Paper

An amortized approach to non-linear mixed-effects modeling based on neural posterior estimation

Jonas Arruda[1], Yannik Schälte[1,2], Clemens Peiter[1], Olga Teplytska[3], Ulrich Jaehde[3] and Jan Hasenauer[1,2]

[1]University of Bonn, Life and Medical Sciences Institute, 53115 Bonn, Germany, [2]Helmholtz Zentrum München, Computational Health Center, 85764 Neuherberg, Germany, [3]University of Bonn, Pharmaceutical Institute, 53121 Bonn, Germany

UNIVERSITÄT BONN

Overview

- Individuals and cells are **heterogeneous**
- **Complexity** is an intrinsic property of many diseases
- Dataset **size** is growing and current methods are not scalable

→ need for **scalable** estimation of non-linear mixed effect models

$$p(\mathscr{D} \mid \theta) = \prod_{i=1}^{N} \int p\left(\tilde{\mathbf{y}}^{(i)} \mid \phi\right) p_{\text{pop}}(\phi \mid \theta)\, \mathrm{d}\phi$$

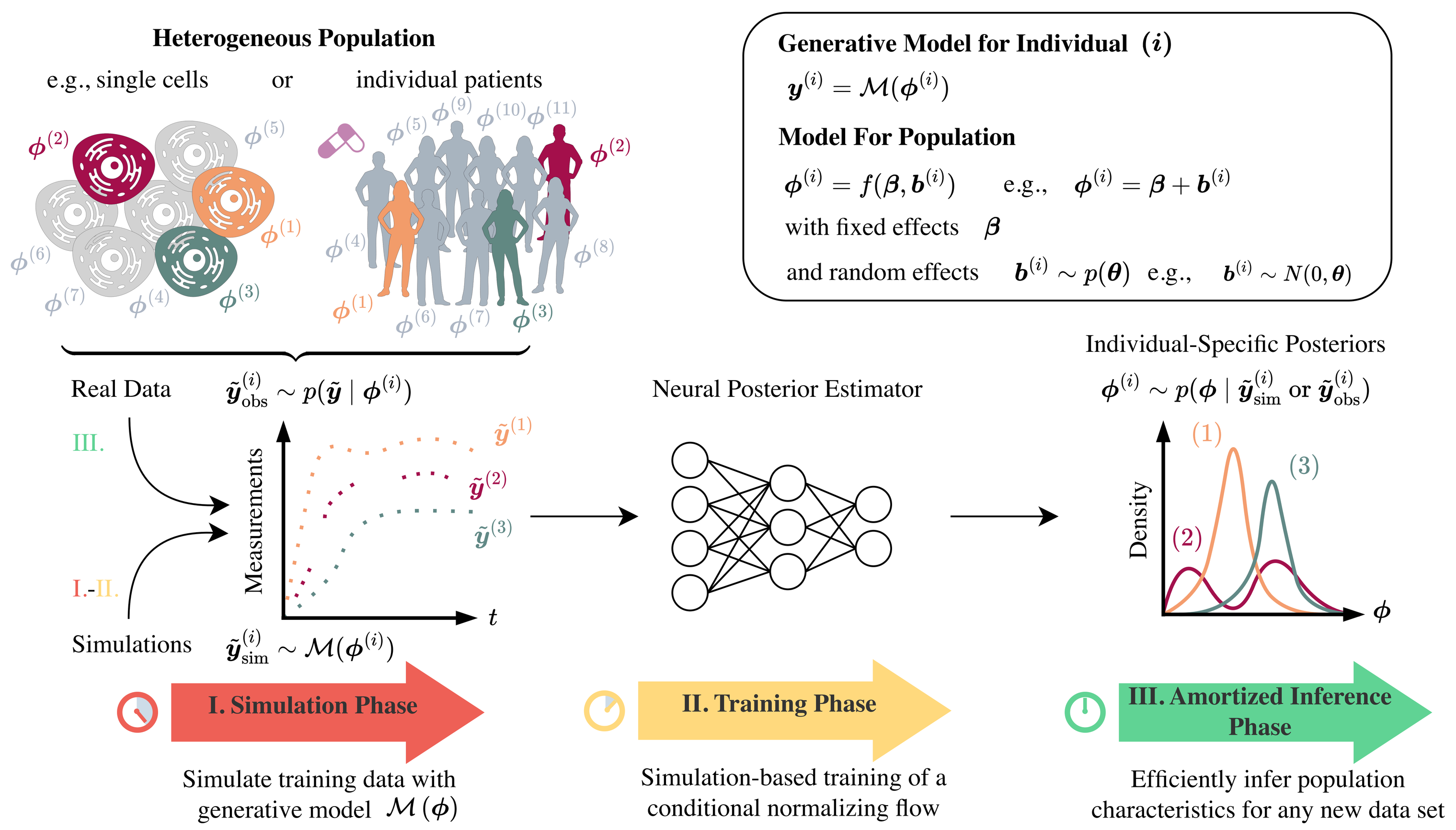
Abstract

- Non-linear mixed-effects models are a powerful tool for studying heterogeneous populations in various fields, including biology, medicine, economics, and engineering
- The aim is to find a distribution over the parameters that describe the whole population using a model that can generate simulations for an individual of that population.
- Fitting these distributions to data is computationally challenging if the description of individuals is complex and the population is large.
- We propose a novel machine learning-based approach: We exploit neural density estimation based on conditional normalizing flows to approximate individual-specific posterior distributions in an amortized fashion, thereby allowing for efficient inference of population parameters.
- Applying this approach to problems from cell biology and pharmacology, we demonstrate its unseen flexibility and scalability to large data sets compared to established methods.

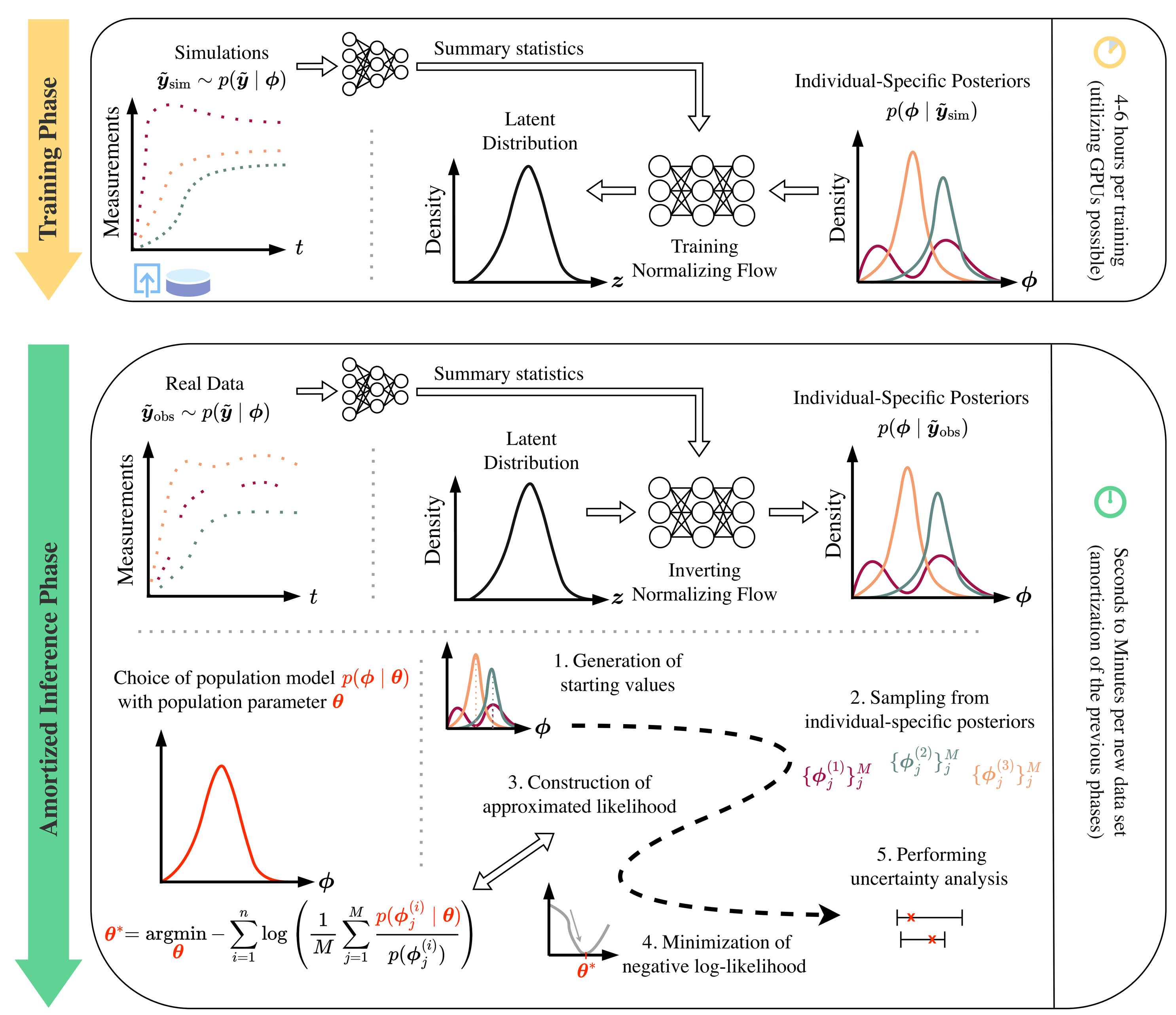
Basic Definitions

- The **generative model** M can generate simulations for a given set of parameters $\phi^{(i)} \in \mathbb{R}^k$ and time points. As a generative model, we understand any parametric model, such as linear models, the solution of (stochastic) differential equations, or Markov jump processes, which can produce simulations for an individual given some parameters $\phi^{(i)}$.
- A **non-linear mixed-effects (NLME)** model describes observations of the entire population using the generative model M and individual-specific parameters $\phi^{(i)} \in \mathbb{R}^k$. Individual parameters, which need to be marginalized out, come from a **population model** parameterized by population parameters θ, i.e., $\phi^{(i)} \sim p(\theta)$.
- **Conditional normalizing flows** can transform a complicated conditional density, such as a posterior probability, into a simpler density from which we know how to sample.

Concept



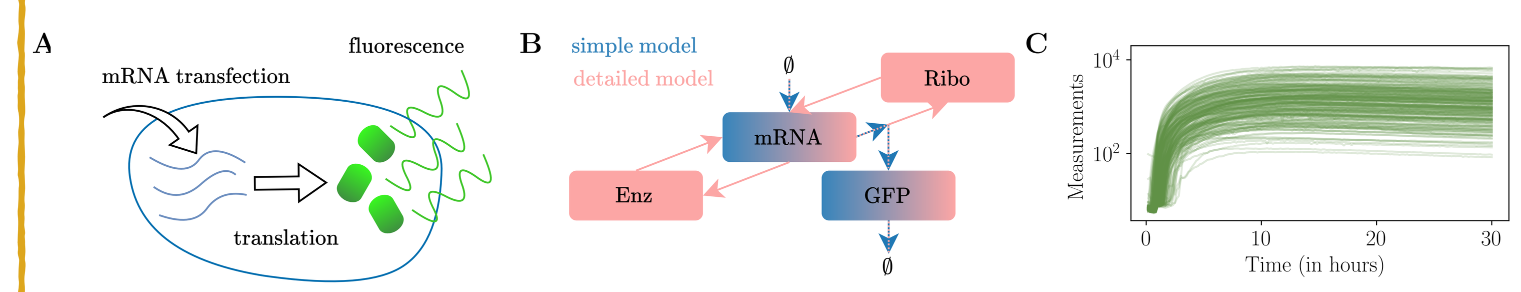
Our approach employs amortizing inference, based on **invertible neural networks**, to sample from the **individual-specific posterior distribution**, which is then used to infer the **population-level** parameters. After training, we **amortize** the cost of training by repeatedly applying the trained neural networks (potentially from different data sets) for inference.



Results

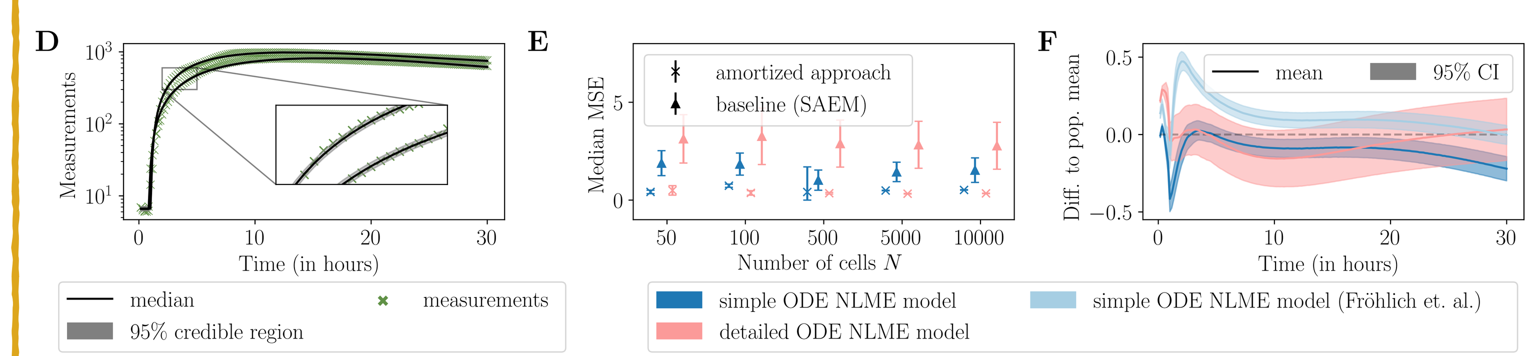
Example problems

Single-cell models describing translation kinetics after transfection with mRNA (model from Fröhlich et. al. (2018)).



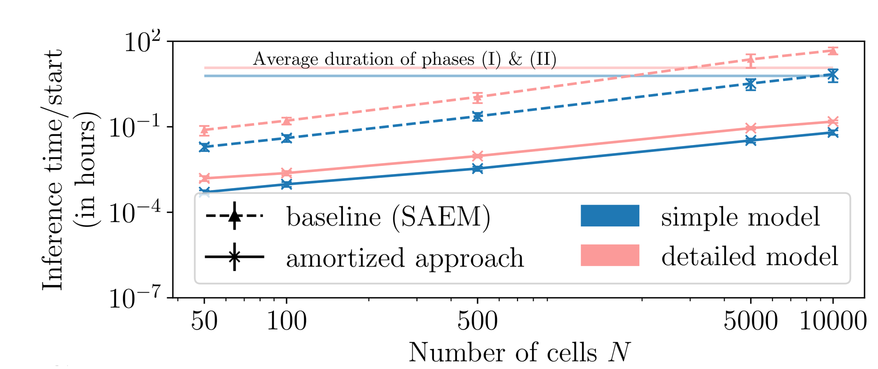
A Schematic representation of GFP translation after mRNA transfection in a single cell.
B Visualization of the simple and detailed single-cell ODE models (color indicates the states included in model).
C Fluorescence intensity time courses of 200 single cells out of 5488.

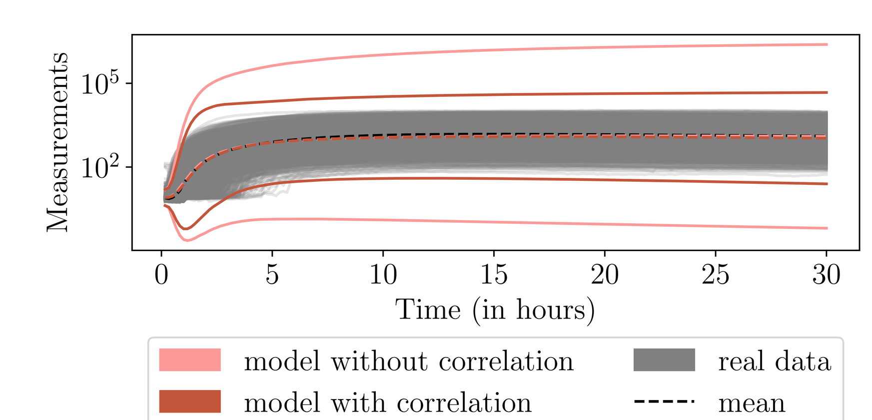
We show that we get fits on single cell level and population level at least as good as the state of the art method (SAEM - Kuhn, E. and Lavielle, M. (2005)).



D Credible regions of trajectories estimated by the neural posterior estimator for two real cells.
E Median of the mean squared error (MSE) of the estimated compared to the true parameters of the synthetic data for both single-cell NLME models is shown for different numbers of cells.
F The difference in the population mean estimated from real trajectories and simulations generated with the estimated population parameters is shown.

Our approach is more **scalable** and flexible in terms of choosing the population model (e.g. including correlations) and can work with **arbitrary generative models** (e.g. ODEs or SDEs, only simulation based). Furthermore, **no retraining** is required to include changes or an entire new population model.





Conclusion

The amortized approach is **more scalable** than state of the art methods, it is **flexible** w.r.t. the population and the individual model, it **facilitates** Bayesian inference, uncertainty analysis and the use of more complex models.

Cons

- Training the neural posterior estimator is simulation hungry. Hence, method more relevant for large data sets or when the likelihood for the generative model is not available.

Pros

- You can use **any generative model** for the individual (as long we can perform simulations).
- You can include missing data, censoring directly in the generative model.
- You can estimate point estimates of the population parameters or perform a **full Bayesian analysis.**
- No assumptions of the population model are needed to train the neural posterior estimator. Hence, the **population model can be changed at any time** without requiring retraining.
- Sampling from the neural posterior estimator is fast, therefore, inference time of the population parameters **scales almost constant** with respect to the number of individuals in a population.

Main References

- **Arruda, J., Schälte, Y., Peiter, C., Teplytska, O., Jaehde, U., and Hasenauer, J. An amortized approach to non-linear mixed-effects modeling based on neural posterior estimation. *Proceedings of the 41st International Conference on Machine Learning,* 2024.**
- Cranmer, K., Brehmer, J., and Louppe, G. The frontier of simulation-based inference. *Proceedings of the National Academy of Sciences*, 117(48):30055–30062, 2020. doi: 10.1073/pnas.1912789117. *(overview simulation based inference)*
- Fröhlich, F., Reiser, A., Fink, L., Woschée, D., Ligon, T., Theis, F.J., Rädler, J.O., and Hasenauer, J. Multi- experiment nonlinear mixed effect modeling of single-cell translation kinetics after transfection. *npj Systems Biology and Applications*, 5(1):1, 2018. doi: 10.1038/s41540-018-0079-7. *(data & model)*
- Kuhn, E. and Lavielle, M. Maximum likelihood estimation in nonlinear mixed effects models. *Comput. Stat. Data. Anal.*, 49(4):1020 – 1038, 2005. doi: http://dx.doi.org/10. 1016/j.csda.2004.07.002. *(SAEM, baseline method)*
- Radev, S. T., Mertens, U. K., Voss, A., Ardizzone, L., and Köthe, U. Bayesflow: Learning complex stochastic models with invertible neural networks. IEEE transactions on neural networks and learning systems, 2020. *(Bayesflow)*

arrjon/Amortized-NLME-Models

jonas.arruda@uni-bonn.de

BIGS Bonn International Graduate School of Mathematics
Federal Ministry of Education and Research
Emune
IML IRU Mathematics & Life Sciences