Form Uncertainty with Bayesian Statistics and Neural Differential Equations

Portone, T., Acquesta, E., Dandekar, R., Rackauckas, C.

ECCOMAS 2022, June 2022, Oslo, Norway

U.S. DEPARTMENT OF ENERGY

NNSA National Nuclear Security Administration

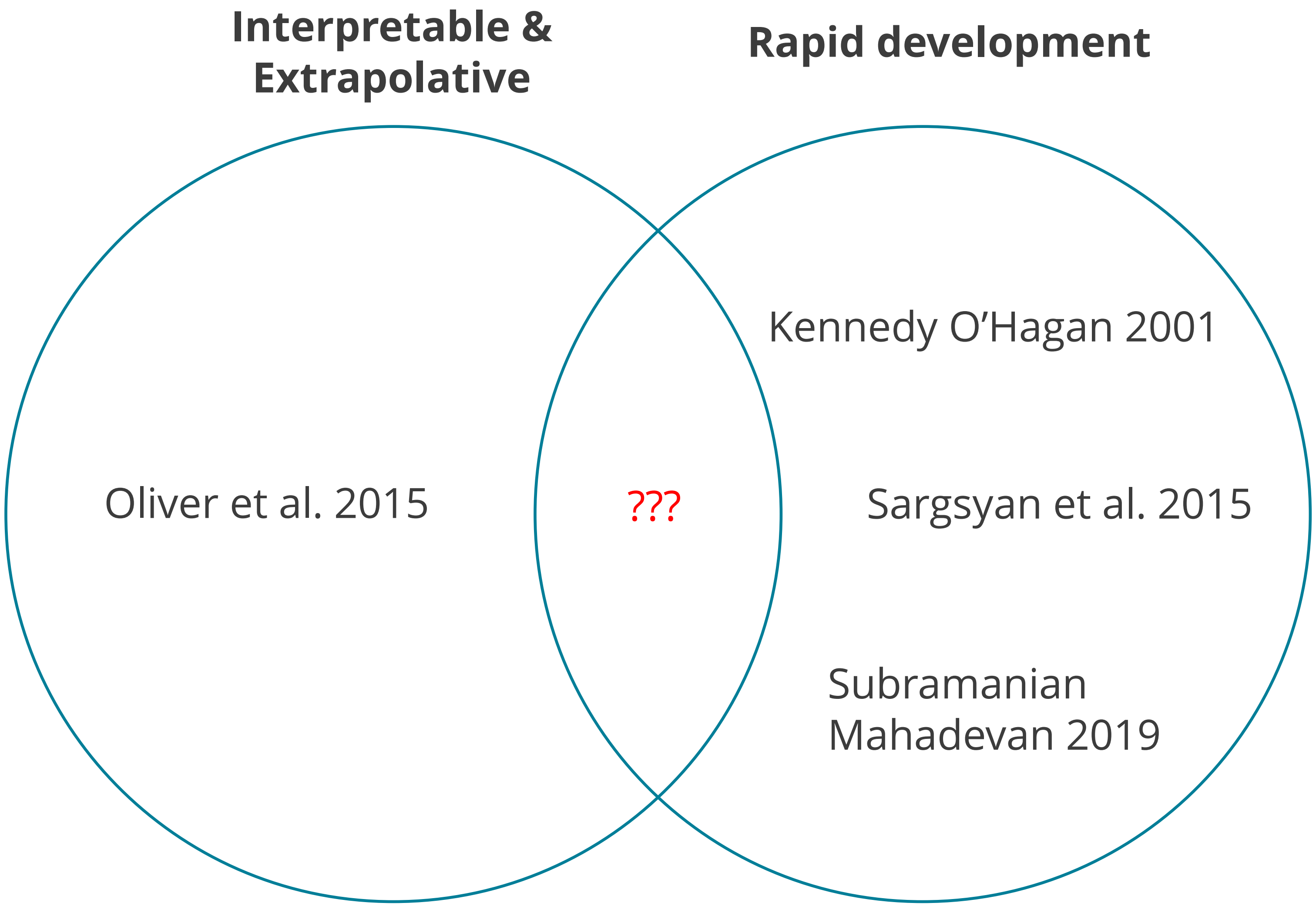
Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Motivation: Develop an

- agile
- interpretable
- extrapolative

method to represent model-form uncertainty (MFU) for predictions



Idea: combine universal differential equations (UDEs) with Bayesian statistics to represent MFU

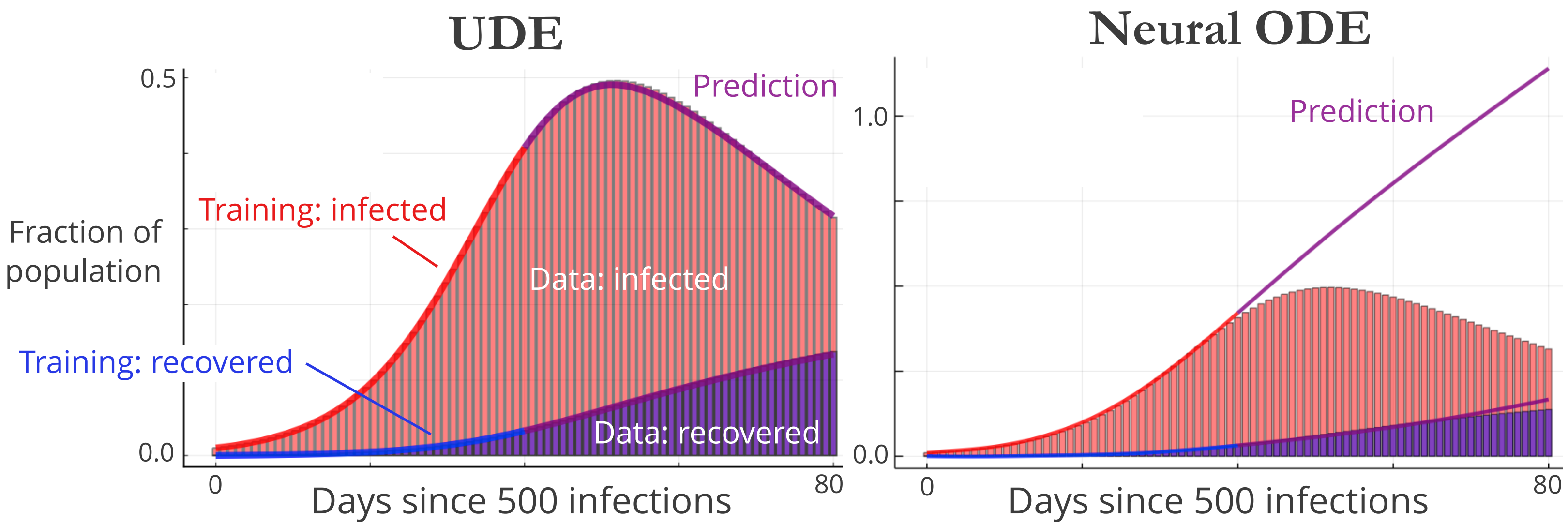
UDEs embed ML models, e.g. neural nets (NNs) within existing scientific models:

$$\boldsymbol{u}' = F\big(\boldsymbol{u}, t, NN_\theta(\boldsymbol{u})\big)$$

$$\min_\theta \|\boldsymbol{d} - \boldsymbol{u}(\theta)\|$$

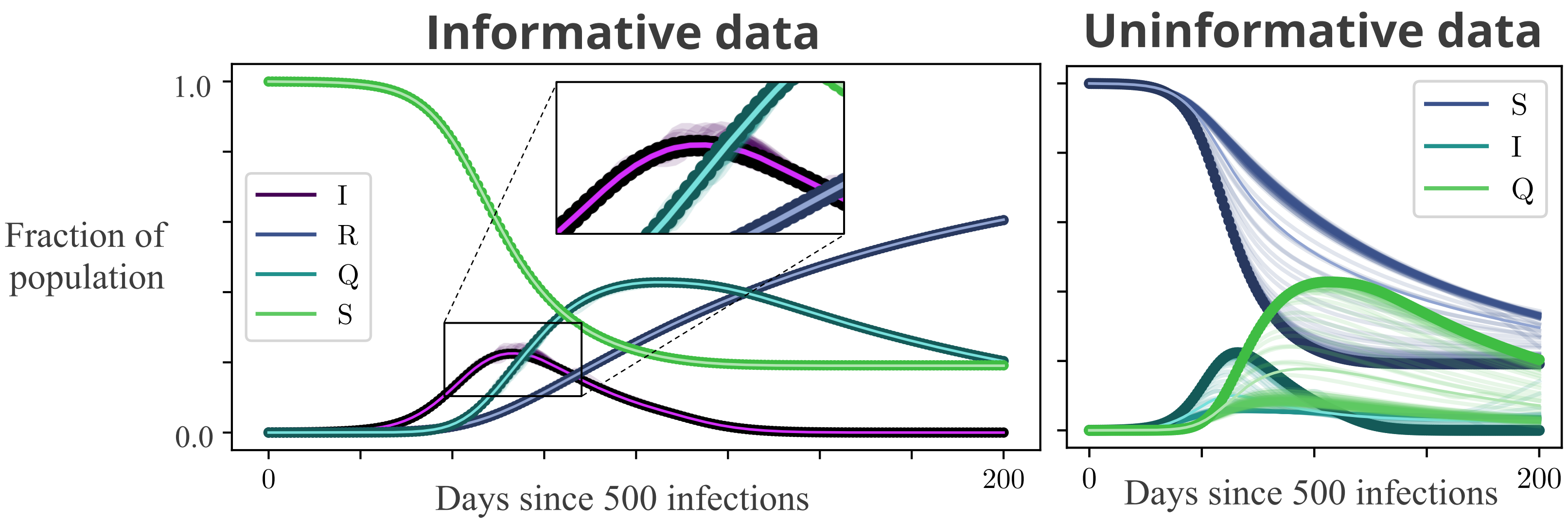
- Data driven, **BUT**
- Time-independent parameterization
- Can respect physical principles by construction
- Can be more predictive than Neural ODEs:

$$\boldsymbol{u}' = NN_\theta(\boldsymbol{u})$$



Idea: combine universal differential equations (UDEs) with Bayesian statistics to represent MFU

- UDEs used in a deterministic setting to find "model corrections" or "missing physics."
- **Problem:** If data does not adequately inform the UDE, there is no single correction: the appropriate model form is uncertain.



By endowing the NN with a Bayesian parameterization, can we use them to represent model-form uncertainty?

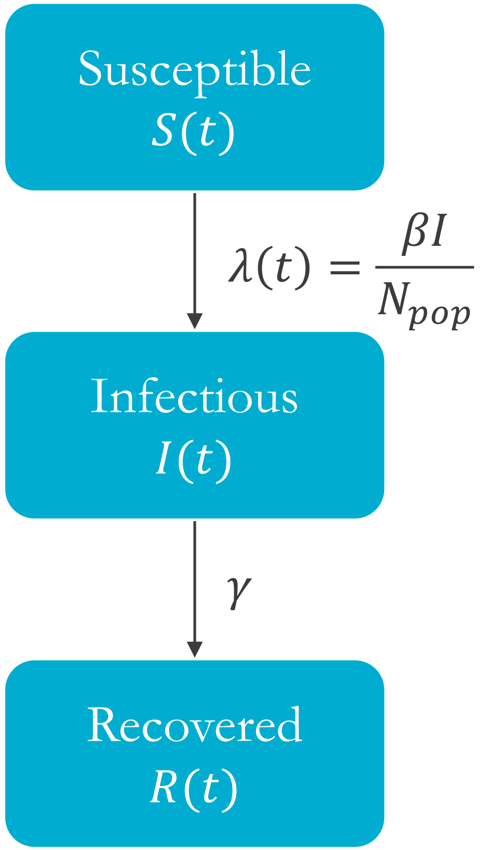
Exemplar problem: compartment-based disease models

Let $N_{pop} = S(t) + I(t) + R(t)$.

$$\frac{dS}{dt} = -\frac{\beta I S}{N_{pop}}$$

$$\frac{dI}{dt} = \frac{\beta I S}{N_{pop}} - \gamma I$$

$$\frac{dR}{dt} = \gamma I$$



- SIR a common, simple model of disease spread.
- Doesn't account for infected population quarantine as we saw for COVID-19.
- Quarantine dynamics could depend nonlinearly on state variables.

Exemplar problem: compartment-based disease models

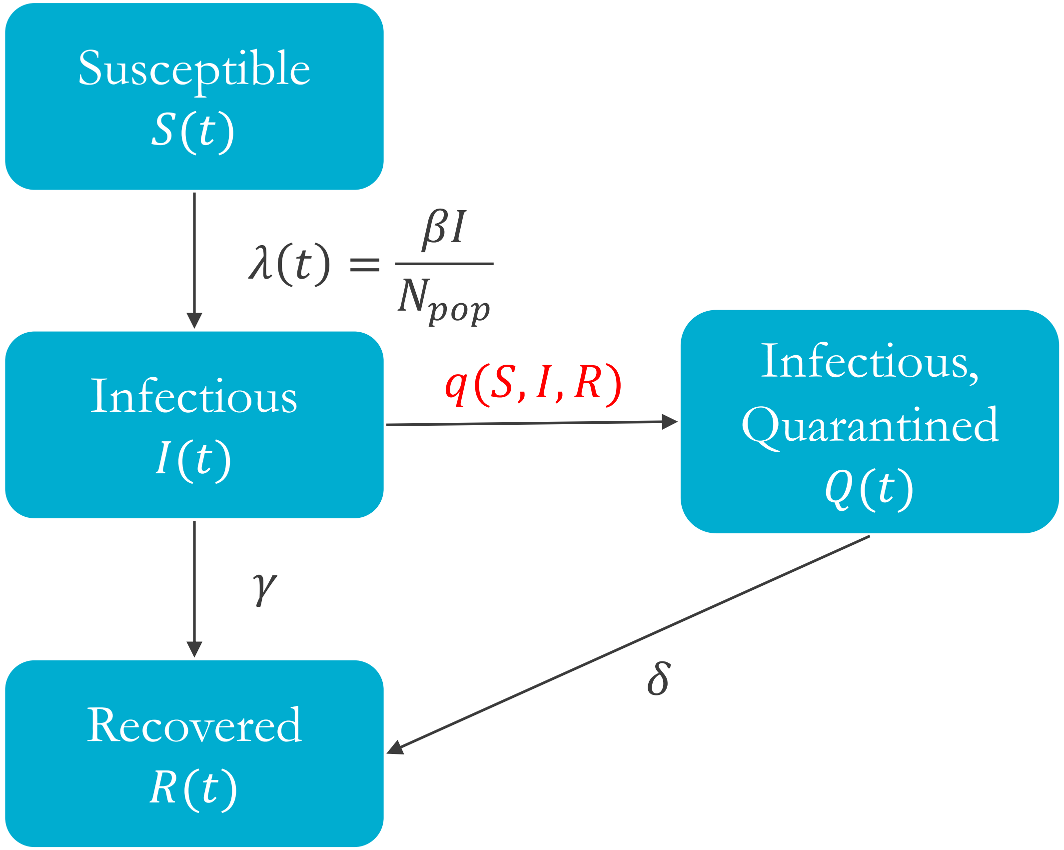
Let $N_{pop} = S(t) + I(t) + R(t) + Q(t)$.

$$\frac{dS}{dt} = -\frac{\beta I S}{N_{pop}}$$

$$\frac{dI}{dt} = \frac{\beta I S}{N_{pop}} - \gamma I - q(S, I, R) I$$

$$\frac{dR}{dt} = \gamma I + \delta Q$$

$$\frac{dQ}{dt} = q(S, I, R) I - \delta Q$$



Represent nonlinear transition into quarantine with a small neural network.

Constrained to conserve population by construction.

Bayesian UDEs

Limitations of deterministic UDEs to address model-form uncertainty:

- Data not always informative enough to identify a single ”model correction”
- There may not exist a deterministic model correction that depends only on modeled states

Idea: define a Bayesian probabilistic representation of NN parameterization to represent uncertainty in model form.

Bayesian UDEs

Challenges:

- NNs notoriously challenging to train even in deterministic setting.
- Traditional Bayesian methods computationally challenging & suffer from curse of dimensionality.

Are Bayesian UDEs a feasible approach to representing MFU?

Initially explored this question in the context of a Bayesian NN embedded in the SIRQ model.

Bayesian NN study

Inferring disease parameters $[\beta, \gamma, \delta]$ along with NN parameters

Prior

- Disease parameters ~ $U(0,2)$
- 51 NN parameters ~ $N(0, (50)^2)$

- **Likelihood**
- Synthetic data generated from SIRQ model
- Calibration data = observations of I, R, Q first 50 days after 500 infections.
- No noise added; likelihood assumes 95% confidence bound of $\pm 10\%$ error, i.e.

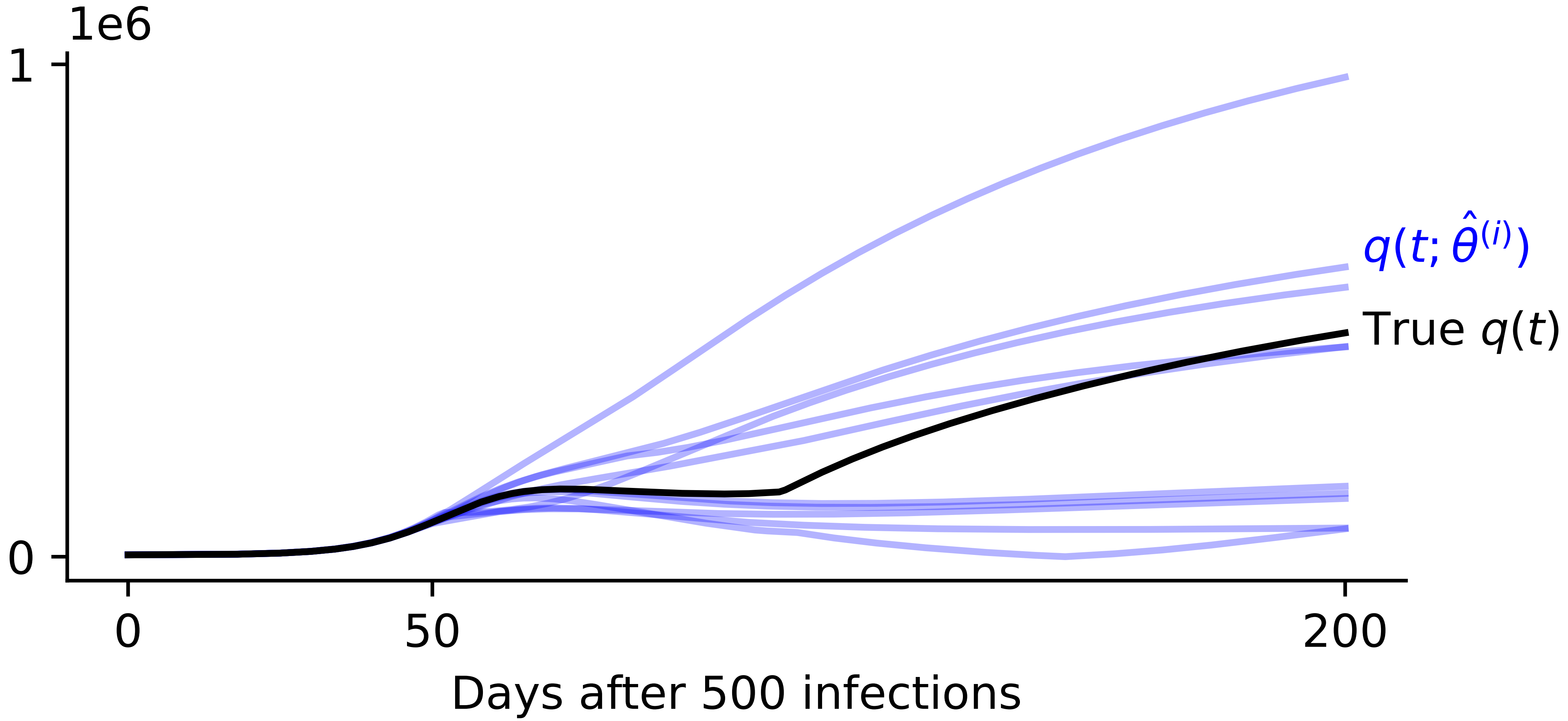
$$d = s + \epsilon, \qquad \epsilon \sim N(0, \sigma^2), \qquad 2\sigma = \pm 0.1s$$

Bayesian NN study

Questions:

- Can we achieve agreement with calibration data using existing Bayesian methods?
- Can we extrapolate beyond calibration data time horizon?

MAP estimates differ by initial guess



Multiple parameter combinations reproduce calibration data.

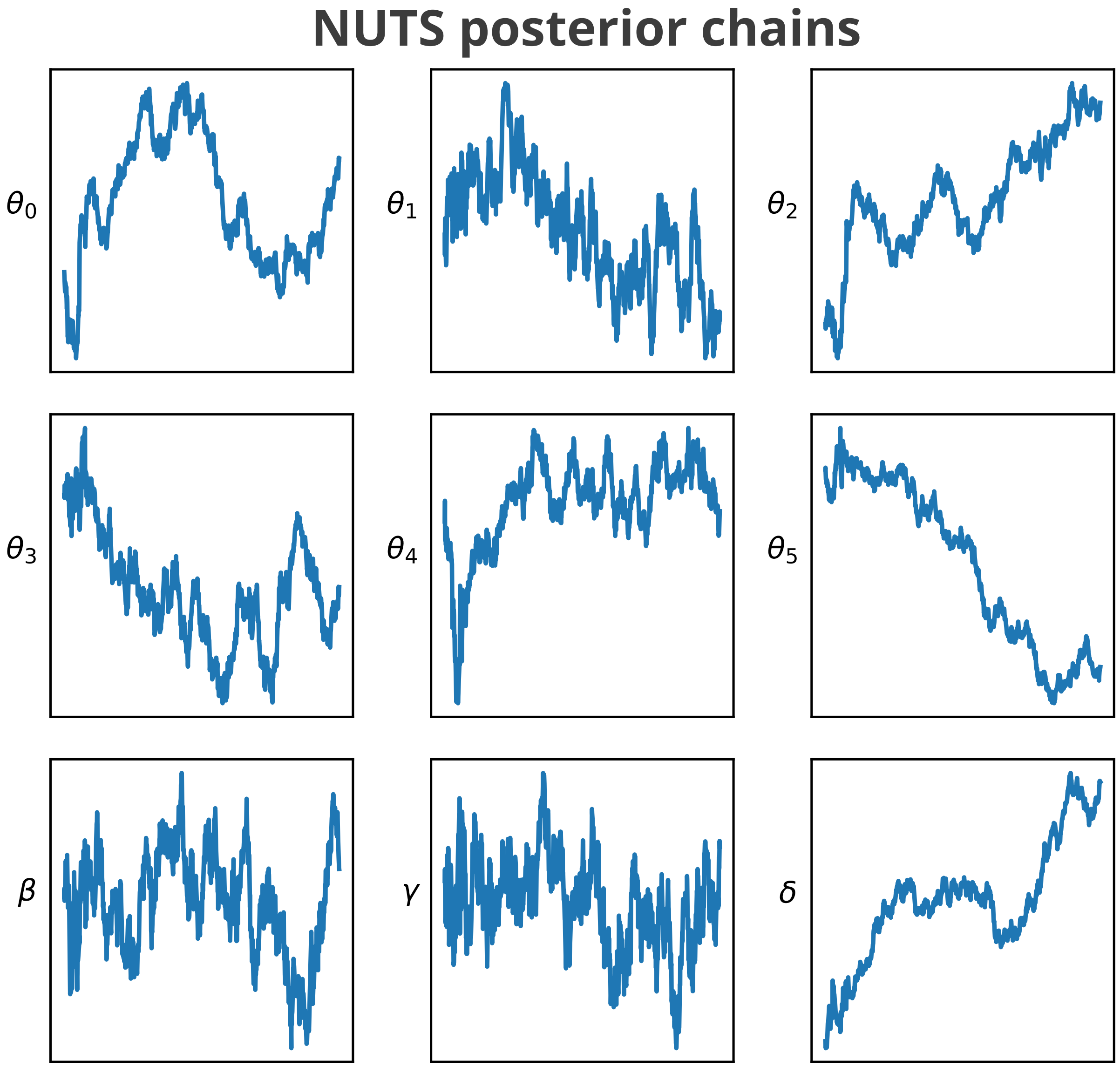
Posterior approximation

Seeded posterior approximations at best MAP point.

Methods:

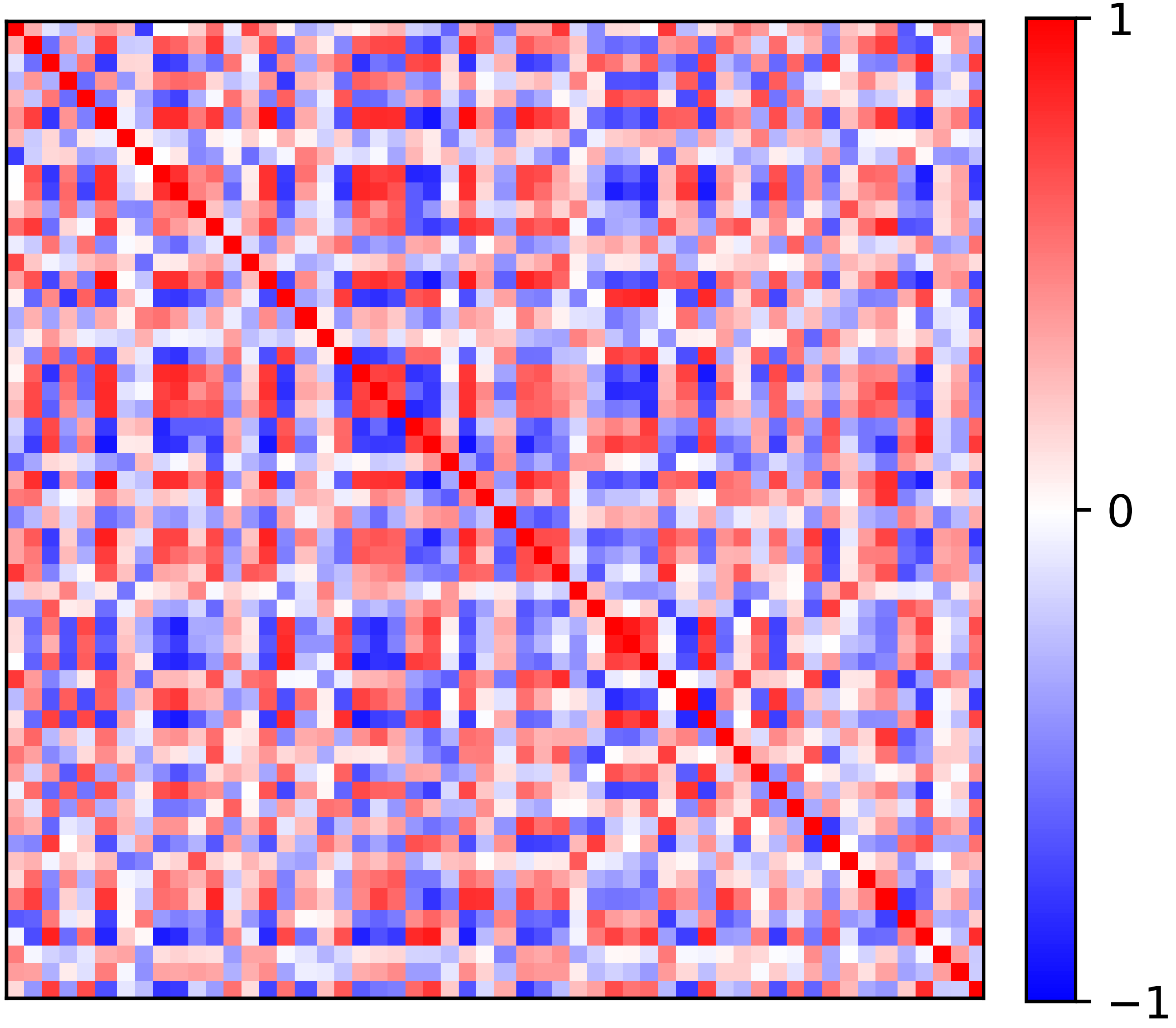
- NUTS
  - HMC variant, derivative based
- ADVI assuming posterior $N\left(\boldsymbol{\mu}, \mathrm{diag}(\boldsymbol{\sigma}^2)\right)$
  - derivative-based, variational inference

Even advanced MCMC algorithms struggle with BNN



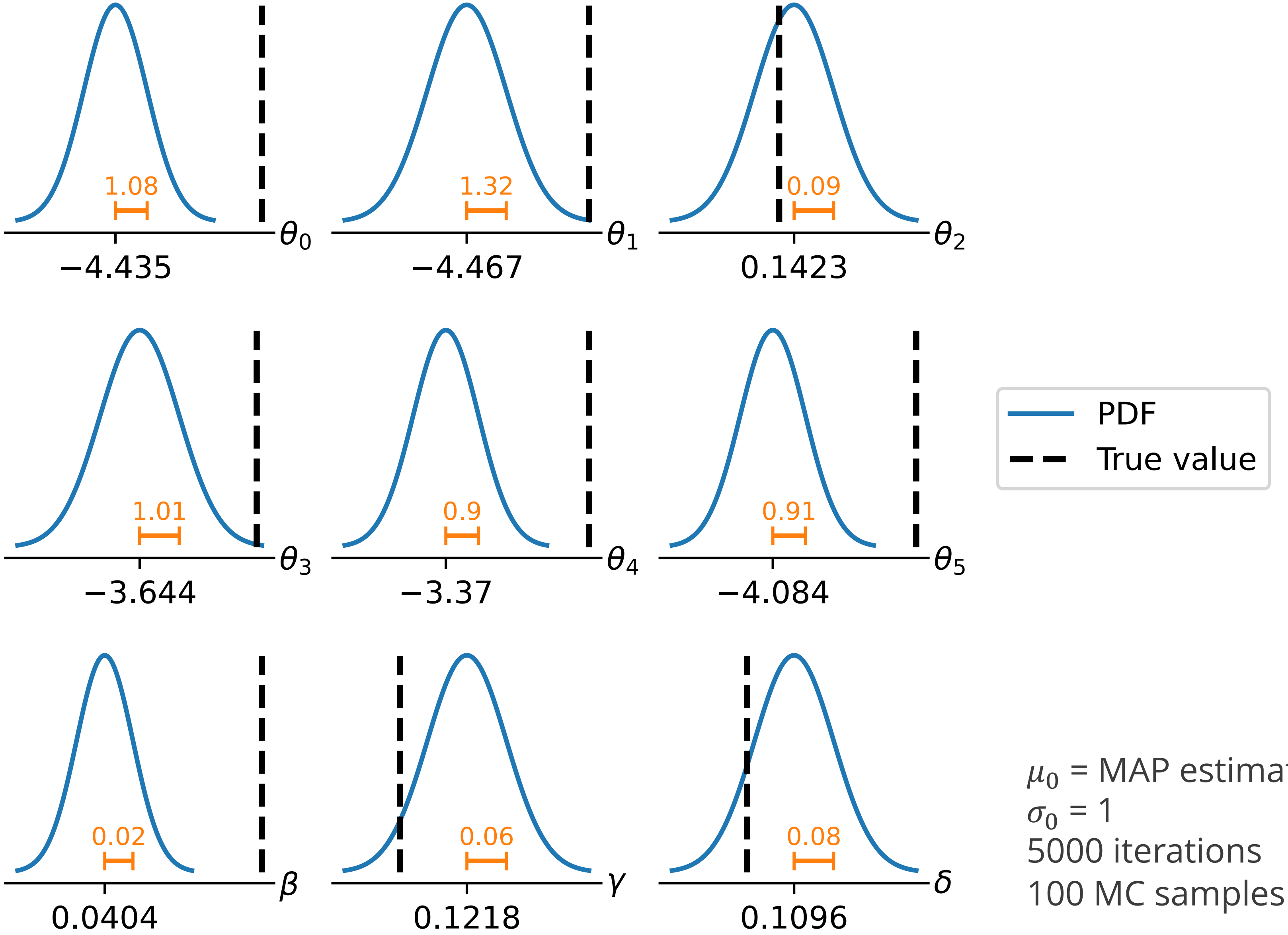
2000 steps
Average acceptance rate: 0.86
Adaptive step size

Correlations indicate complex posterior structure



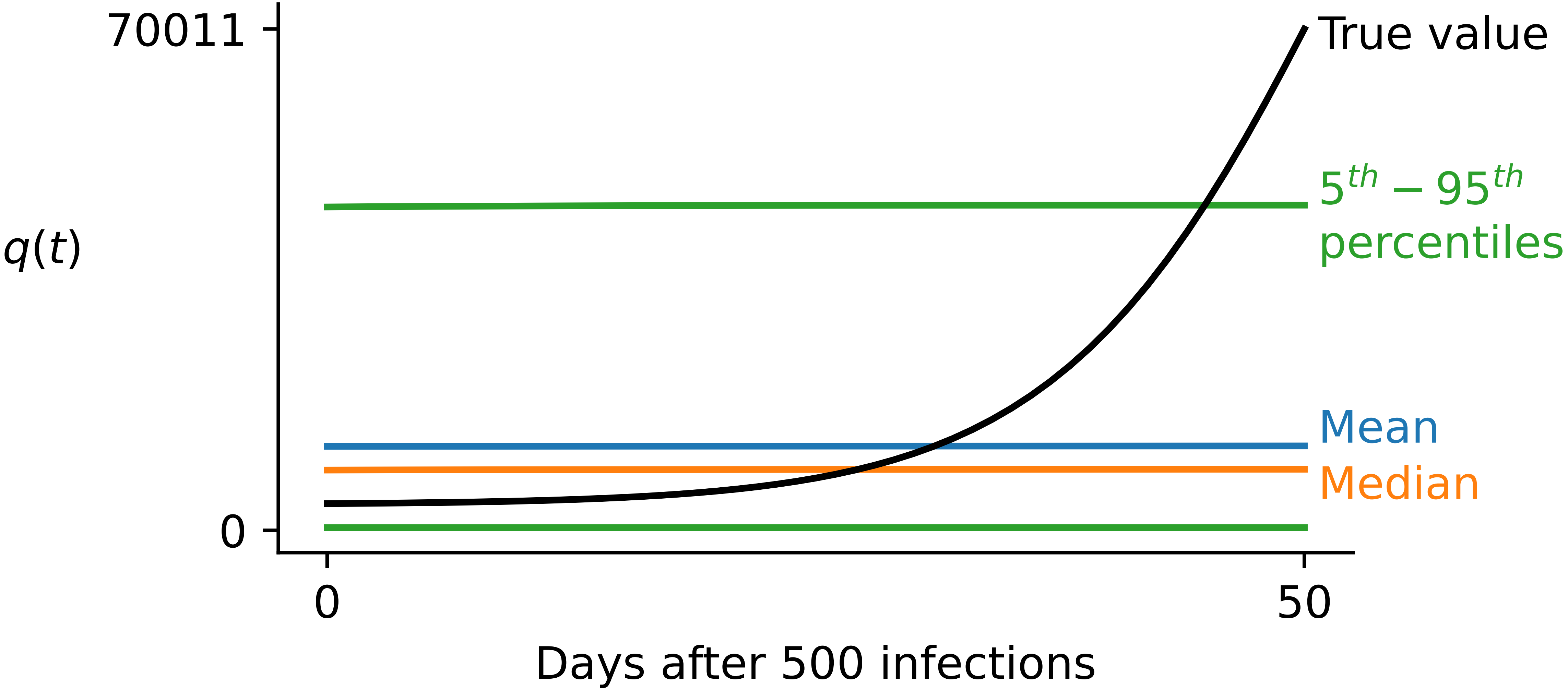
Performance with current ADVI specification even worse

Posterior approximation marginals

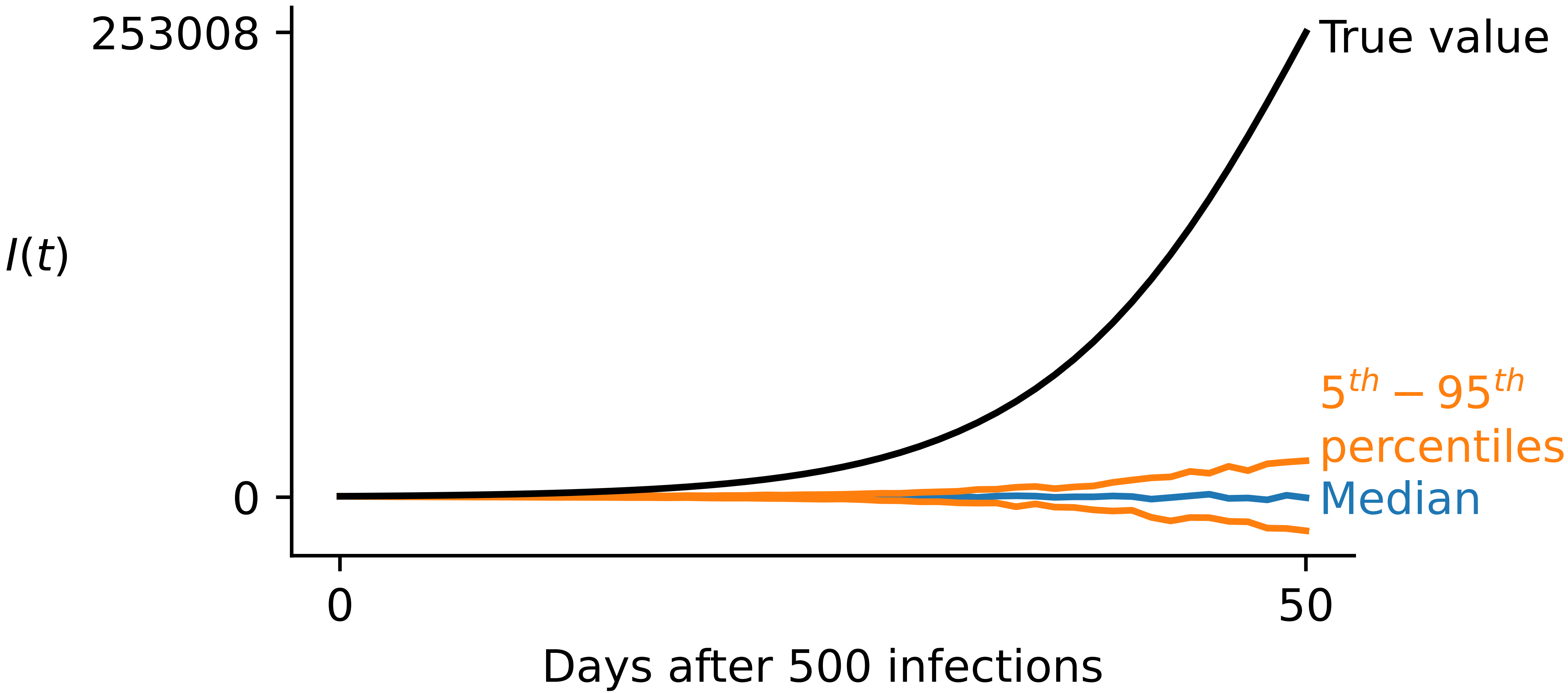


μ_0 = MAP estimate
σ_0 = 1
5000 iterations
100 MC samples for ELBO

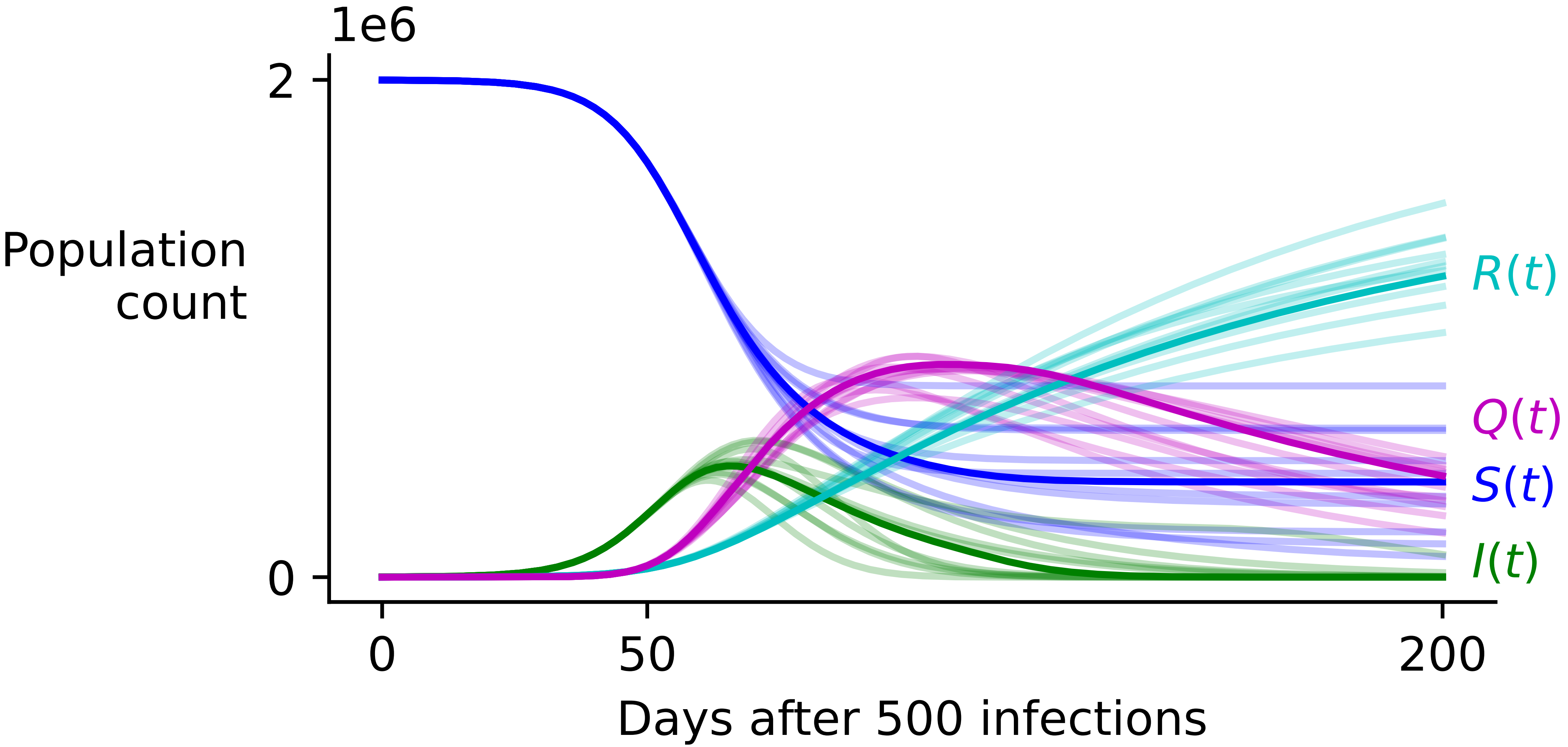
Performance with current ADVI specification even worse



Performance with current ADVI specification even worse



Ensemble of MAP estimates encompass truth



Conclusions and future work

- Despite lower-d NN, Bayesian inference challenging.
- Posterior likely multimodal, non-Gaussian.

- Next steps:
  - Hierarchical model for BNN parameters
  - Estimate posterior with Gaussian mixture model
  - Goal-oriented Bayesian inference

teresaportone.com

SciML.ai

Turing.jl