Comparing Amortized to MCMC-based Bayesian Inference for Cognitive Models of the Stop-Signal Paradigm

Background: Amortized Bayesian Inference (ABI) is an emerging technique that improves on the ideas of approximate Bayesian computation by integrating learning the data-parameter mapping via deep learning, thus only requiring a data-generative model. We present comparisons for and draw conclusions and take-aways for the application of ABI for (ExGaussian) models for stop-signal detection tasks.

Priors & Model

Setup

Software

- ABI as implemented in BayesFlow [1]
- MCMC using *Dynamic Models of Choice* (DMC) [2]

Prior

Empirical parameter fits on stop-signal task data ([3]; DMC)

- (Marg.) Box-Cox-transformed to normal distribution

Distributions (retransformed)

- Marginally (univariate) normal (UVN)
- Multivariate normal (MVN; covariance of parameters)

Sampling schemes

- Random sampling from distributions
- Latin Hypercube Sampling (LHS) using Cholesky decomp.

Model

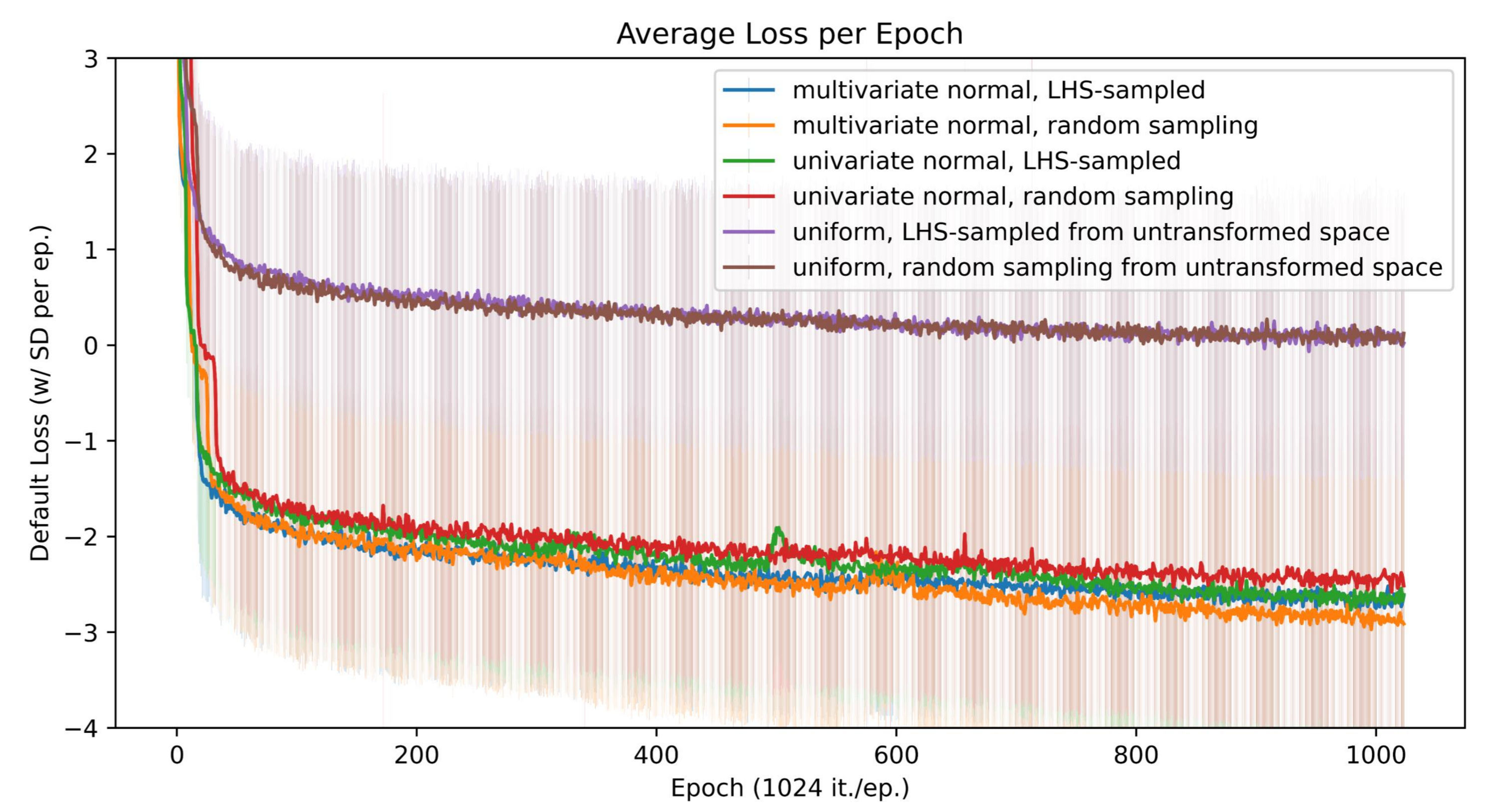
Described in BEESTS [4]

Two Ex-Gaussian processes, each with density

$$f(t \mid \mu, \sigma, \tau) = \frac{1}{\tau} \exp\left(\frac{\mu - t}{\tau} + \frac{\sigma^2}{2\tau^2}\right) \times \Phi\left(\frac{t - \mu}{\sigma} - \frac{\sigma}{\tau}\right); \mu, \sigma, \tau > 0$$

In stop trials (25%), $t_s = t + SSD$ that is determined with a staircase that varies in $50ms$ intervals

BayesFlow Convergence across Schemes

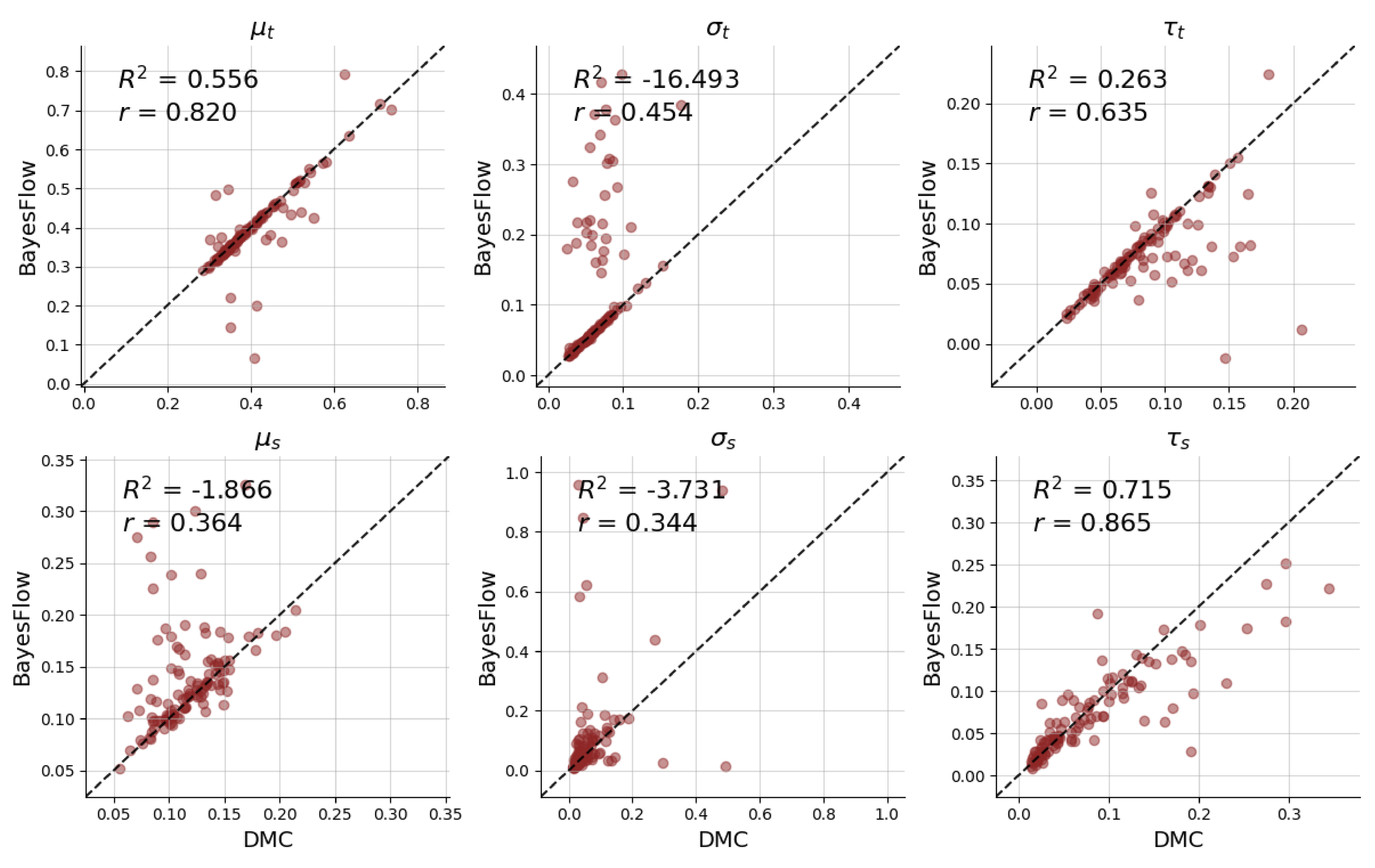


Presents (not converged) models differing in prior sampling scheme

Explorations

Comparison with DMC

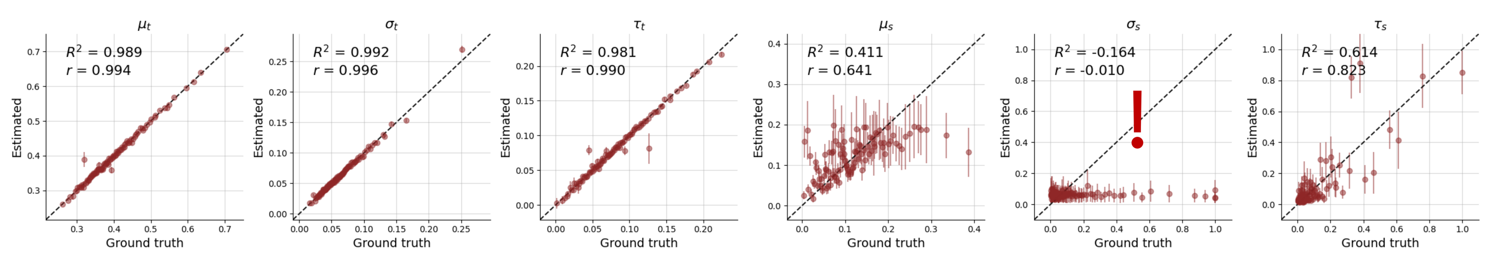
Fitting data with DMC and BayesFlow, using UVN priors



DMC estimates skewed given the sampling distributions

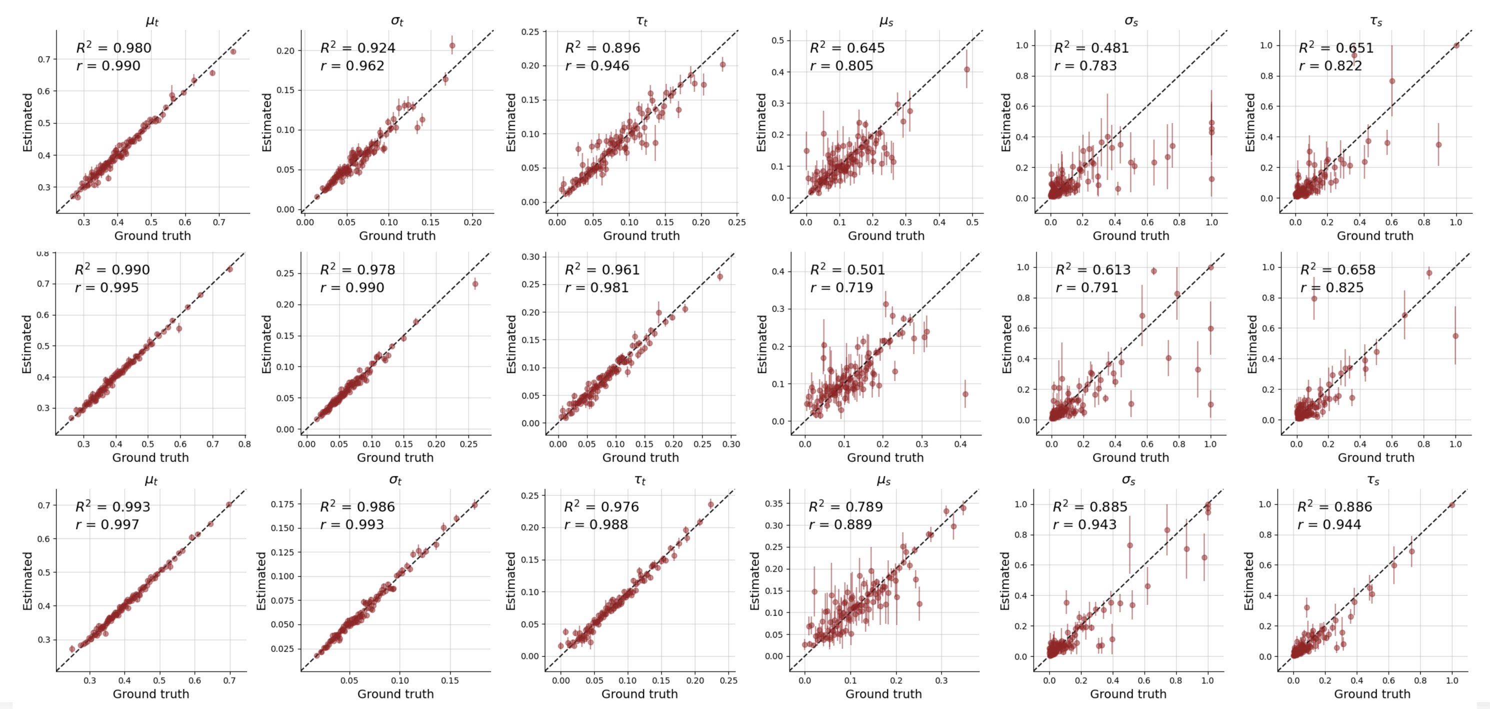
BayesFlow parameter recovery

Fixed sample size (n = 10,000)



See issues with convergence, especially with σ_s -> computing time

Varying sample size (n = 200, 600, 1000)



Takeaways:

- ABI can somewhat reliably recover reaction time parameters in the stop signal paradigm. The unobserved data is much more challenging (as expected), so convergence must be insured when fitting models
- Constructing appropriate prior distributions is a challenge (relevant for training convergence). Informative priors reduce training effort considerably, so we recommend using them whenever empirically / theoretically possible

[1] Radev, S. T., et al. (2020). BayesFlow: Learning complex stochastic models with invertible neural networks. *IEEE transactions on neural networks and learning systems*, *33*(4), 1452-1466.
[2] Heathcote, A., et al. (2019). Dynamic models of choice. *Behavior research methods*, *51*, 961-985.
[3] White, C. N., et al.. (2014). Decomposing decision components in the stop-signal task: a model-based approach to individual differences in inhibitory control. *Journal of cognitive neuroscience*, *26*(8), 1601-1614.
[4] Matzke, D., et al. (2013). Release the BEESTS: B ayesian E stimation of E x-Gaussian ST op-S ignal reaction time distributions. *Frontiers in Psychology*, *4*, 918.

Leonhard Volz — l.volz@uva.nl | @levolz@qoto.org

Dora Matzke, Michael D. Nunez, Andrew Heathcote