Supplementary material for paper: Constraining the Dynamics of Deep Probabilistic Models

Marco Lorenzi [1] **Maurizio Filippone** [2]

1. Description of the ODE systems considered in this work

Lotka-Volterra (Goel et al., 1971). This ODE describes a two-dimensional process with the following dynamics:

$$\frac{df_1}{dt} = \alpha f_1 - \beta f_1 f_2; \qquad \frac{df_2}{dt} = -\gamma f_2 + \delta f_1 f_2,$$

and is identified by the parameters $\boldsymbol{\theta} = \{\alpha, \beta, \gamma, \delta\}$. Following (Niu et al., 2016) we generated a ground truth from numerical integration of the system with parameters $\boldsymbol{\theta} = \{0.2, 0.35, 0.7, 0.4\}$ over the interval $[0, 30]$ and with initial condition $[1, 2]$. We generated two different configurations, composed by respectively 34 and 51 observations sampled at uniformly spaced points, and corrupted by zero mean Gaussian noise with standard deviation $\sigma = 0.25$ and $\sigma = 0.4$ respectively.

FitzHugh-Nagumo (FitzHugh, 1955). This system describes a two-dimensional process governed by 3 parameters, $\boldsymbol{\theta} = \{a, b, c\}$:

$$\frac{df_1}{dt} = c(f_1 - b\frac{(f_1)^3}{3} + f_2); \qquad \frac{df_2}{dt} = -\frac{1}{c}(f_1 - a + b * f_2).$$

We reproduced the experimental setting proposed in (Macdonald & Husmeier, 2015), by generating a ground truth with $\boldsymbol{\theta} = \{3, 0.2, 0.2\}$, and by integrating the system numerically with initial condition $[-1, 1]$. We created two scenarios; in the first one, we sampled 401 observations at equally spaced points within the interval $[0, 20]$, while in the second one we sampled only 20 points. In both cases we corrupted the observations with zero-mean Gaussian noise with $\sigma = 0.5$.

Biopathways (Vyshemirsky & Girolami, 2007). These equations describe a five-dimensional process associated with 6 parameters $\boldsymbol{\theta} = \{k_1, k_2, k_3, k_4, V, K_m\}$ as follows:

$$\begin{aligned}
\frac{df_1}{dt} &= -k_1 f_1 - k_2 f_1 f_3 + k_3 f_4; \\
\frac{df_2}{dt} &= k_1 f_1; \\
\frac{df_3}{dt} &= -k_2 f_1 f_3 + k_3 f_4 + \frac{V f_5}{K_m + f_5}; \\
\frac{df_4}{dt} &= k_2 f_1 f_3 - k_3 f_4 - k_4 f_4; \\
\frac{df_5}{dt} &= k_4 f_4 - \frac{V f_5}{K_m + f_5}.
\end{aligned}$$

We generated data by sampling 15 observations at times $\mathbf{t} = \{0, 1, 2, 4, 5, 7, 10, 15, 20, 30, 40, 50, 60, 80, 100\}$ (Macdonald & Husmeier, 2015). The ODE parameters were set to $\boldsymbol{\theta} = \{k_1 = 0.07, k_2 = 0.6, k_3 = 0.05, k_4 = 0.3, V = 0.017, K_m = 0.3\}$, and the initial values were $[1, 0, 1, 0, 0]$. We generated two different scenarios, by adding Gaussian noise with $\sigma^2 = 0.1$ and $\sigma^2 = 0.05$, respectively.

[*]Equal contribution [1]University of Cote d'Azur, INRIA Sophia Antipolis, EPIONE research group, France [2]EURECOM, Department of Data Science, Sophia Antipolis, France. Correspondence to: Marco Lorenzi <marco.lorenzi@inria.fr>, Maurizio Filippone <maurizio.filippone@eurecom.fr>.

Proceedings of the 35th International Conference on Machine Learning, Stockholm, Sweden, PMLR 80, 2018.

2. Detailed results of the benchmark on ODE parameter inference

In figures 1 and 2, we report the detailed estimate/posterior distribution obtained by the competing methods on the three ODE systems considered in this study.

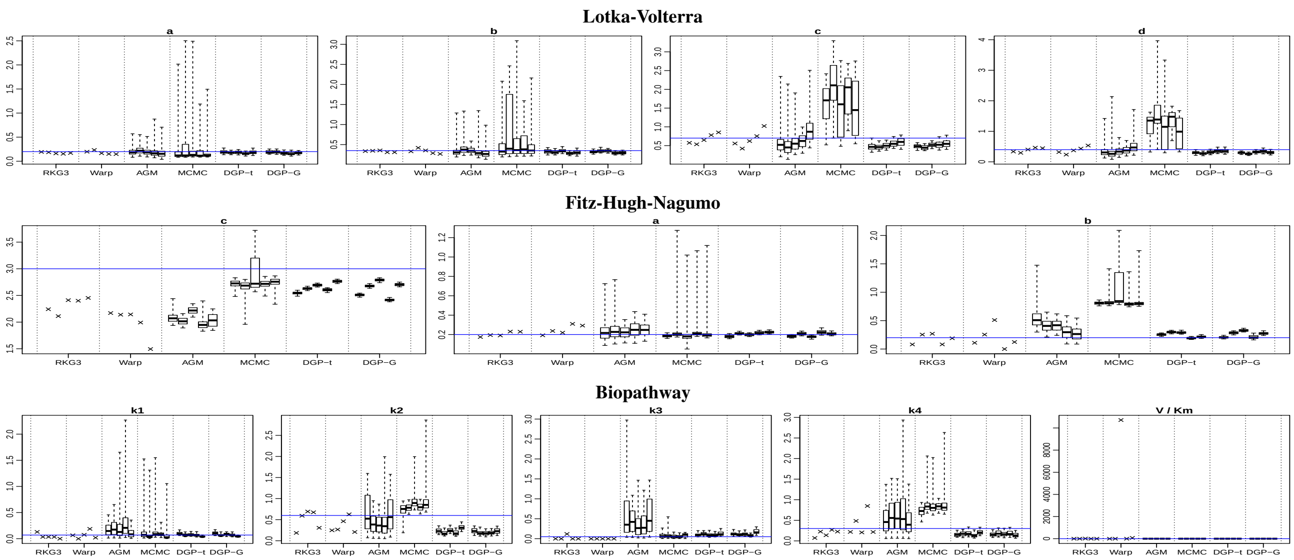

Figure 1. Box-plot of posteriors over model parameters. The five box-plots for each method indicate five different repetitions of the instantiation of the noise.

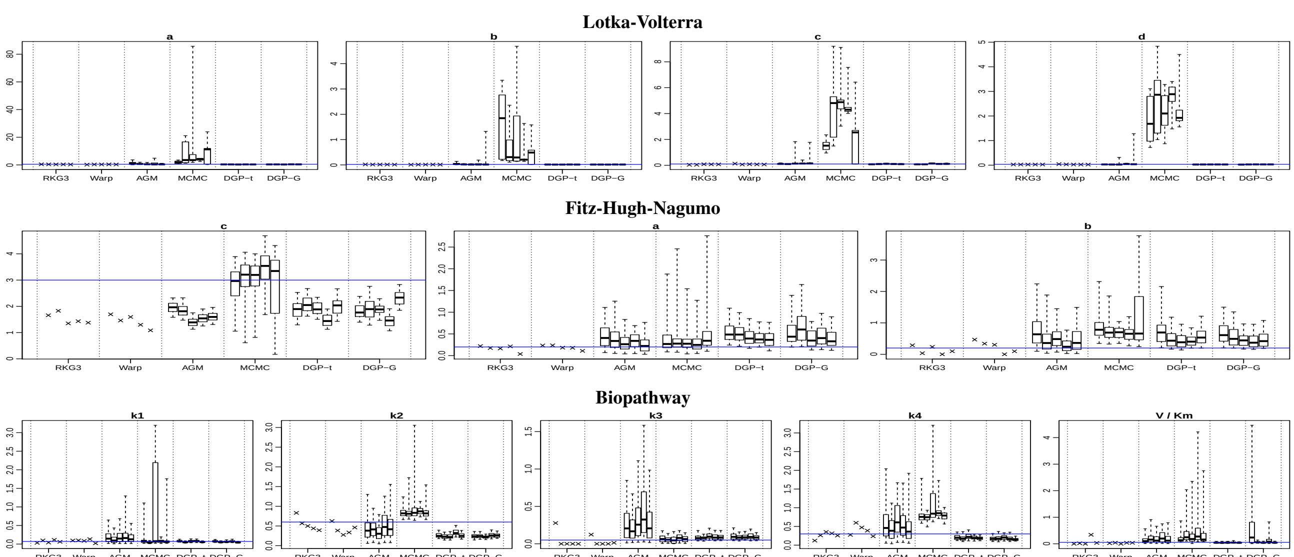

Figure 2. Box-plot of posteriors over model parameters. The five box-plots for each method indicate five different repetitions of the instantiation of the noise.

3. Interpolation results using Matérn covariance in shallow GPs

We report here the result of interpolating FitzHugh-Nagumo ODE with GPs with Matérn covariance. Note that the process is still stationary, but it allows for the modeling of non-smooth functions when ν is small.

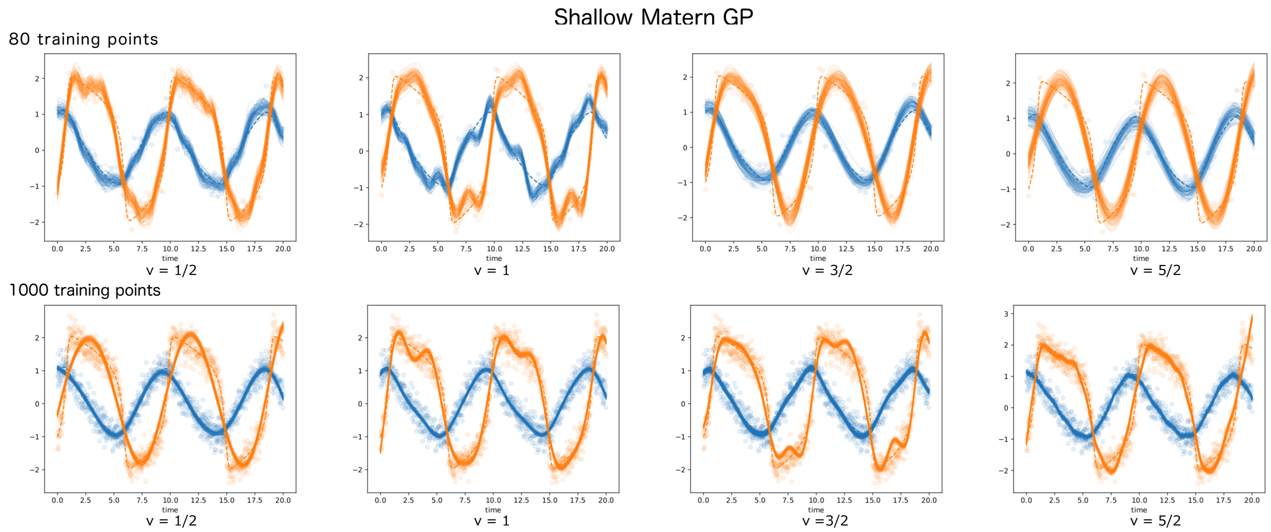

Figure 3. Modeling sampling points from the FitzHugh-Nagumo ODE with GPs with Matérn covariance.

References

FitzHugh, R. Mathematical models of threshold phenomena in the nerve membrane. *The bulletin of mathematical biophysics*, 17(4):257–278, 1955.

Goel, N. S., Maitra, S. C., and Montroll, E. W. On the volterra and other nonlinear models of interacting populations. *Reviews of modern physics*, 43(2):231, 1971.

Macdonald, B. and Husmeier, D. Gradient Matching Methods for Computational Inference in Mechanistic Models for Systems Biology: A Review and Comparative Analysis. *Frontiers in Bioengineering and Biotechnology*, 3:180, 2015. doi: 10.3389/fbioe.2015.00180. URL https://www.frontiersin.org/article/10.3389/fbioe.2015.00180.

Niu, M., Rogers, S., Filippone, M., and Husmeier, D. Fast Parameter Inference in Nonlinear Dynamical Systems using Iterative Gradient Matching. In Balcan, M. F. and Weinberger, K. Q. (eds.), *Proceedings of The 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*, pp. 1699–1707, New York, New York, USA, June 2016. PMLR. URL http://proceedings.mlr.press/v48/niu16.html.

Vyshemirsky, V. and Girolami, M. A. Bayesian ranking of biochemical system models. *Bioinformatics*, 24(6):833–839, 2007.