

Statistical methods for mechanistic models in neuroscience

Likelihood-Free Inference Workshop, Flatiron Institute

Jan-Matthis Lueckmann

Computational Neuroengineering (Jakob Macke)

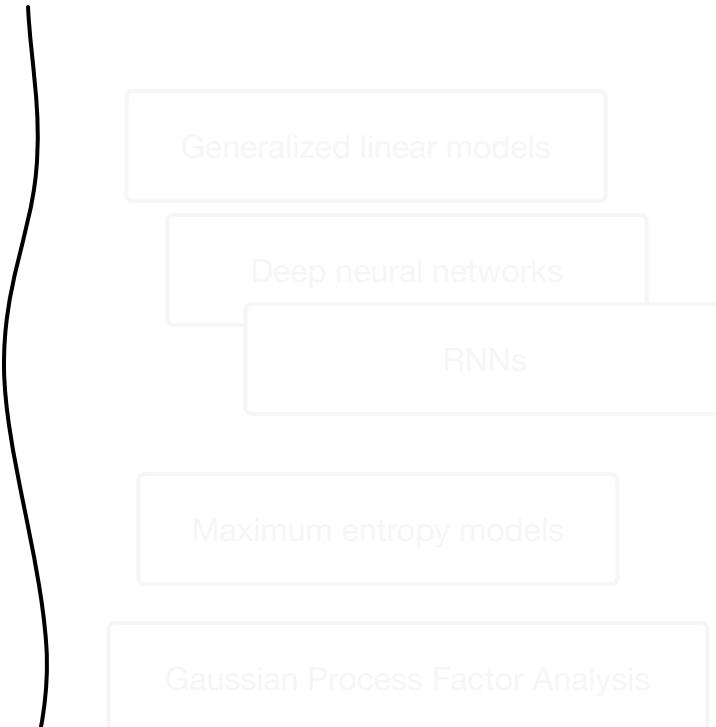
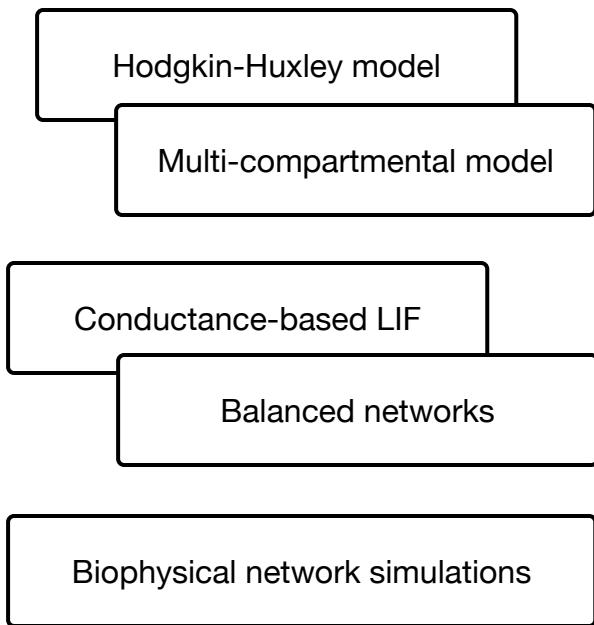
March 18th, 2019



Mechanistic models

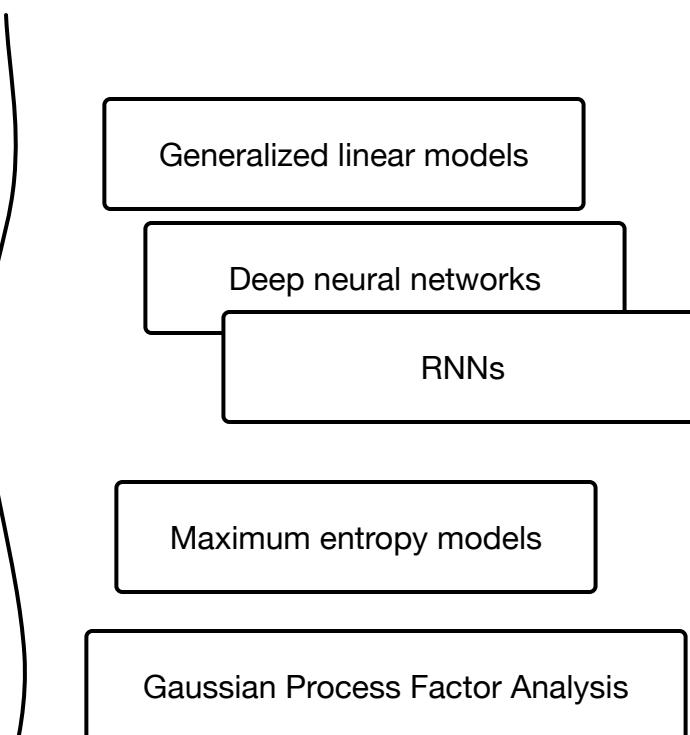
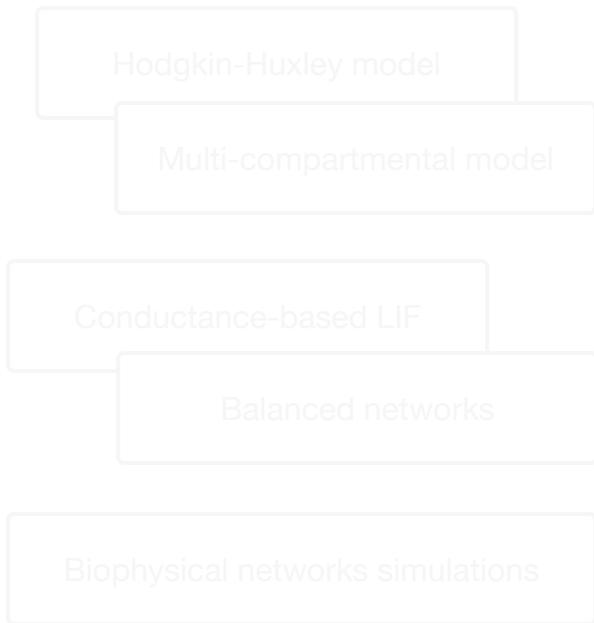
vs.

statistical/ML models



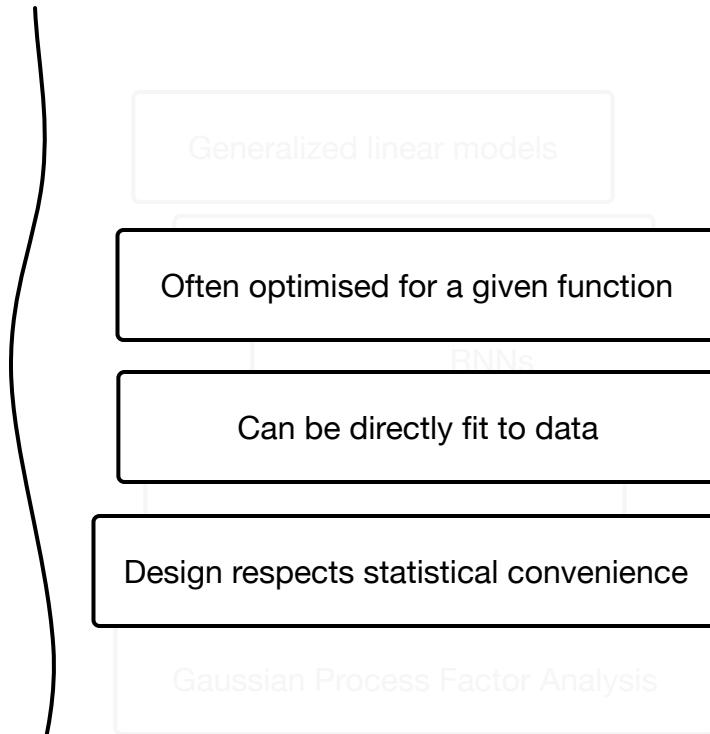
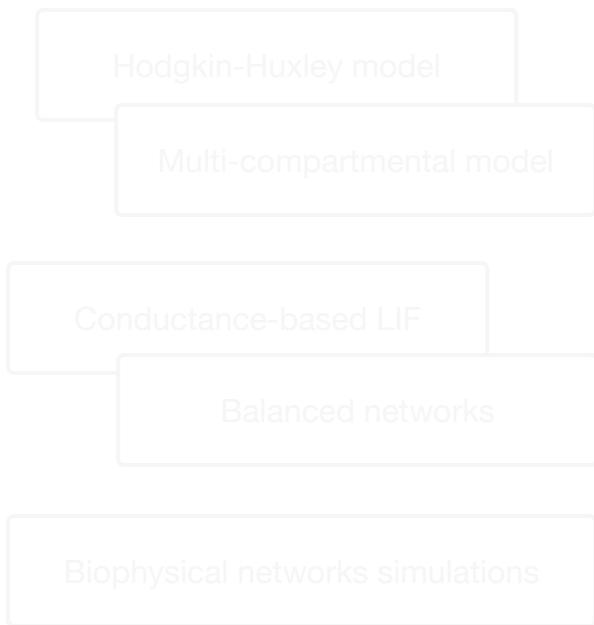
Mechanistic models

vs. statistical/ML models

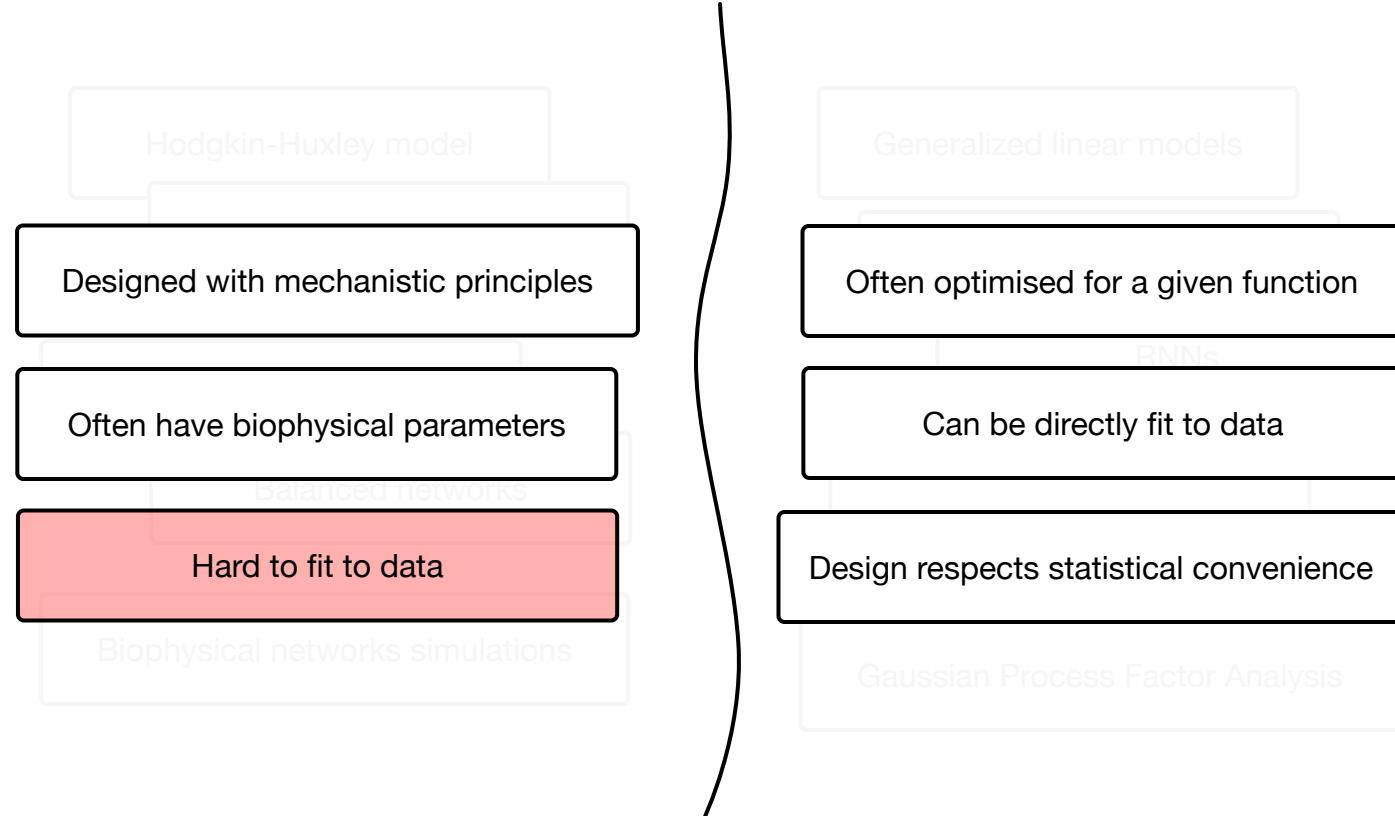


Mechanistic models

vs. statistical/ML models



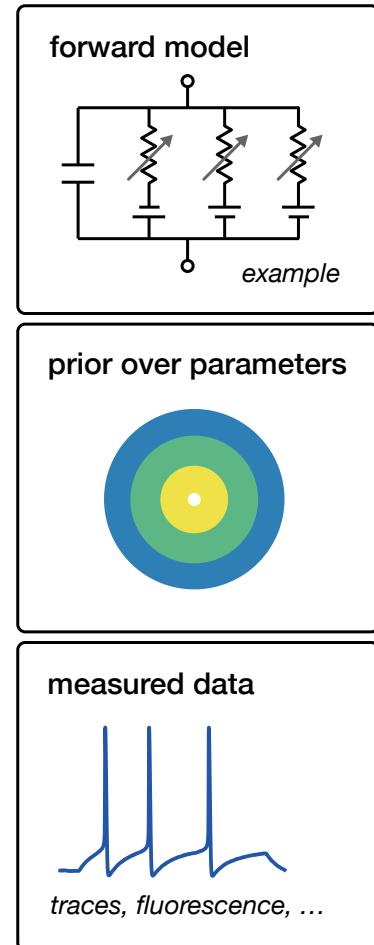
Mechanistic models vs. statistical/ML models



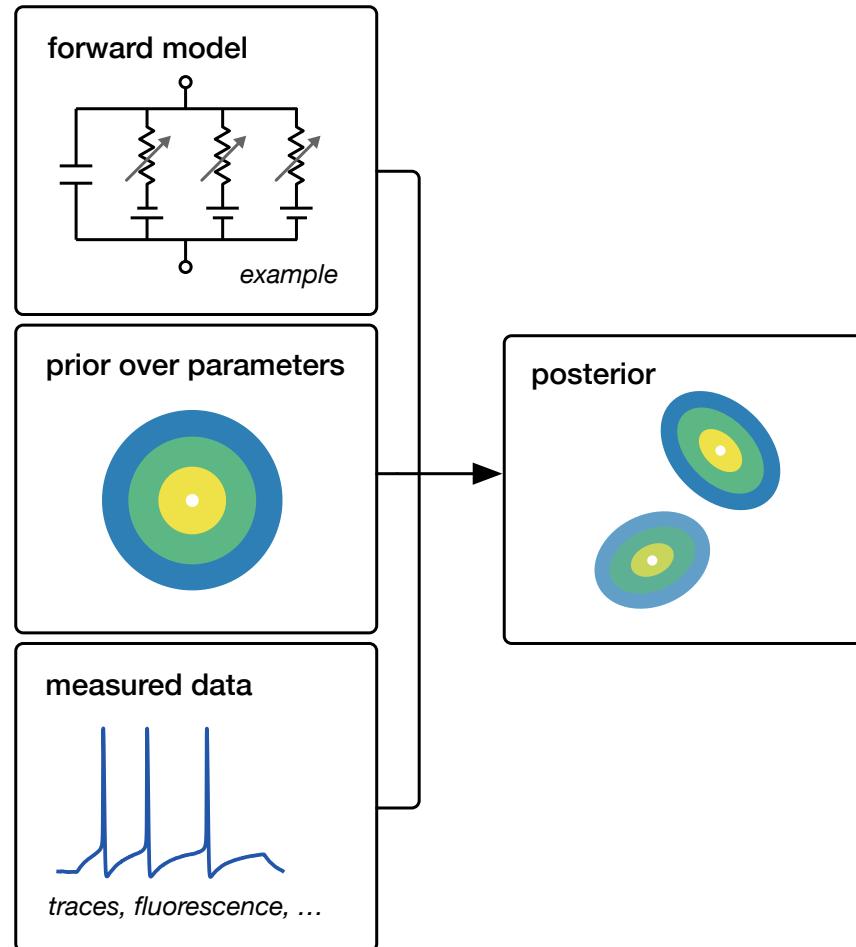
**How can we make
statistical inference tractable
for mechanistic models?**

**How can we find parameters
that are consistent with
data and prior knowledge?**

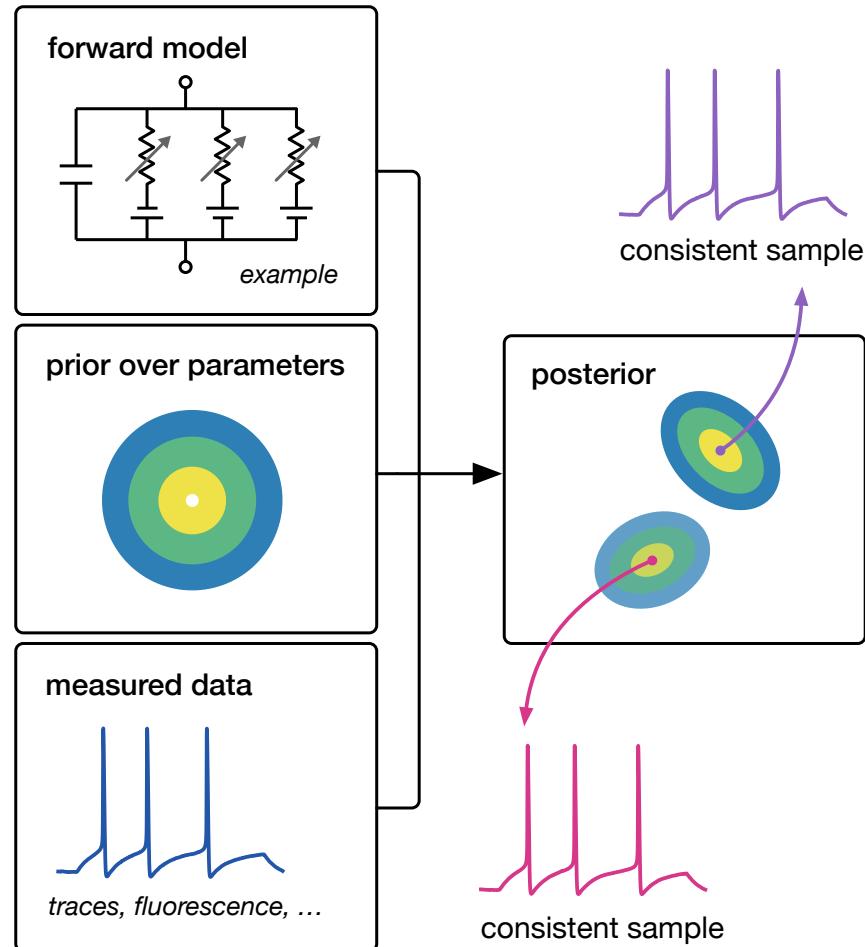
Goal: Bayesian inference for mechanistic models



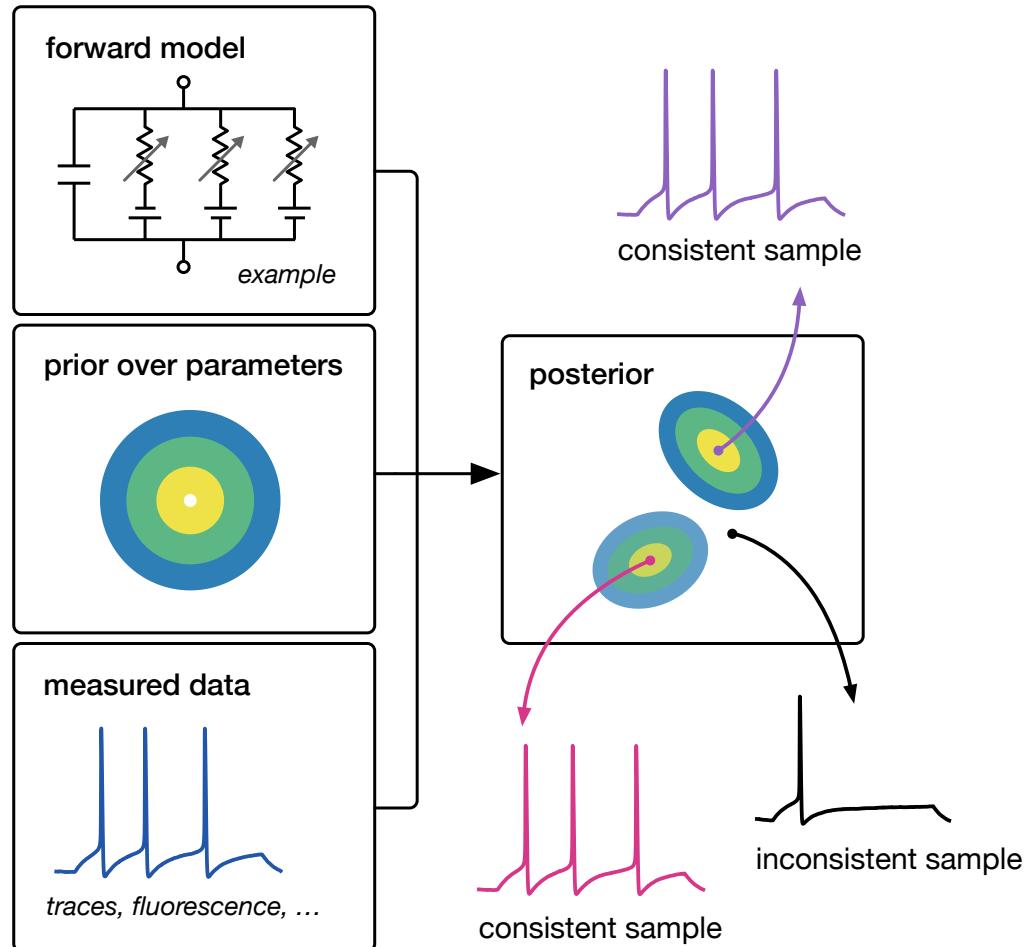
Goal: Bayesian inference for mechanistic models



Goal: Bayesian inference for mechanistic models



Goal: Bayesian inference for mechanistic models



Nobel Prize in Physiology or Medicine 1963

"for their discoveries concerning the ionic mechanisms involved in excitation and inhibition in the peripheral and central portions of the nerve cell membrane."

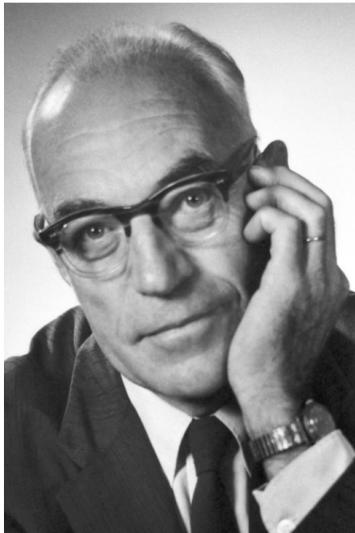


Photo from the Nobel Foundation archive.

Sir John Carew Eccles
Prize share: 1/3

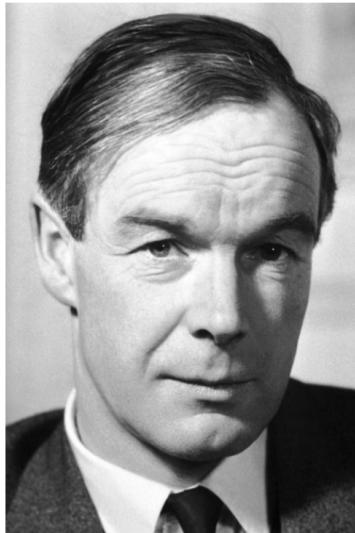


Photo from the Nobel Foundation archive.

Alan Lloyd Hodgkin
Prize share: 1/3

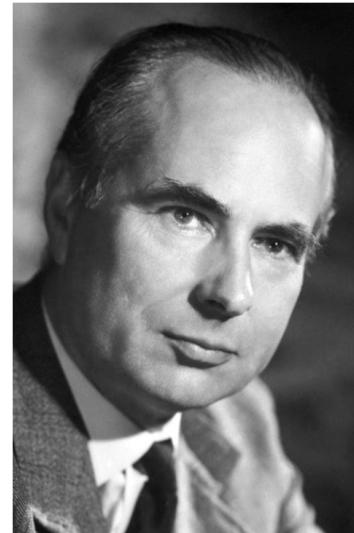
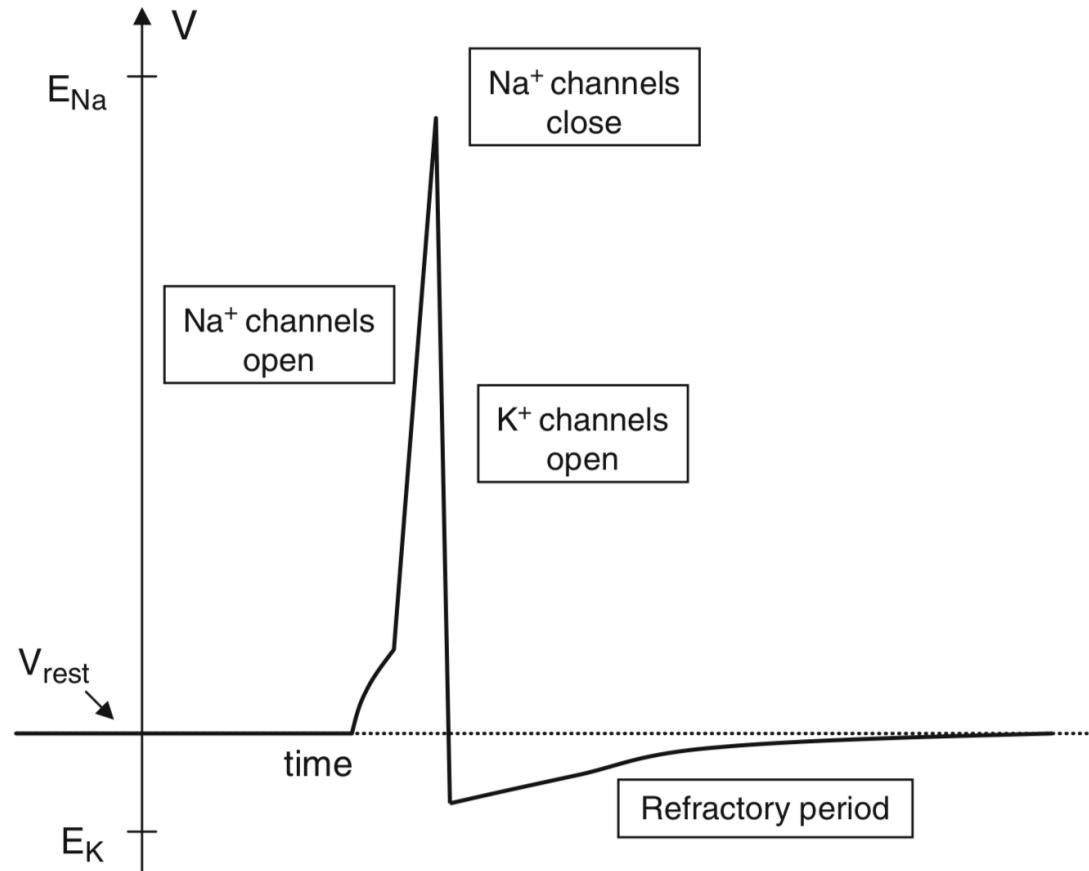


Photo from the Nobel Foundation archive.

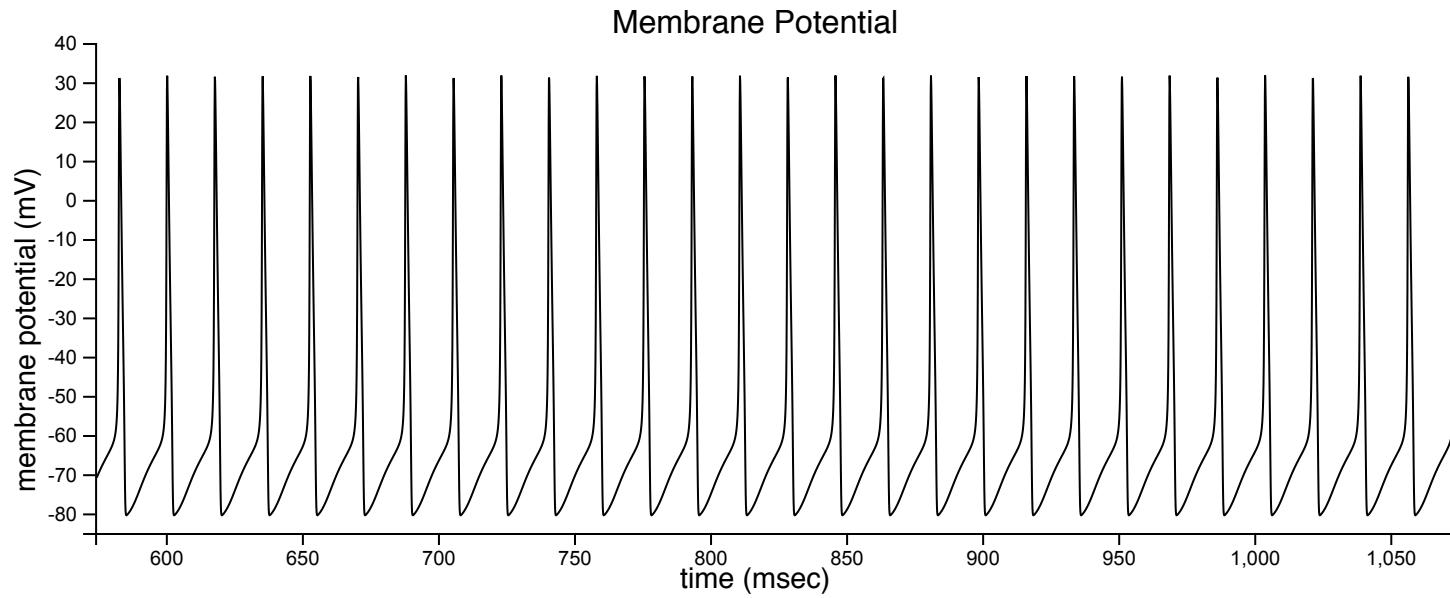
Andrew Fielding Huxley
Prize share: 1/3

Action potential

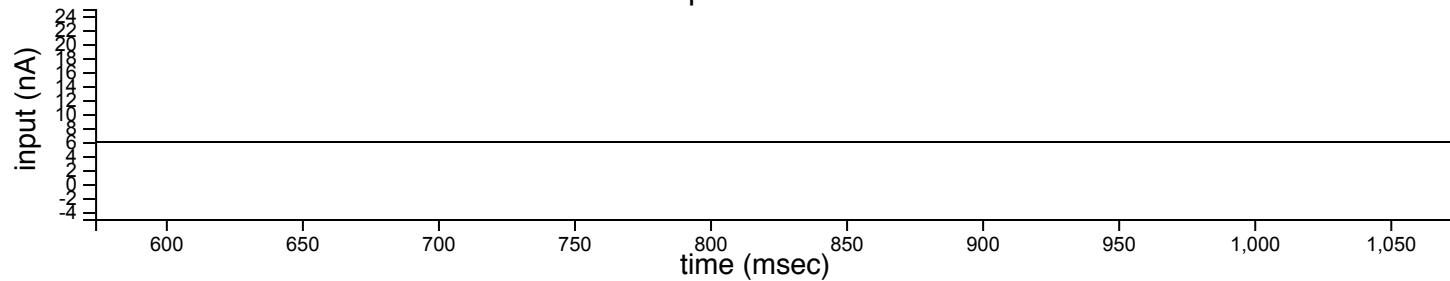


Example: Hodgkin-Huxley model

input current (nA): 6



Input Current



Simulator by Jack Terwilliger

<http://jackterwilliger.com/biological-neural-networks-part-i-spiking-neurons/>

Example: Hodgkin-Huxley model

$$c_M \frac{dV}{dt} = -\bar{g}_{Na} m^3 h (V - E_{Na}) - \bar{g}_K n^4 (V - E_K) - \bar{g}_L (V - E_L),$$

Hodgkin, A. L., & Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol*, 117(4).

Example: Hodgkin-Huxley model

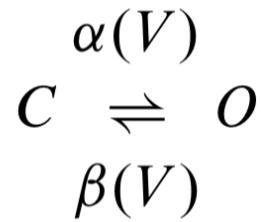
$$c_M \frac{dV}{dt} = -\bar{g}_{Na} m^3 h (V - E_{Na}) - \bar{g}_K n^4 (V - E_K) - \bar{g}_L (V - E_L),$$

$$\begin{array}{ccc} \alpha(V) \\ C \rightleftharpoons O \\ \beta(V) \end{array}$$

Hodgkin, A. L., & Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. J Physiol, 117(4).

Example: Hodgkin-Huxley model

$$c_M \frac{dV}{dt} = -\bar{g}_{Na} m^3 h (V - E_{Na}) - \bar{g}_K n^4 (V - E_K) - \bar{g}_L (V - E_L),$$



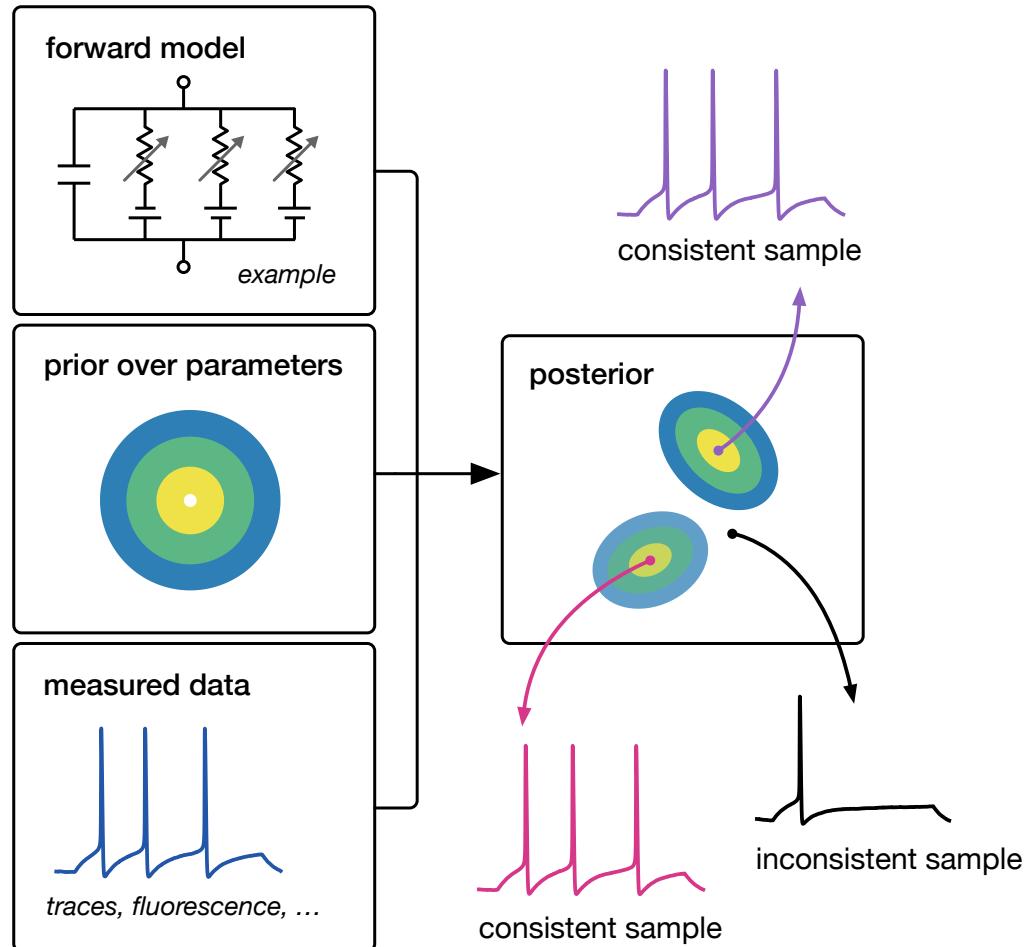
$$\frac{dn}{dt} = \phi[\alpha_n(V)(1-n) - \beta_n(V)n],$$

$$\frac{dm}{dt} = \phi[\alpha_m(V)(1-m) - \beta_m(V)m],$$

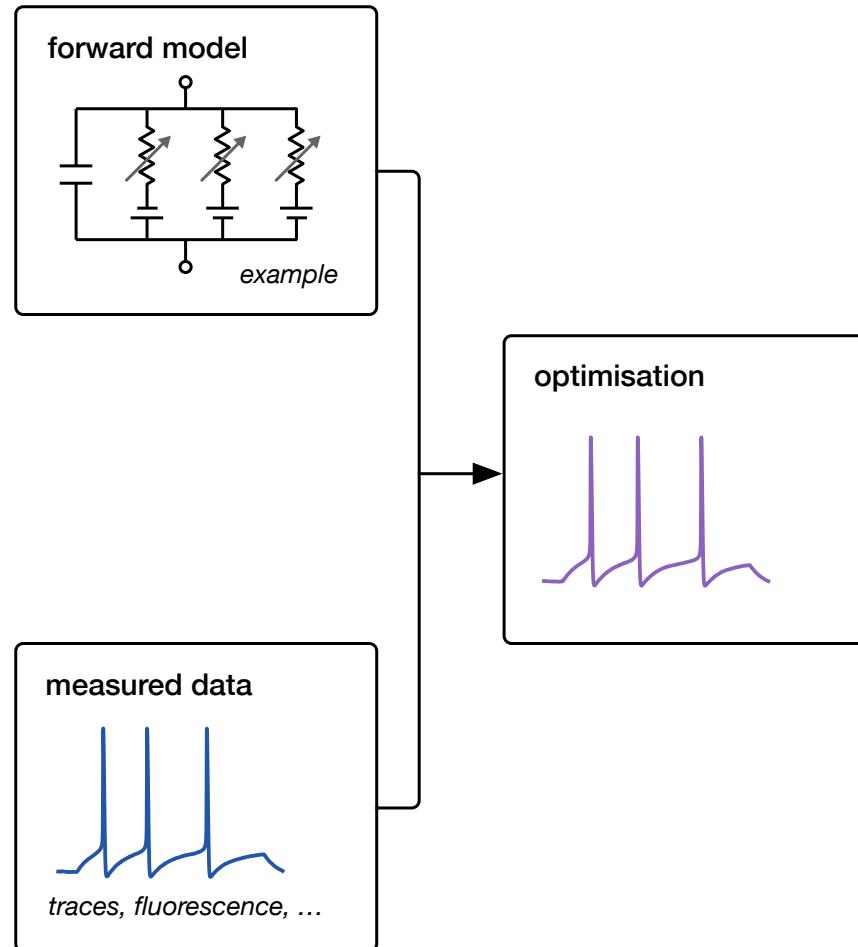
$$\frac{dh}{dt} = \phi[\alpha_h(V)(1-h) - \beta_h(V)h].$$

Hodgkin, A. L., & Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol*, 117(4).

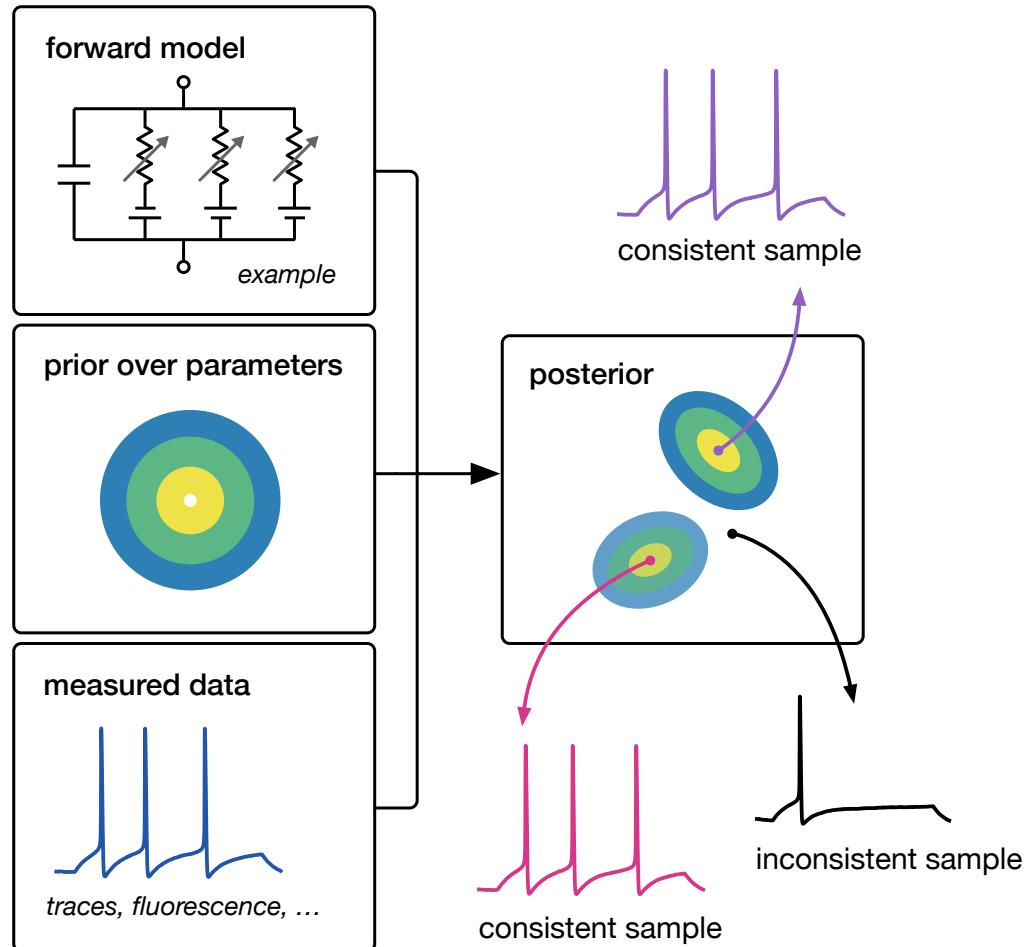
Goal: Bayesian inference for mechanistic models



Goal: Bayesian inference for mechanistic models

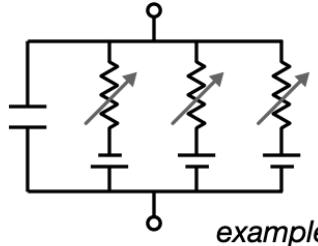


Goal: Bayesian inference for mechanistic models

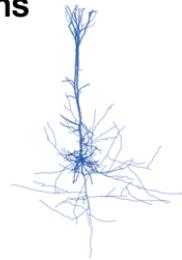


Goal: Bayesian inference for mechanistic models

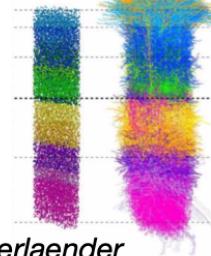
forward model



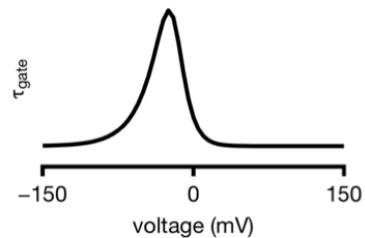
multi-compartmental neurons



connectomics



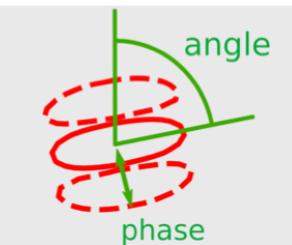
ion-channel kinetics



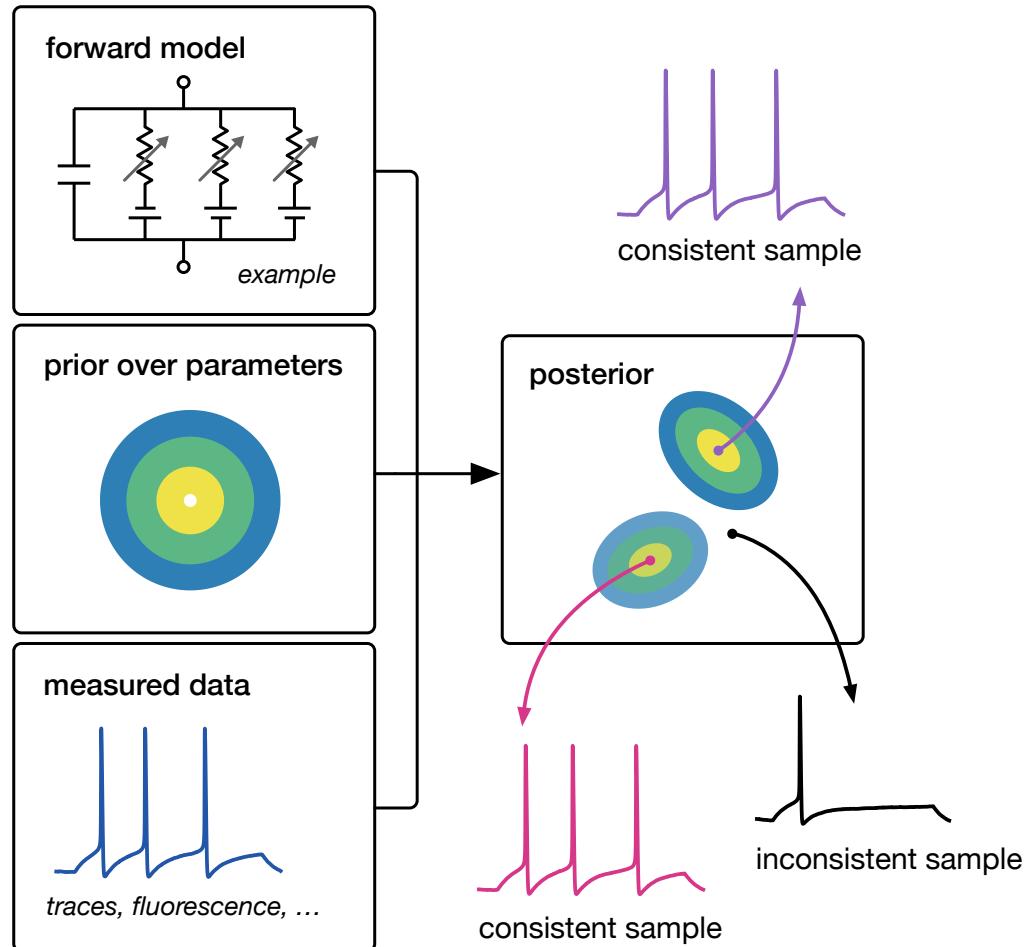
spiking networks



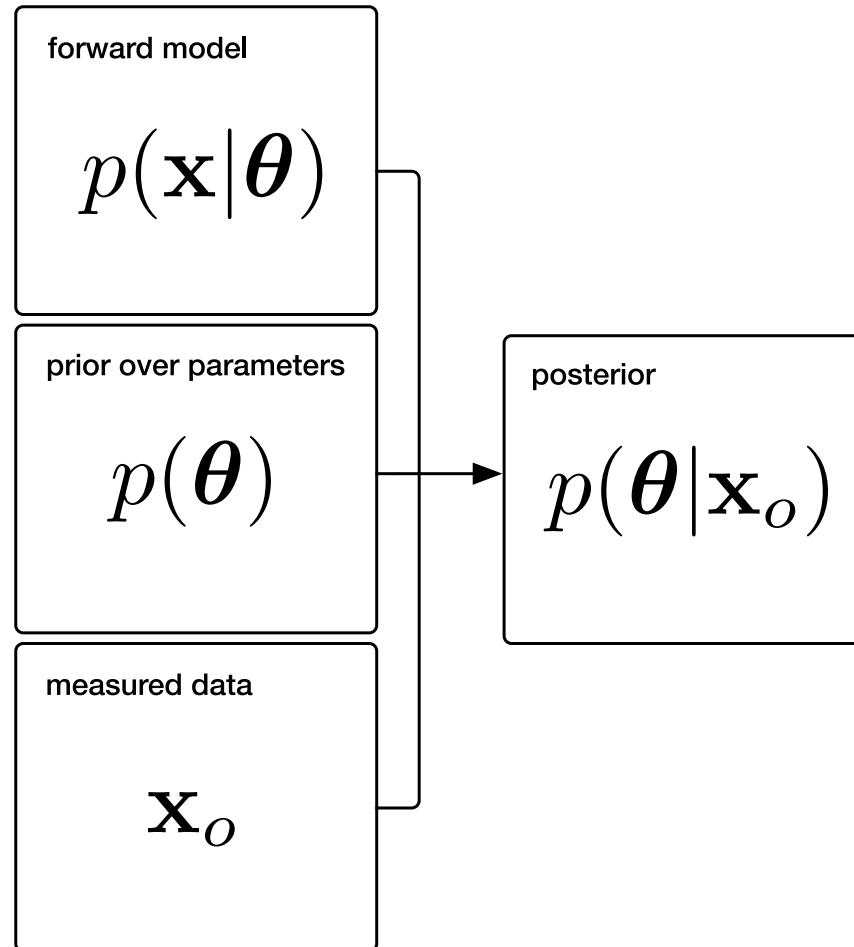
receptive field models



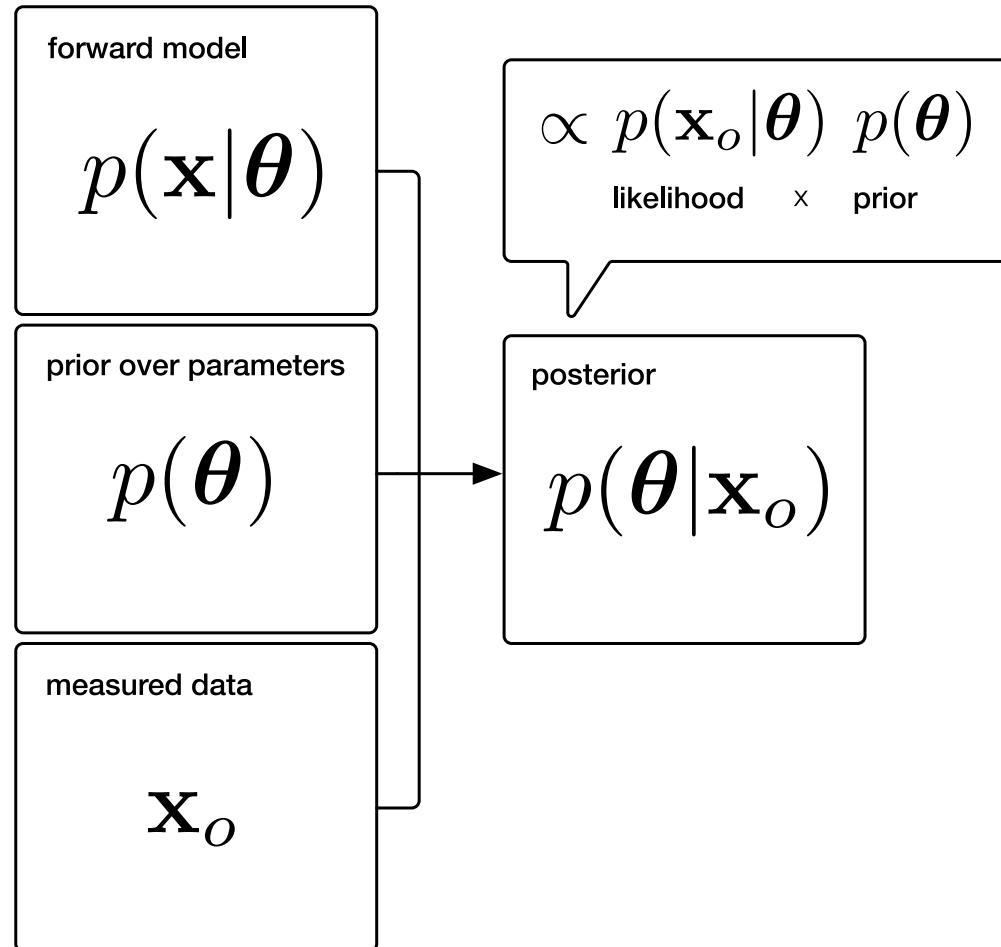
Goal: Bayesian inference for mechanistic models



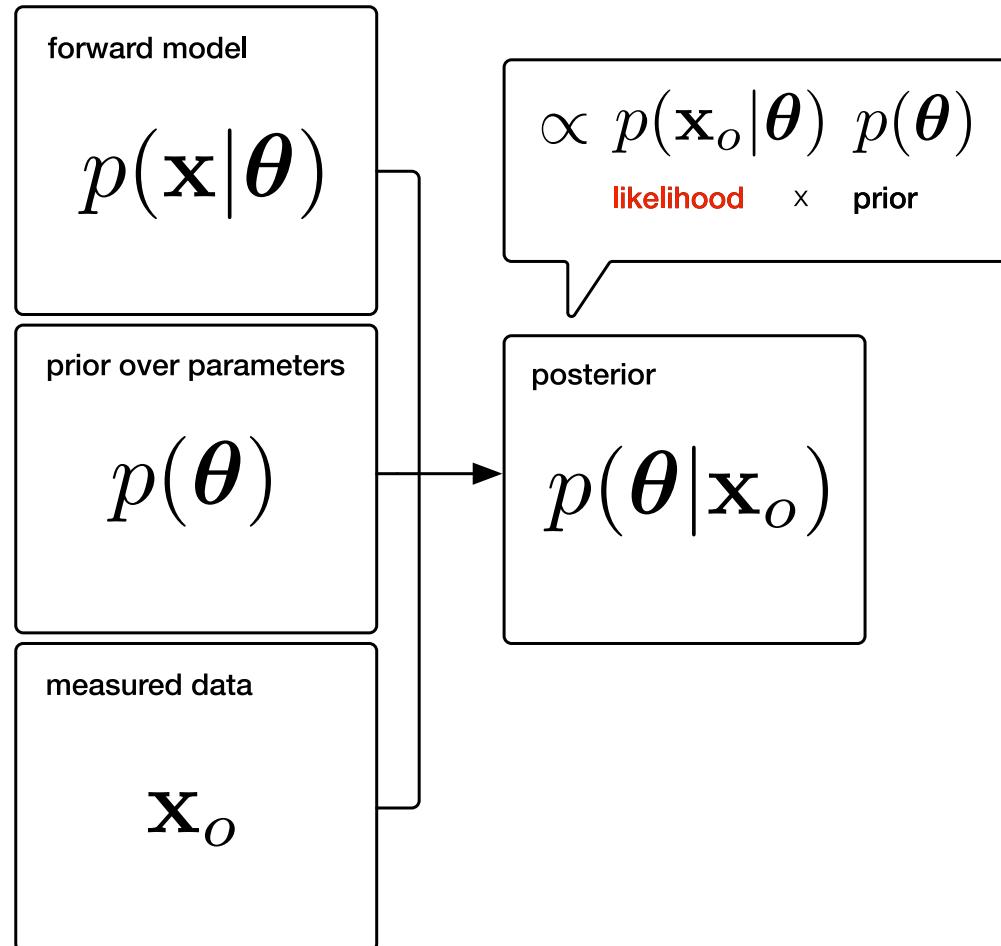
Goal: Bayesian inference for mechanistic models



Goal: Bayesian inference for mechanistic models



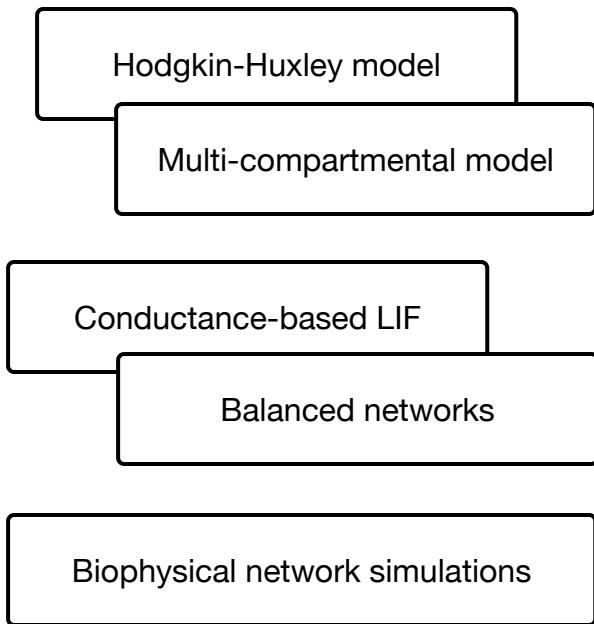
Goal: Bayesian inference for mechanistic models



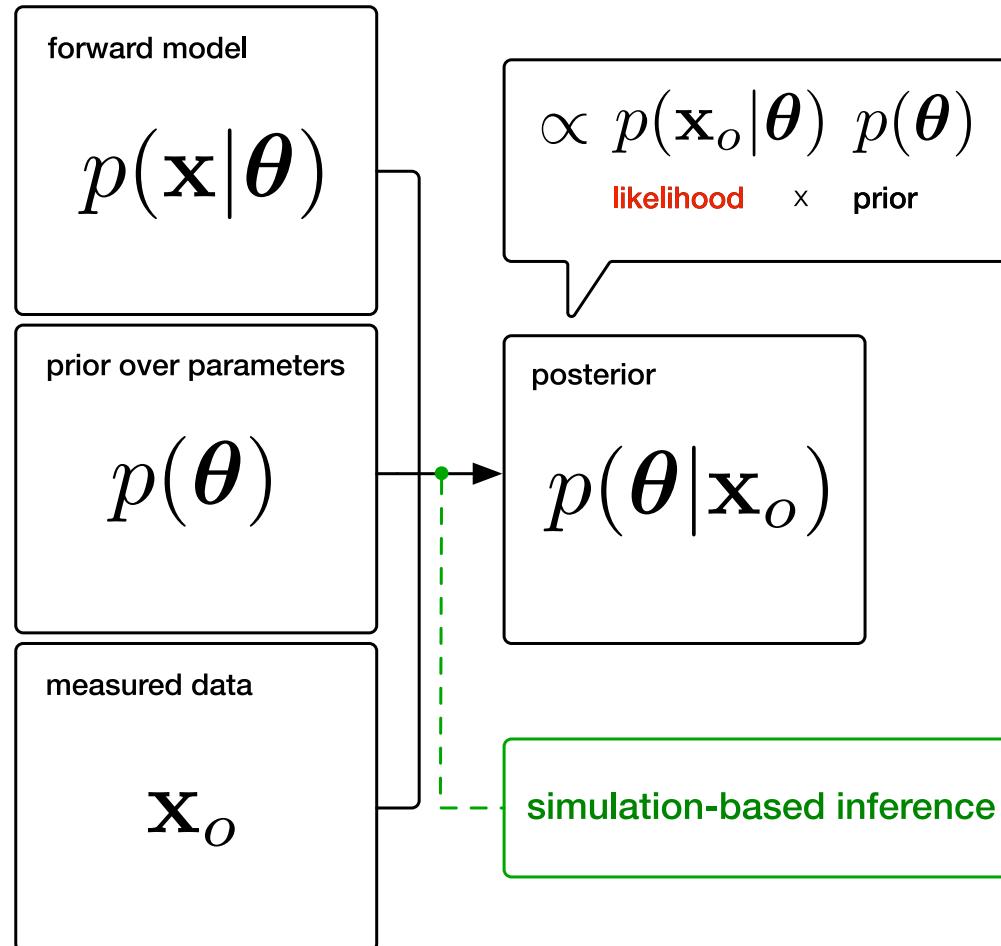
Mechanistic models

vs.

statistical/ML models



Goal: Bayesian inference for mechanistic models

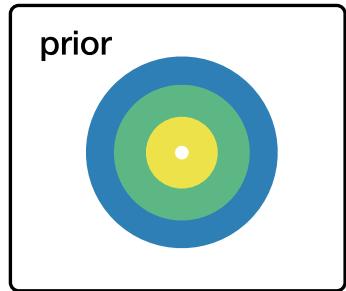


Learning the posterior

Simple algorithm:

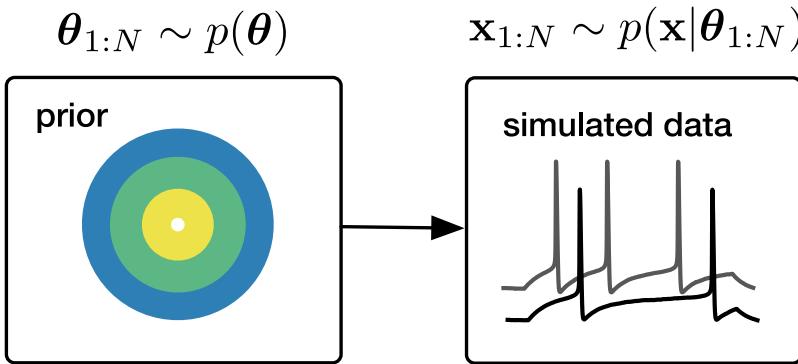
Simulation-based inference by density estimation

$$\theta_{1:N} \sim p(\theta)$$



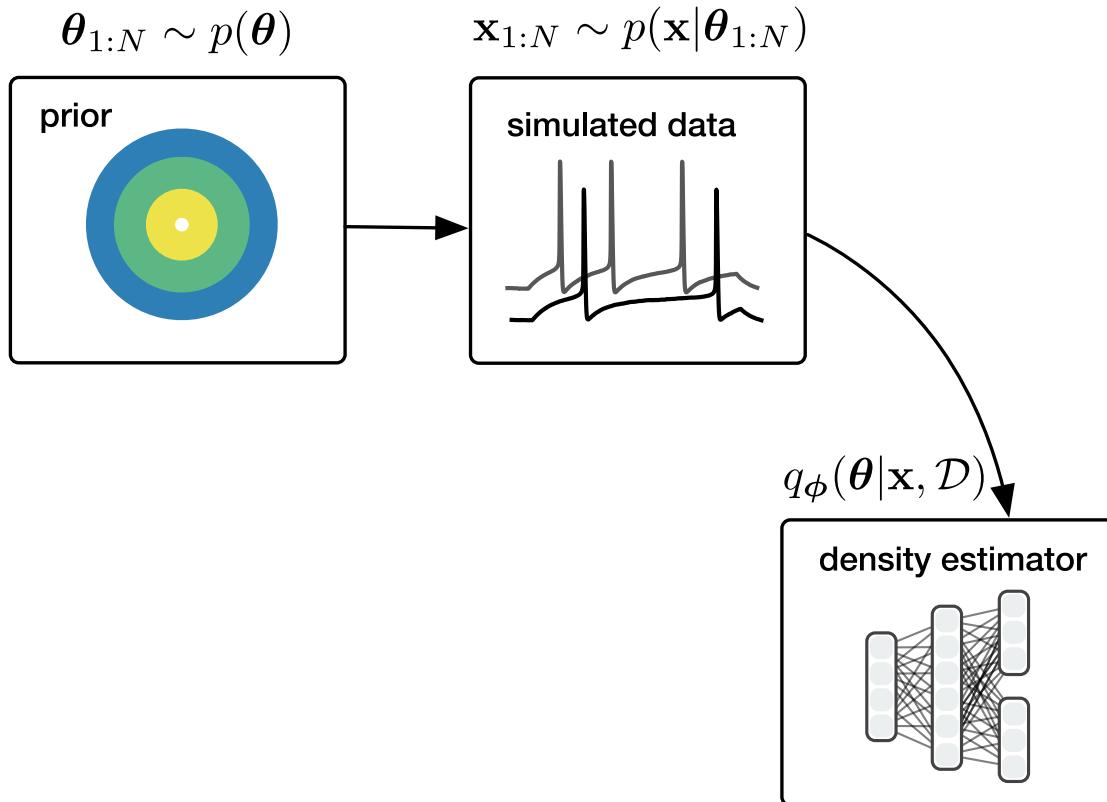
Simple algorithm:

Simulation-based inference by density estimation



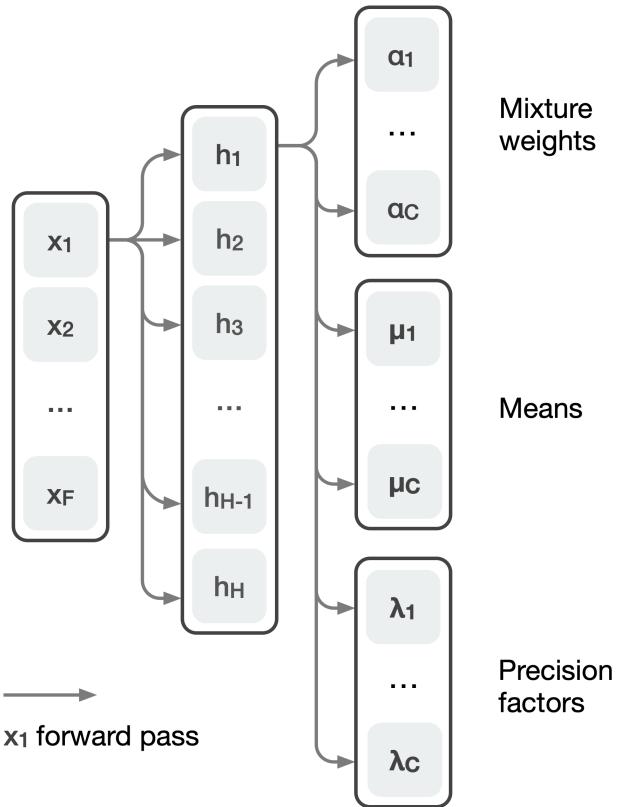
Simple algorithm:

Simulation-based inference by density estimation



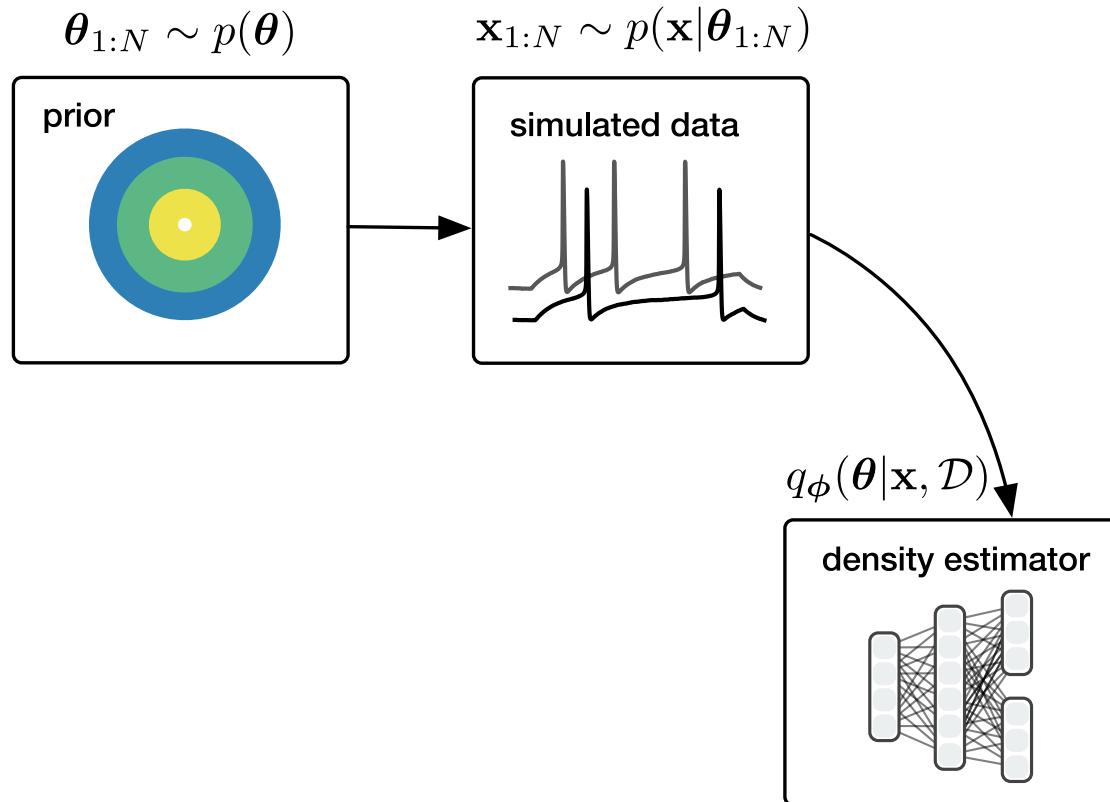
Mixture density network

$$q_{\phi}(\theta|x) = \sum_c \alpha_c(x) \mathcal{N}(\theta|\mu_c(x), \Sigma_c(x))$$



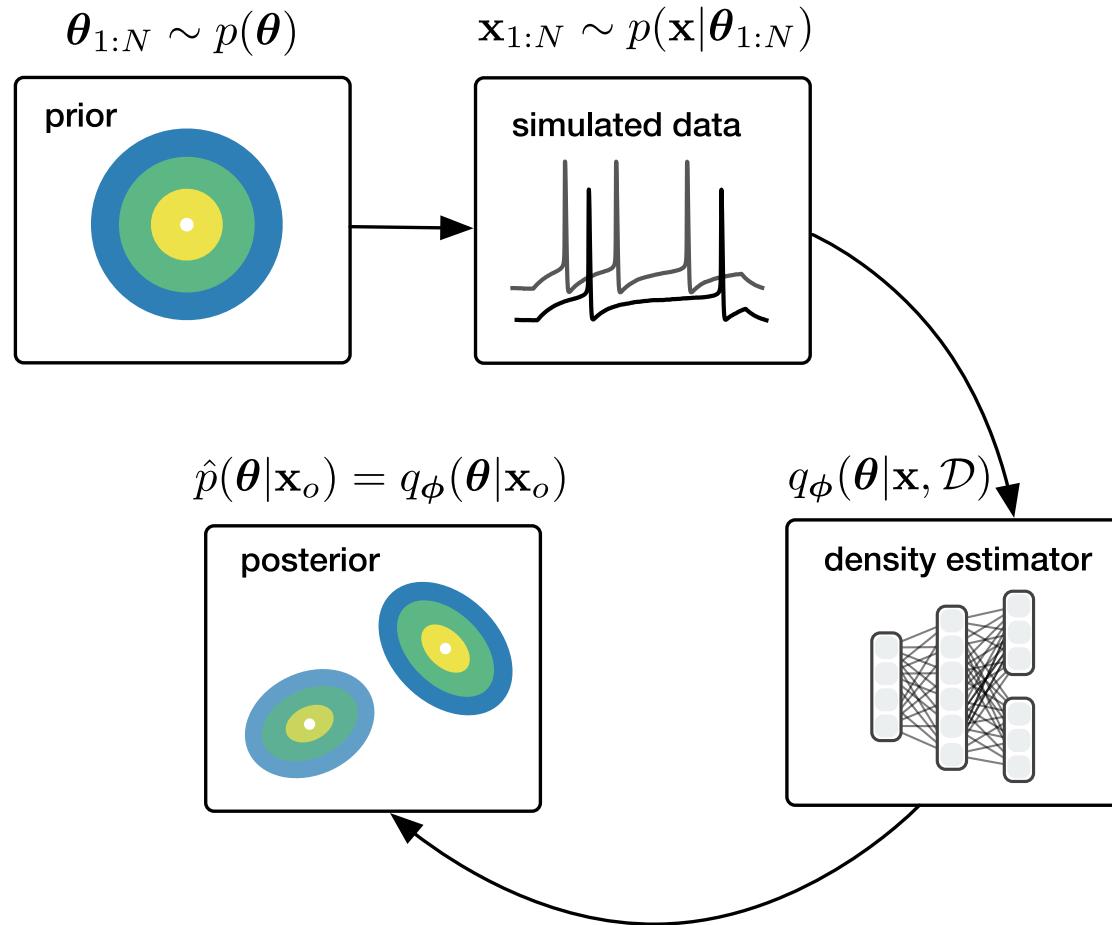
Simple algorithm:

Simulation-based inference by density estimation



Simple algorithm:

Simulation-based inference by density estimation

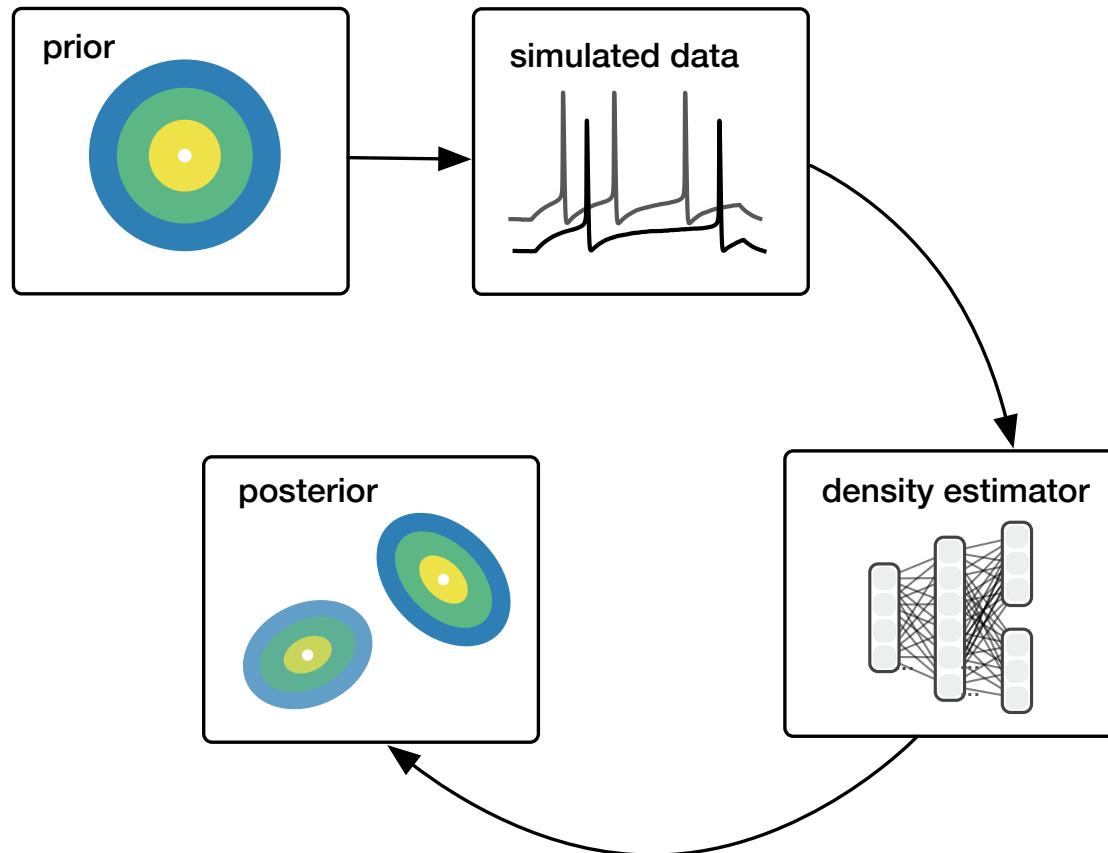


We got limited computational budget for running simulations.

**How can we be efficient
in the # of simulations?**

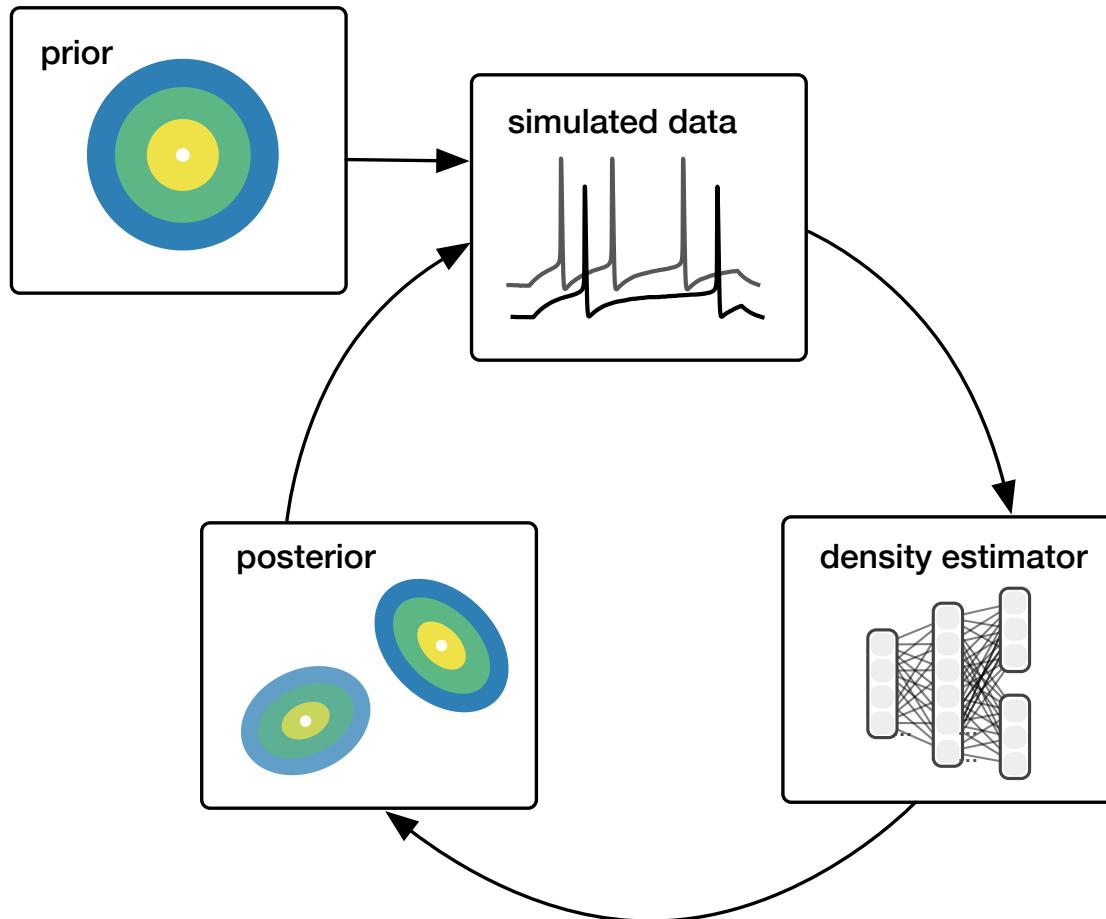
Simple algorithm:

Simulation-based inference by density estimation



Sequential algorithm:

Simulation-based inference by density estimation



1) How do we correct for using a proposal distribution $\tilde{p}(\theta)$?

SNPE: Sequential Neural Posterior Estimation

$$\mathcal{L}(\phi) = -\frac{1}{N} \sum_n \frac{p(\theta_n)}{\tilde{p}(\theta_n)} \log q_\phi(\theta_n | x_n)$$

- correct for proposal distribution in the loss function

Lueckmann*, J.-M., Goncalves*, P. J., Bassetto, G., Öcal, K., Nonnenmacher, M., & Macke, J. H. (2017). Flexible statistical inference for mechanistic models of neural dynamics. In Advances in Neural Information Processing Systems 30 (pp. 1289–1299). Curran Associates, Inc.

Papamakarios, G., & Murray, I. (2016). Fast ϵ -free inference of simulation models with bayesian conditional density estimation. In Advances in Neural Information Processing Systems 29 (pp. 1028–1036). Curran Associates, Inc.

SNPE: Sequential Neural Posterior Estimation

$$\mathcal{L}(\phi) = -\frac{1}{N} \sum_n \frac{p(\theta_n)}{\tilde{p}(\theta_n)} \log q_\phi(\theta_n | x_n)$$

- correct for proposal distribution in the loss function
- previous work: post-hoc analytical correction (Papamakarios & Murray, 2016), which can lead to numerical issues in practice

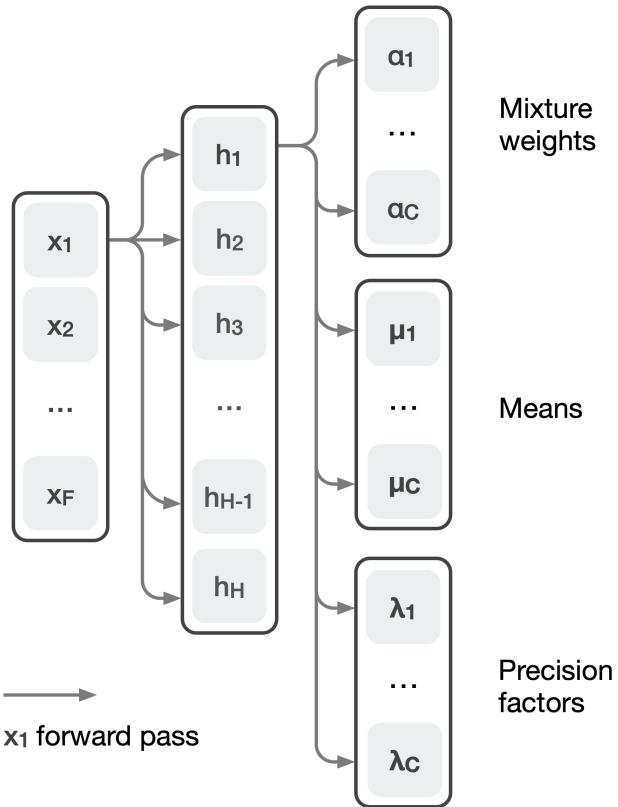
Lueckmann*, J.-M., Goncalves*, P. J., Bassetto, G., Öcal, K., Nonnenmacher, M., & Macke, J. H. (2017). Flexible statistical inference for mechanistic models of neural dynamics. In Advances in Neural Information Processing Systems 30 (pp. 1289–1299). Curran Associates, Inc.

Papamakarios, G., & Murray, I. (2016). Fast ϵ -free inference of simulation models with bayesian conditional density estimation. In Advances in Neural Information Processing Systems 29 (pp. 1028–1036). Curran Associates, Inc.

2) How do we update the density estimator with new training data?

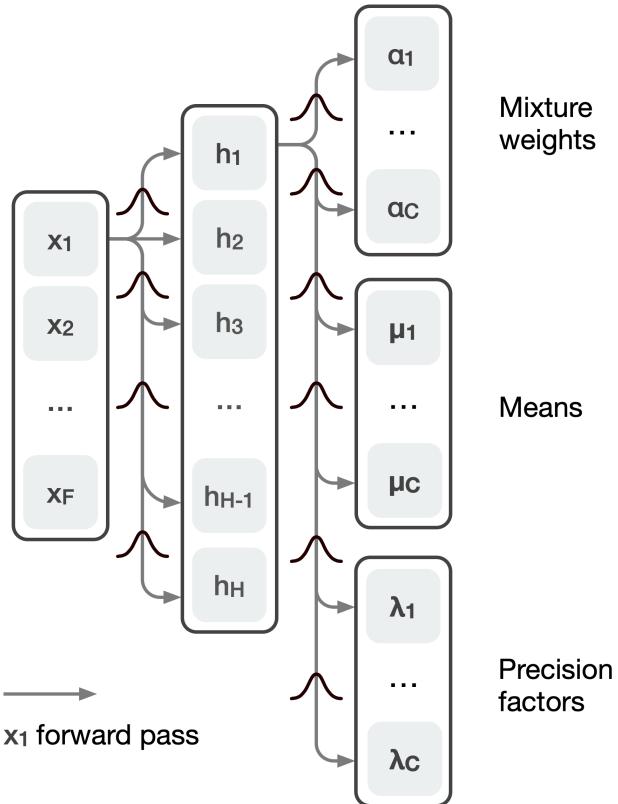
Mixture density network

$$q_{\phi}(\theta|x) = \sum_c \alpha_c(x) \mathcal{N}(\theta|\mu_c(x), \Sigma_c(x))$$



Bayesian Mixture density network

Assume distribution over network weights: $\pi_\phi(w) = \mathcal{N}(w|\phi_m; \phi_s^2)$.



Bayesian Mixture density network

Assume distribution over network weights: $\pi_\phi(w) = \mathcal{N}(w|\phi_m; \phi_s^2)$.

Training via variational inference:

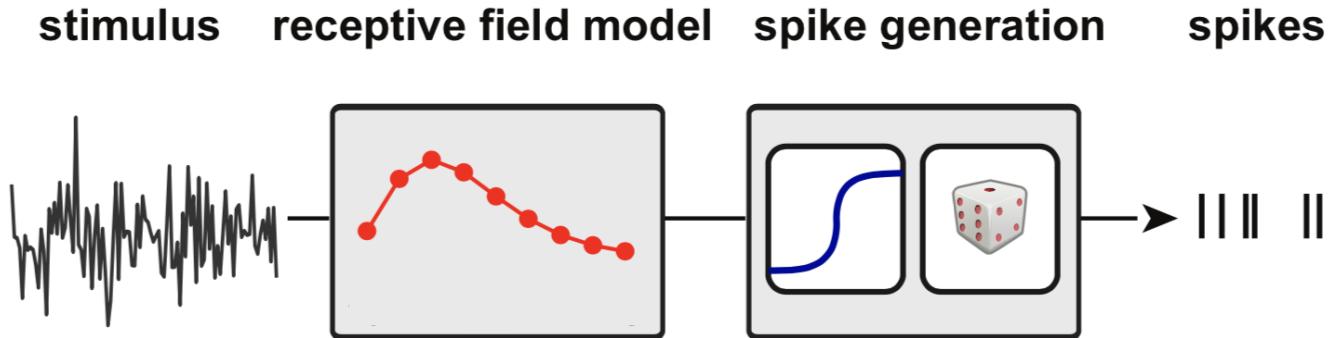
$$\begin{aligned}\phi^* &= \arg \min_{\phi} D_{KL}(\pi_\phi(\mathbf{w}) || p(\mathbf{w} | \mathcal{D})) \\ &= \arg \min_{\phi} \int \pi_\phi(\mathbf{w}) \log \frac{\pi_\phi(\mathbf{w})}{p(\mathbf{w})p(\mathcal{D}|\mathbf{w})} d\mathbf{w} \\ &= \arg \min_{\phi} -\langle \log p(\mathcal{D}|\mathbf{w}) \rangle_{\pi_\phi(\mathbf{w})} + D_{KL}(\pi_\phi(\mathbf{w}) || p(\mathbf{w}))\end{aligned}$$

SNPE: Sequential Neural Posterior Estimation

$$\begin{aligned}\mathcal{L}(\phi^{(r)}) = & -\frac{1}{N} \sum_n \frac{p(\theta_n)}{\tilde{p}^{(r)}(\theta_n)} \langle \log q_w(\theta_n | x_n) \rangle_{\pi_{\phi^{(r)}}(w)} \\ & + \frac{1}{N} \text{KL} (\pi_{\phi^{(r)}}(w) || \pi_{\phi^{(r-1)}}(w))\end{aligned}$$

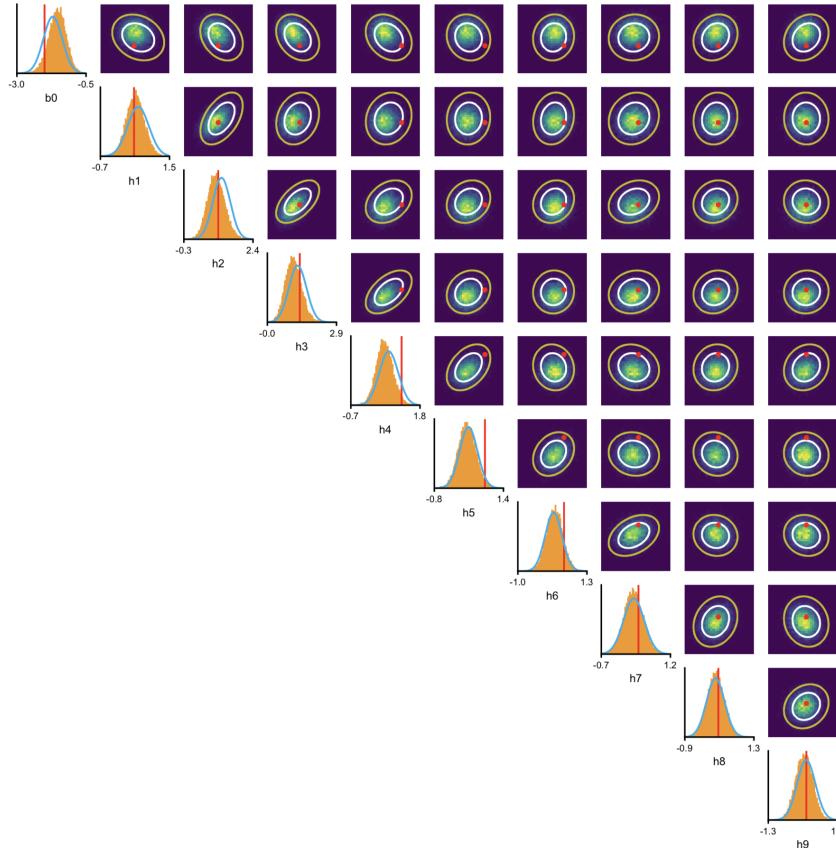
- KL-term acts as a regularizer: Keeps weights close to previous round

Bernoulli GLM (Generalised linear model)



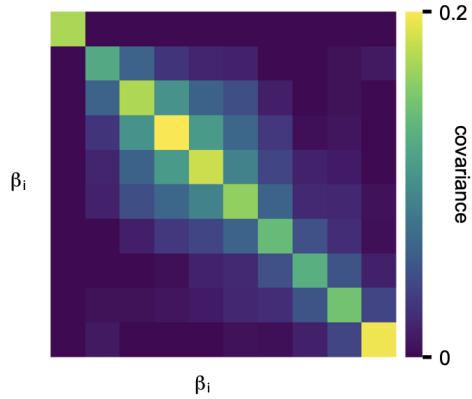
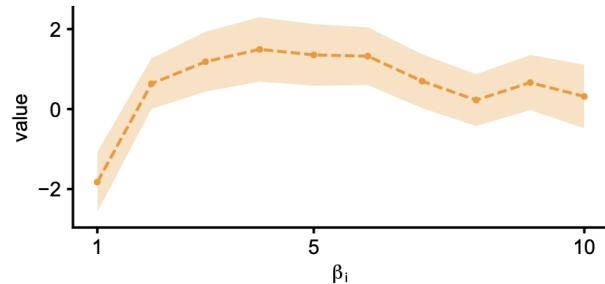
Bernoulli GLM

Posterior: Space of parameters consistent with prior and data



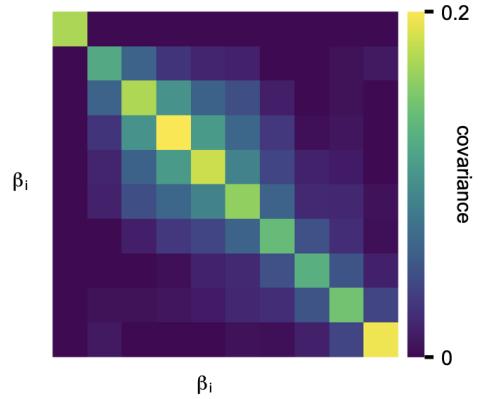
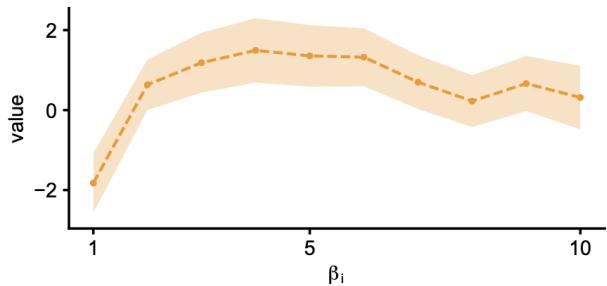
Bernoulli GLM

True posterior

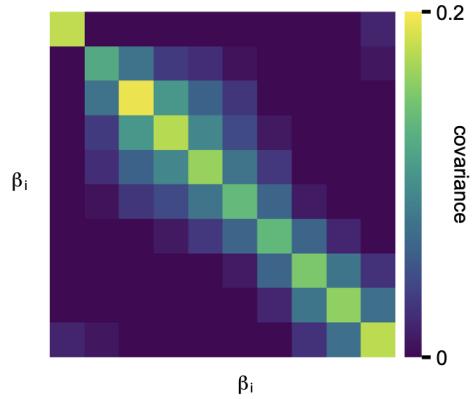
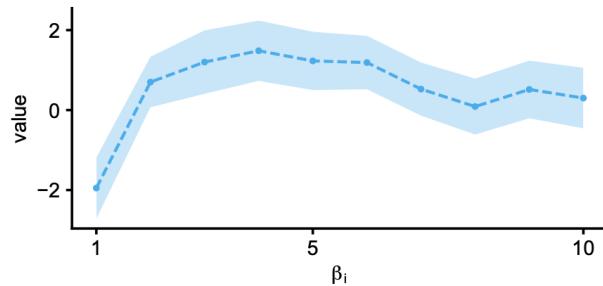


Bernoulli GLM

True posterior

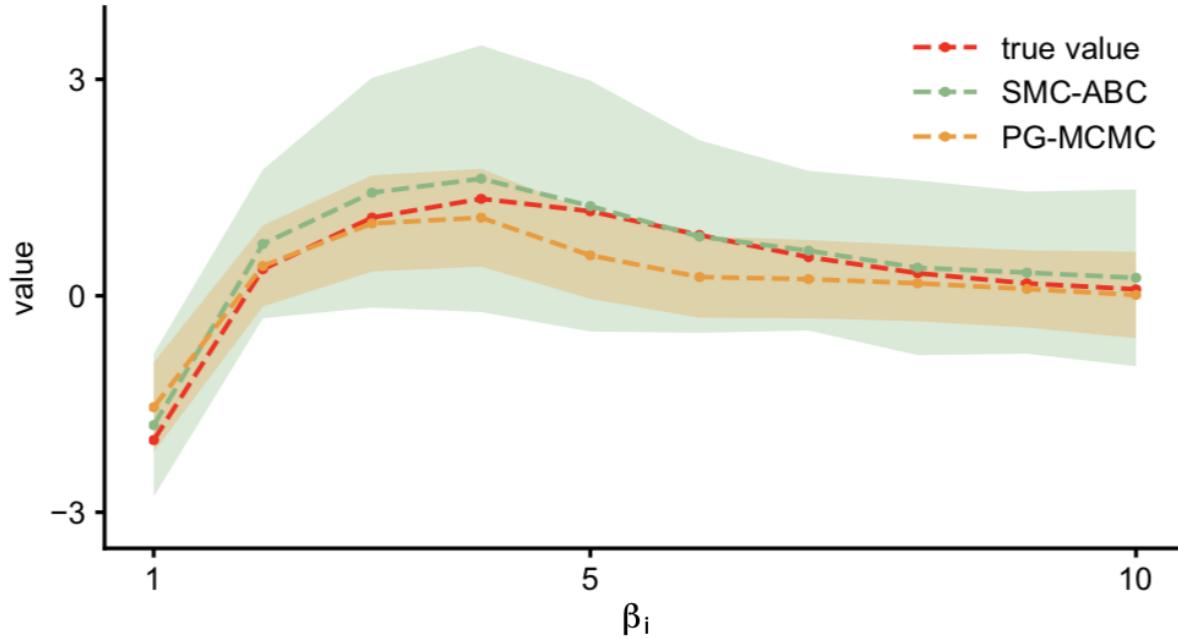


SNPE posterior



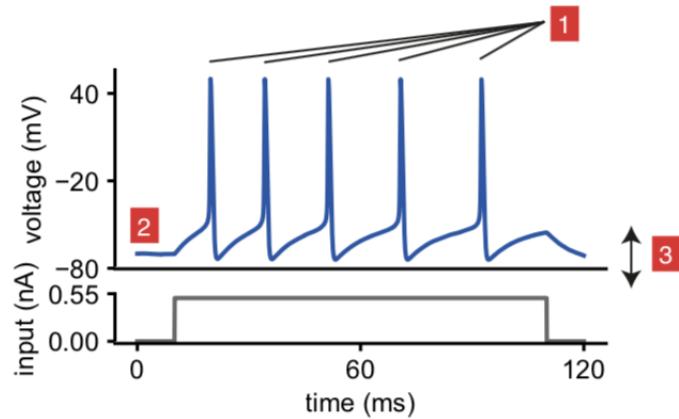
Bernoulli GLM

SMC-ABC needs significantly more simulations



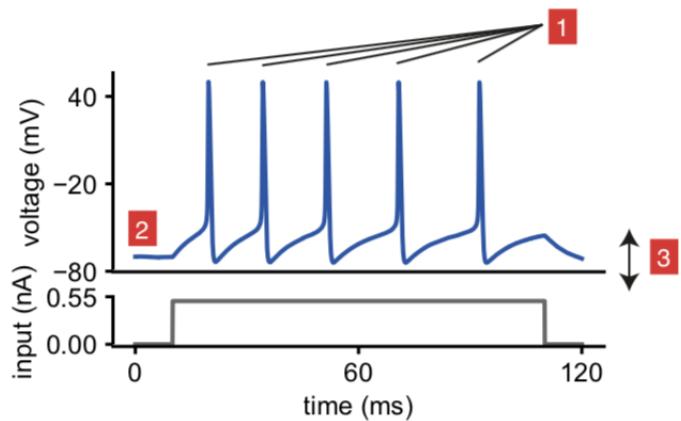
Hodgkin-Huxley

original data (simulated)

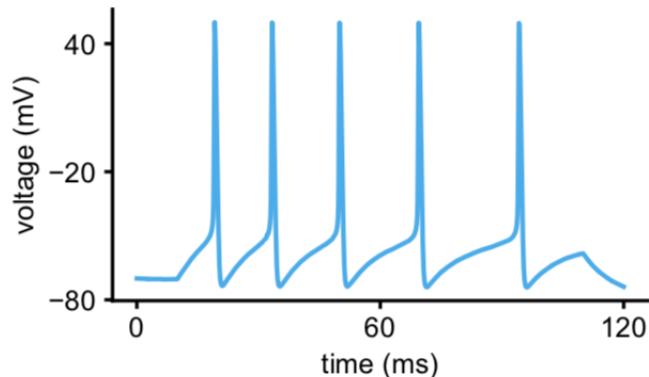


Hodgkin-Huxley

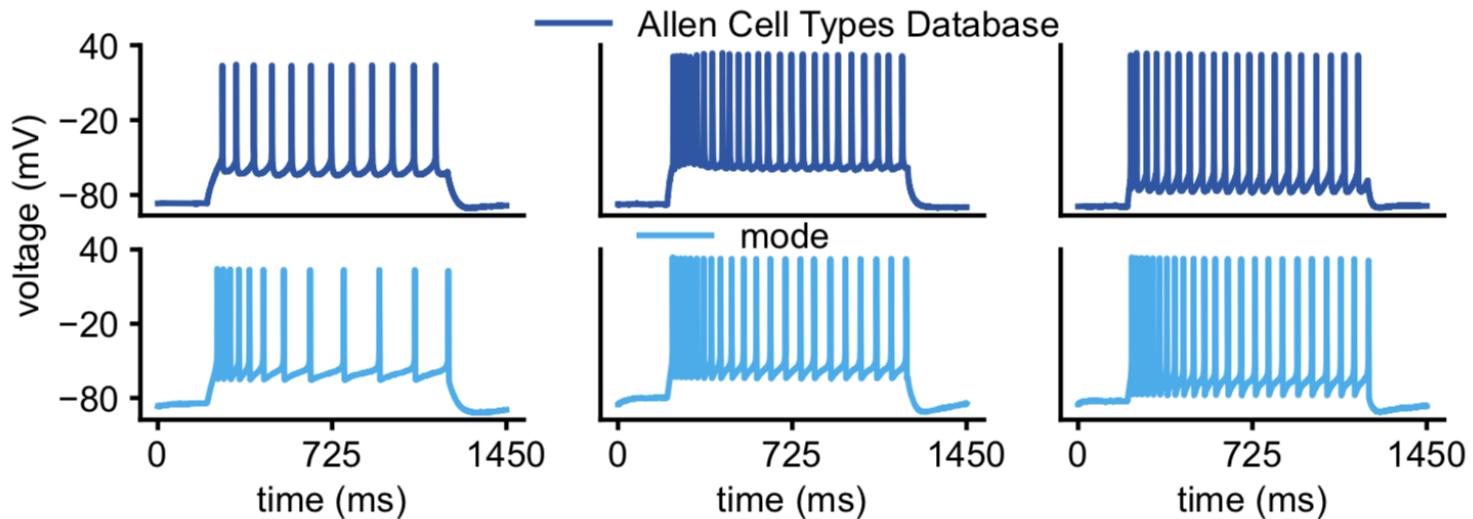
original data (simulated)



data generated by inferred model

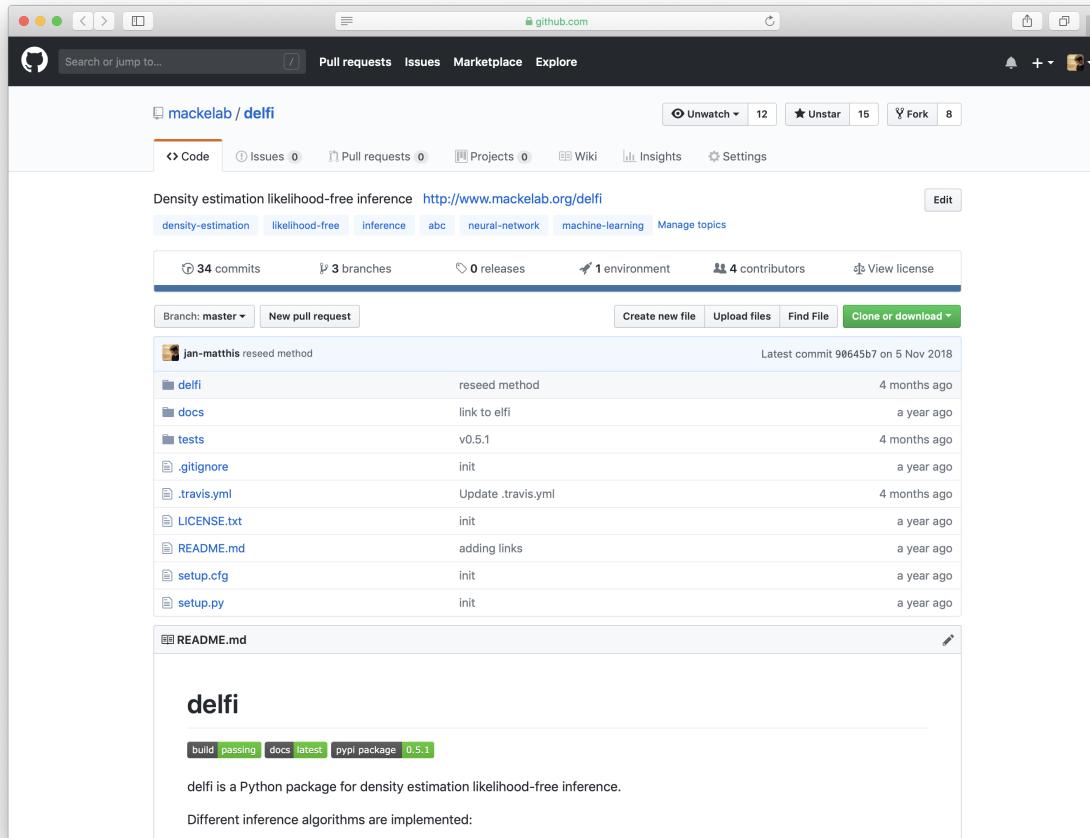


Hodgkin-Huxley



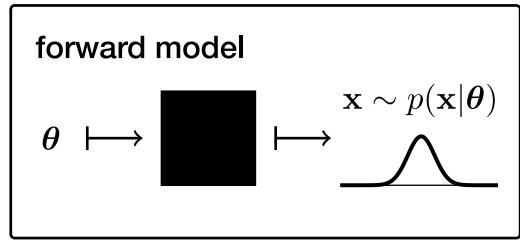
... there's more

- No need for hand-crafted summary statistics
- Avoiding regions of bad parameters

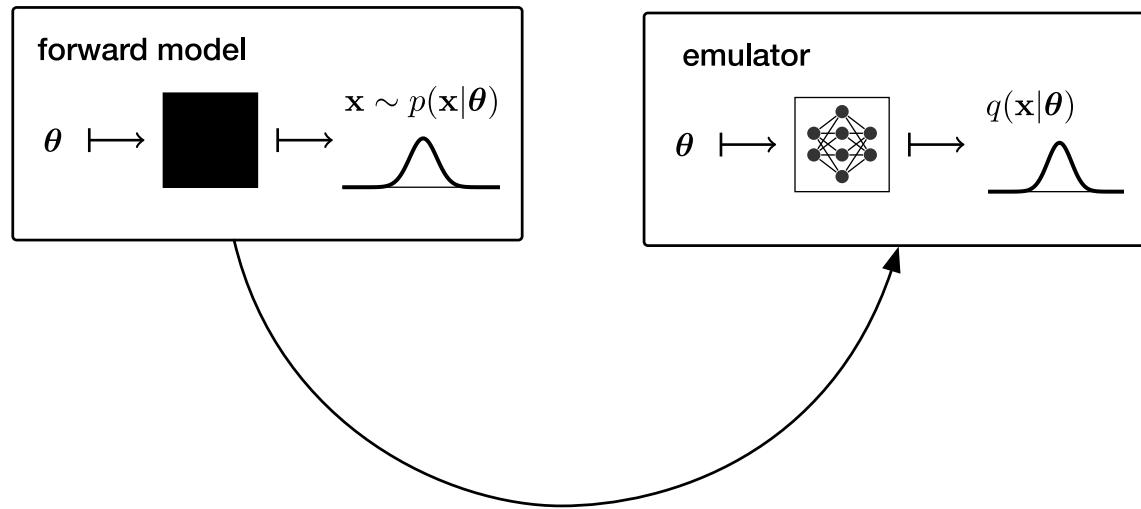


Learning the likelihood

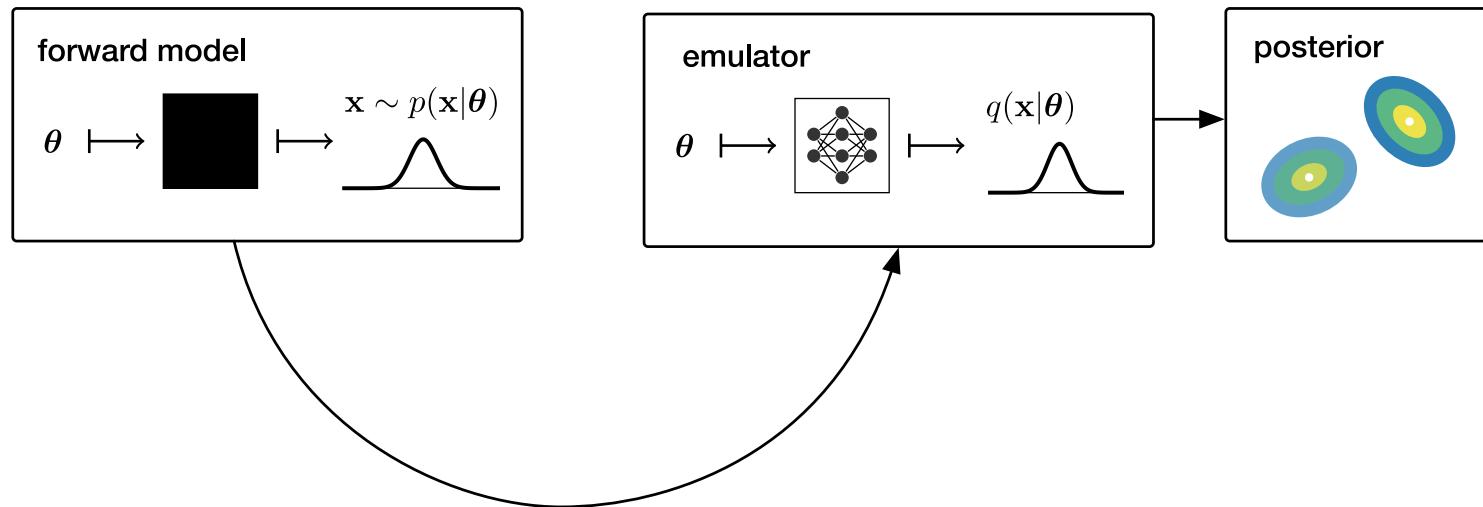
General approach



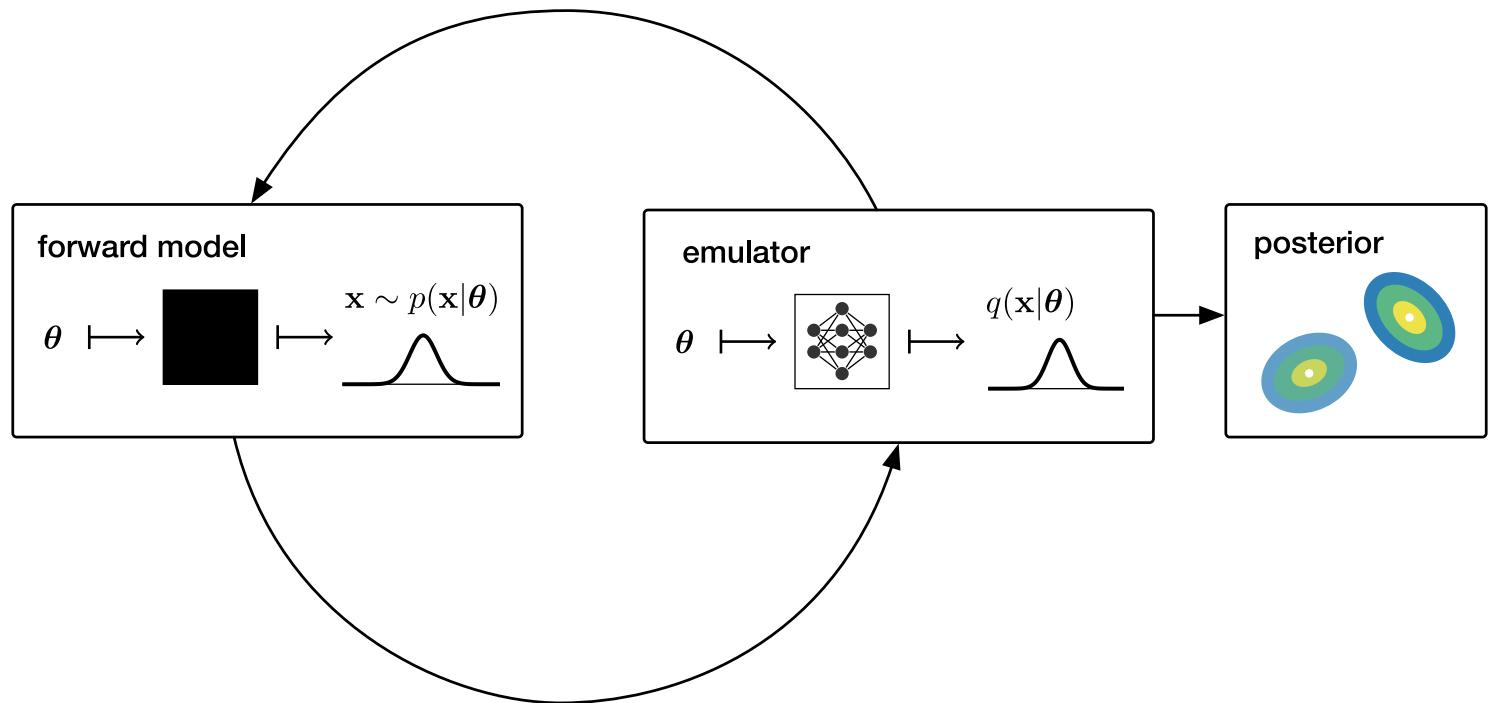
General approach



General approach



Sequential approach



Acquisition rules:

1. Learning local emulators

- Maximum variance in approximate posterior
- Takes into account measured data \mathbf{x}_o
- Good approximation of $p(\mathbf{x}|\boldsymbol{\theta})$ by q around \mathbf{x}_o

$$\begin{aligned}\boldsymbol{\theta}^* &= \arg \max_{\boldsymbol{\theta}} \mathbb{V}_{\phi|\mathcal{D}}[\tilde{p}(\boldsymbol{\theta}|\mathbf{x}_o, \boldsymbol{\phi})] \\ &= \arg \max_{\boldsymbol{\theta}} p(\boldsymbol{\theta})^2 \mathbb{V}_{\phi|\mathcal{D}}[\hat{\mathcal{L}}(\boldsymbol{\theta})] \\ &= \arg \max_{\boldsymbol{\theta}} \log p(\boldsymbol{\theta}) + \log \sqrt{\mathbb{V}_{\phi|\mathcal{D}}[\hat{\mathcal{L}}(\boldsymbol{\theta})]}.\end{aligned}$$

Acquisition rules:

1. Learning local emulators

- Maximum variance in approximate posterior
- Takes into account measured data \mathbf{x}_o
- Good approximation of $p(\mathbf{x}|\boldsymbol{\theta})$ by q around \mathbf{x}_o

$$\begin{aligned}\boldsymbol{\theta}^* &= \arg \max_{\boldsymbol{\theta}} \mathbb{V}_{\phi|\mathcal{D}}[\tilde{p}(\boldsymbol{\theta}|\mathbf{x}_o, \boldsymbol{\phi})] \\ &= \arg \max_{\boldsymbol{\theta}} p(\boldsymbol{\theta})^2 \mathbb{V}_{\phi|\mathcal{D}}[\hat{\mathcal{L}}(\boldsymbol{\theta})] \\ &= \arg \max_{\boldsymbol{\theta}} \log p(\boldsymbol{\theta}) + \log \sqrt{\mathbb{V}_{\phi|\mathcal{D}}[\hat{\mathcal{L}}(\boldsymbol{\theta})]}.\end{aligned}$$

2. Learning global emulators

- Bayesian Active Learning by Disagreement
- Does not take account particular \mathbf{x}_o
- Good approximation of $p(\mathbf{x}|\boldsymbol{\theta})$ by q for all \mathbf{x}

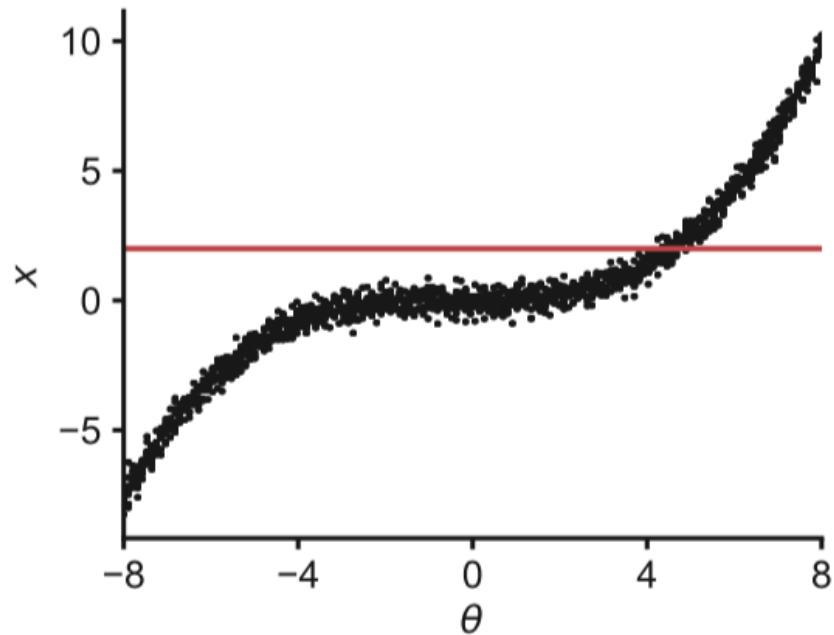
$$\begin{aligned}\boldsymbol{\theta}^* &= \arg \max_{\boldsymbol{\theta}} \mathbb{I}[\mathbf{x}, \boldsymbol{\phi} | \boldsymbol{\theta}, \mathcal{D}] \\ &= \arg \max_{\boldsymbol{\theta}} \underbrace{\mathbb{H}[\mathbf{x} | \boldsymbol{\theta}, \mathcal{D}]}_{\text{entropy}} - \underbrace{\mathbb{E}_{\boldsymbol{\phi}|\mathcal{D}} [\mathbb{H}[\mathbf{x} | \boldsymbol{\theta}, \boldsymbol{\phi}]]}_{\text{expected conditional entropy}}\end{aligned}$$

Lueckmann, J.-M., Bassetto, G., Karaletsos, T., & Macke, J. H. (2018). Likelihood-free inference with emulator networks. In AABI 2018, PMLR volume 96.

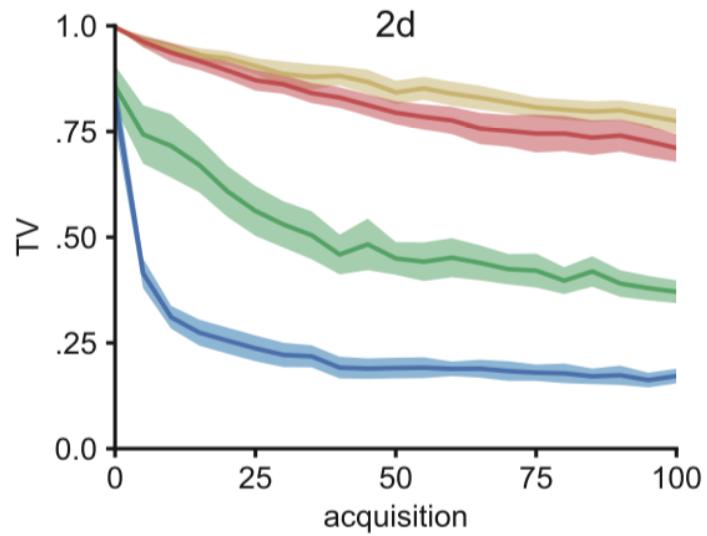
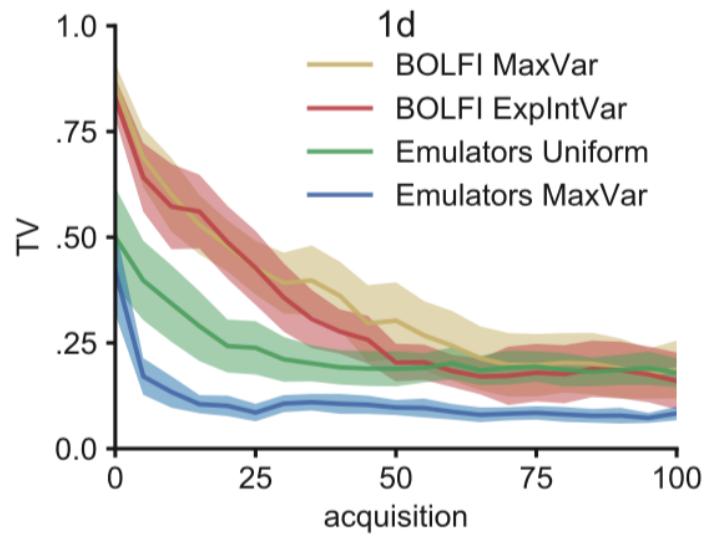
Houlsby, N., Huszár, F., Ghahramani, Z., & Lengyel, M. (2011). Bayesian active learning for classification and preference learning. ArXiv:1112.5745 [Cs, Stat]

Local emulators: Comparison to BOLFI

BOLFI: Bayesian Optimization for Likelihood-Free Inference



Local emulators: Comparison to BOLFI



Local emulators: Sequential neural likelihood

Sequential Neural Likelihood: Fast Likelihood-free Inference with Autoregressive Flows

George Papamakarios

University of Edinburgh

g.papamakarios@ed.ac.uk

David C. Sterratt

University of Edinburgh

david.c.sterratt@ed.ac.uk

Iain Murray

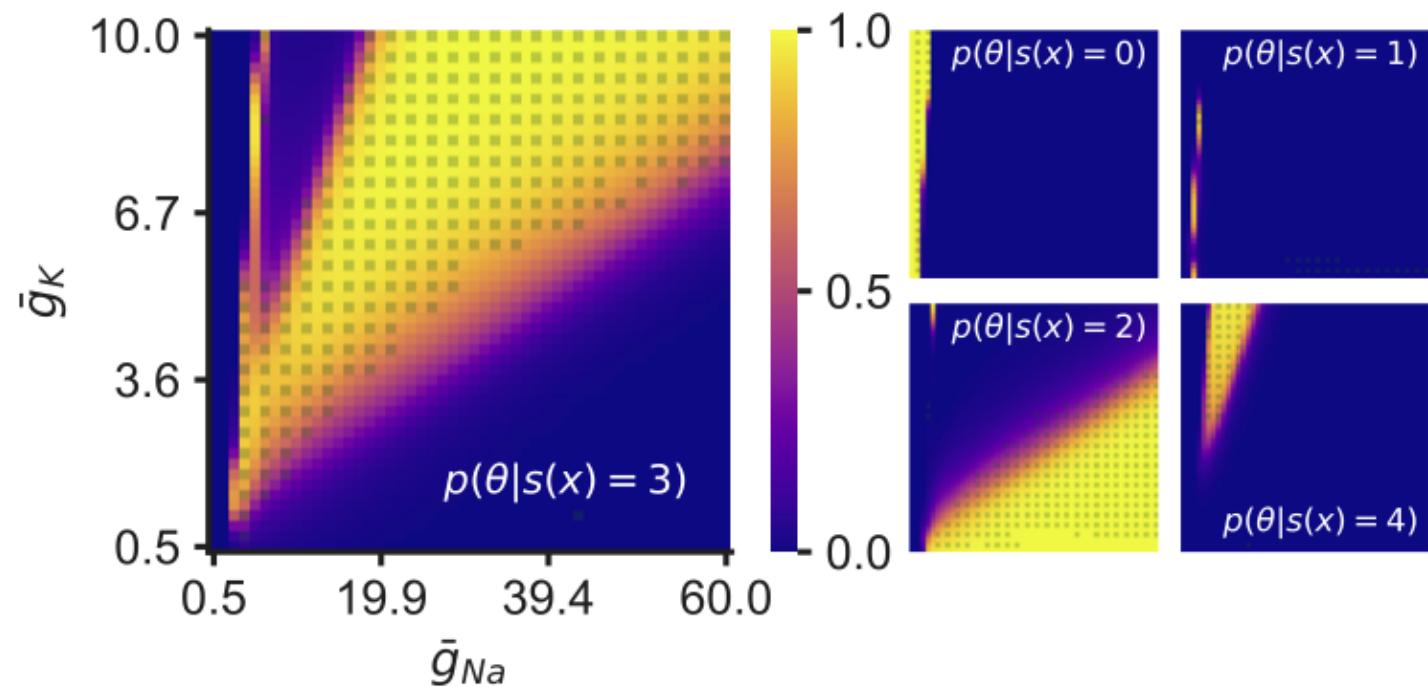
University of Edinburgh

i.murray@ed.ac.uk

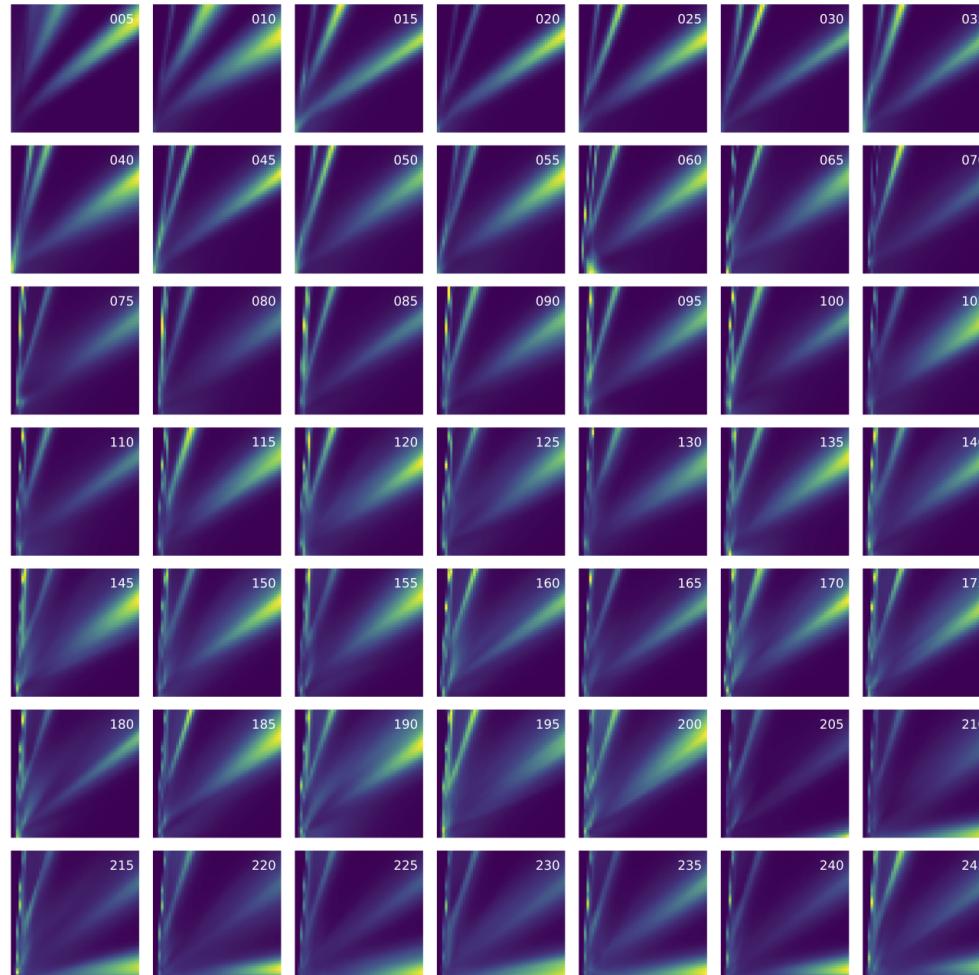
Abstract

We present Sequential Neural Likelihood (SNL), a new method for Bayesian inference in simulator models, where the likelihood is intractable but simulating data from the model is possible. SNL trains an autoregressive flow on simulated data in order to learn a model of the likelihood in the region of high posterior density. A sequential training procedure guides simulations and reduces simulation cost by orders of magnitude. We show that SNL is more robust, more accurate and requires less tuning than related state-of-the-art methods which target the posterior, and discuss diagnostics for assessing calibration, convergence and goodness-of-fit.

Global emulators: Hodgkin-Huxley



Global emulators: Hodgkin-Huxley



Applications

Applications

Channelomics

Collaboration with Tim Vogels' group

Theoretical and Computational Neuroscience



Tim Vogels



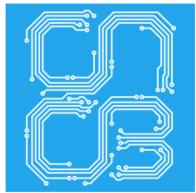
William Podlaski



Chaitanya Chintaluri



UNIVERSITY
OF
OXFORD



Centre
for Neural
Circuits and
Behaviour



Mapping the function of neuronal ion channels in model and experiment

William F Podlaski^{1,2*}, Alexander Seeholzer^{3,4,5†}, Lukas N Groschner^{1,2}, Gero Miesenböck^{1,2}, Rajnish Ranjan⁶, Tim P Vogels^{1,2}

¹Centre for Neural Circuits and Behaviour, University of Oxford, Oxford, United Kingdom; ²Department of Physiology, Anatomy and Genetics, University of Oxford, Oxford, United Kingdom; ³School of Computer and Communication Sciences, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland; ⁴School of Life Sciences, Ecole Polytechnique Federale de Lausanne, Lausanne, Switzerland; ⁵Brain Mind Institute, Ecole Polytechnique Federale de Lausanne, Lausanne, Switzerland;

⁶Blue Brain Project, École Polytechnique Fédérale de Lausanne, Geneva, Switzerland

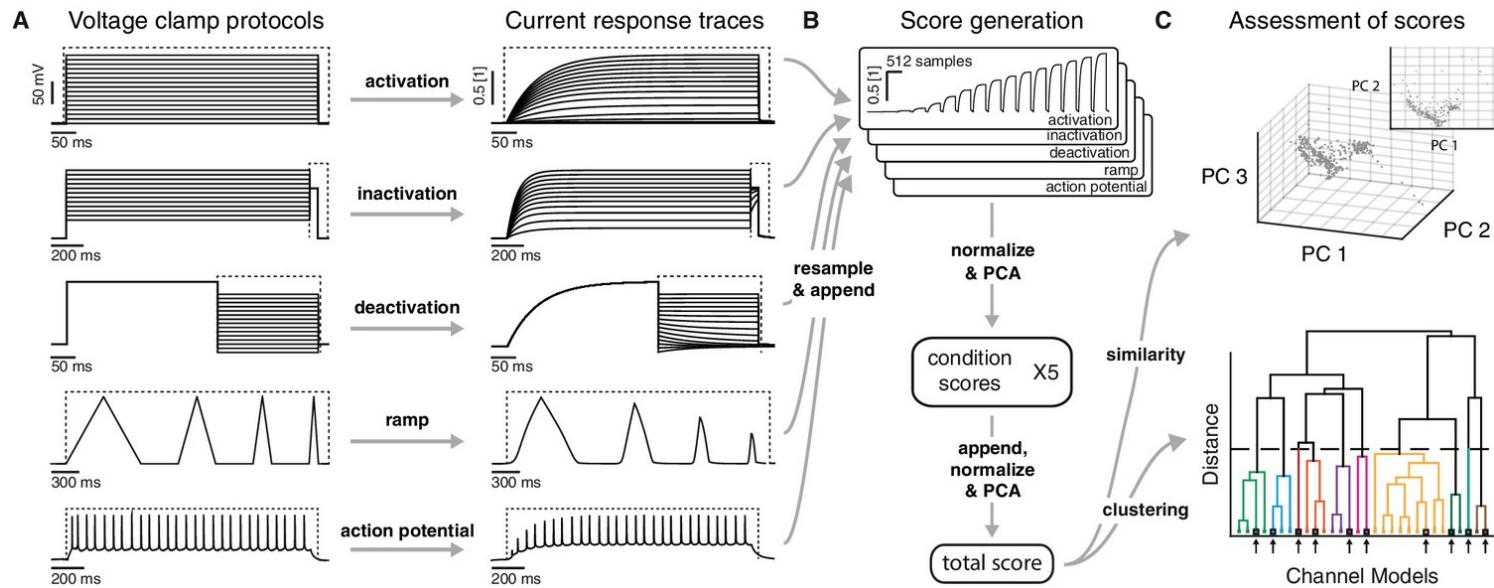
Abstract Ion channel models are the building blocks of computational neuron models. Their biological fidelity is therefore crucial for the interpretation of simulations. However, the number of published models, and the lack of standardization, make the comparison of ion channel models with one another and with experimental data difficult. Here, we present a framework for the automated large-scale classification of ion channel models. Using annotated metadata and responses to a set of voltage-clamp protocols, we assigned 2378 models of voltage- and calcium-gated ion channels coded in NEURON to 211 clusters. The *IonChannelGenealogy* (ICGenealogy) web interface provides an interactive resource for the categorization of new and existing models and experimental recordings. It enables quantitative comparisons of simulated and/or measured ion channel kinetics, and facilitates field-wide standardization of experimentally-constrained modeling.

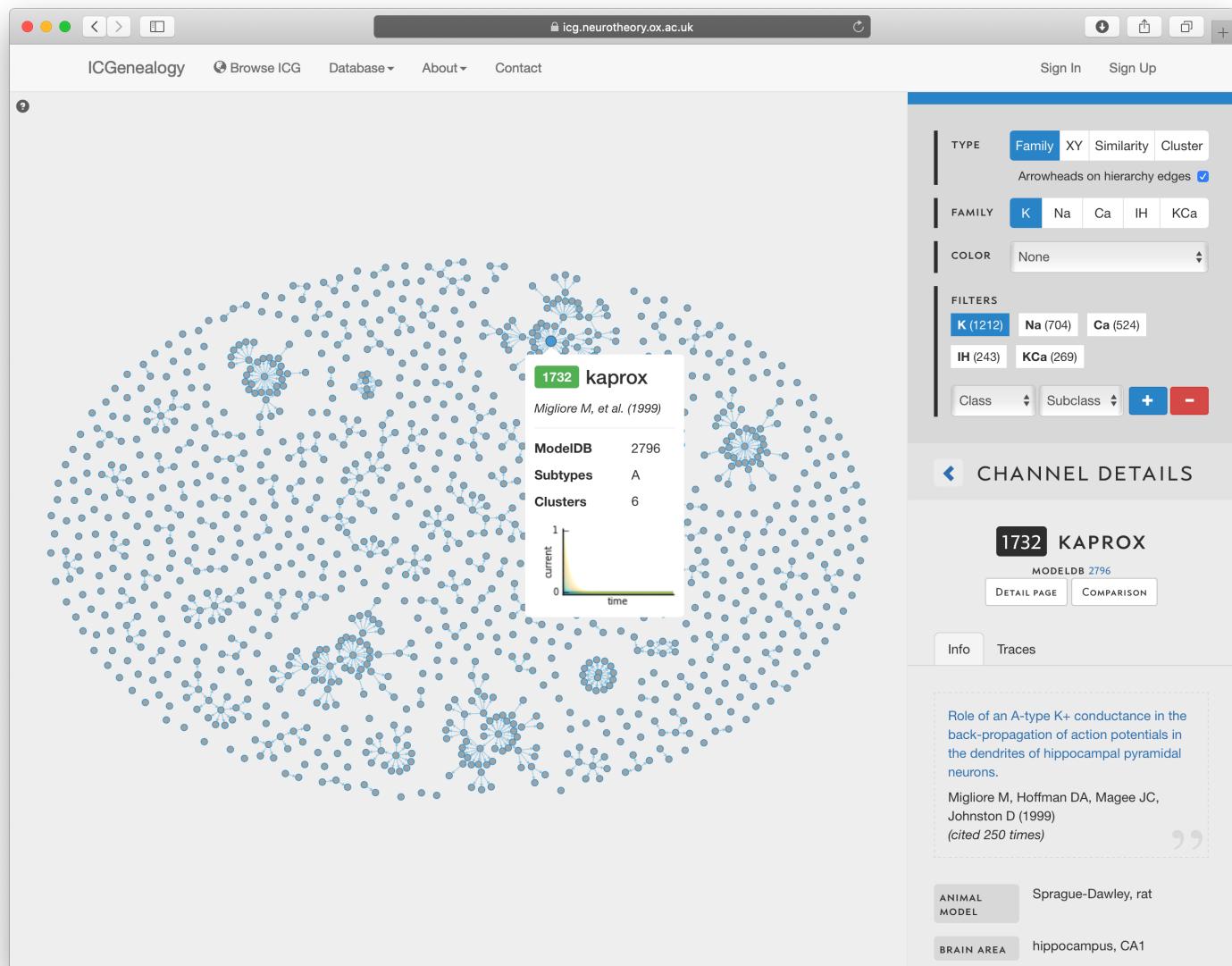
DOI: [10.7554/eLife.22152.001](https://doi.org/10.7554/eLife.22152.001)

*For correspondence: william.podlaski@cncb.ox.ac.uk

†These authors contributed equally to this work

Voltage-clamp protocols for the quantitative analysis of ion channel models





ICCGenealogy Browse ICG Database ▾ About ▾ Contact Sign In Sign Up

Channels / K / 2796-kaprox

Channel 2796-kaprox

[Open in channel browser](#) [Open in comparison](#) [Submit a correction](#)

General data

- ICG id: 1732
- ModelDB id: 2796
- Reference: Migliore M, Hoffman DA, Magee JC, Johnston D (1999): Role of an A-type K⁺ conductance in the back-propagation of action potentials in the dendrites of hippocampal pyramidal neurons.

Metadata classes

- Animal Model: Sprague-Dawley, rat
- Brain Area: hippocampus, CA1
- Neuron Region: soma, dendrites, proximal dendrites
- Neuron Type: pyramidal cell
- Runtime Q: Q3
- Subtype: A

Metadata generic

- Age: 5-8 weeks old
- Authors: M Migliore
- Comments: K-a channel from experiments/model in hoffman et al. (1997), referenced as klee flicker and heinemann (1995) in the mod file, modified to account for dax a current (hoffman et al.). Used in soma, basal dendrites, and apical dendrites less than 100 um from soma. 2796_kadist.mod has an activation curve shifted by -12mv with respect to 2796_kaprox.mod. A temperature of 35 deg c assumed for all simulations. No modeldb ancestors. This is a modeling study based on experimental work from the two previous virtual nodes (hoffman et al. 1997; klee, flicker and heinemann 1995). The animal model info listed here comes from these virtual nodes.
- Runtime: 7.976
- Temperature: Model has standard temperature dependence. The temperature was 35 deg C for all simulations.

Current Response Traces

Activation

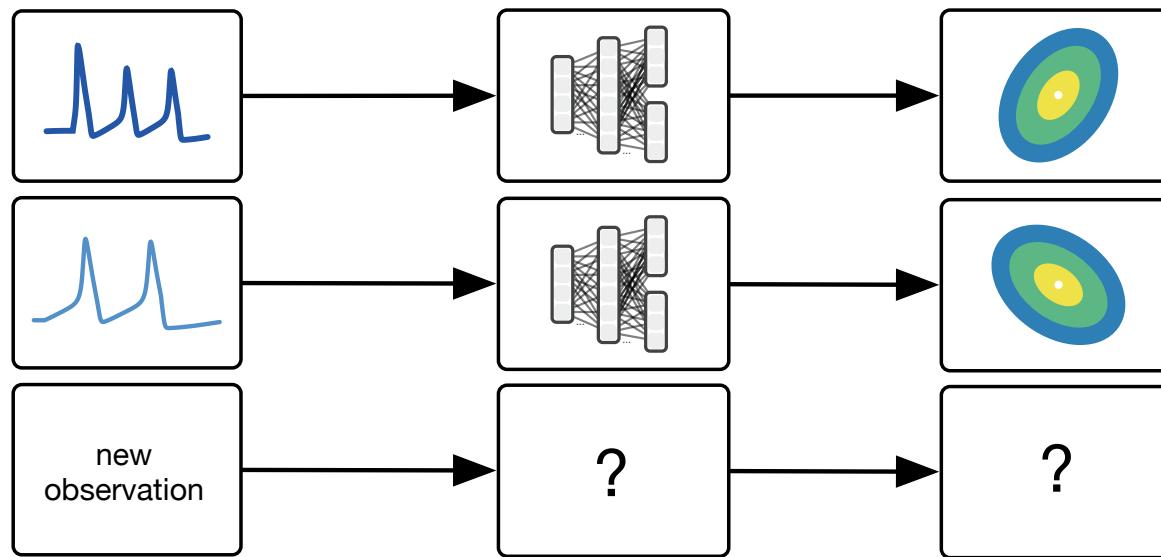
Inactivation

Deactivation

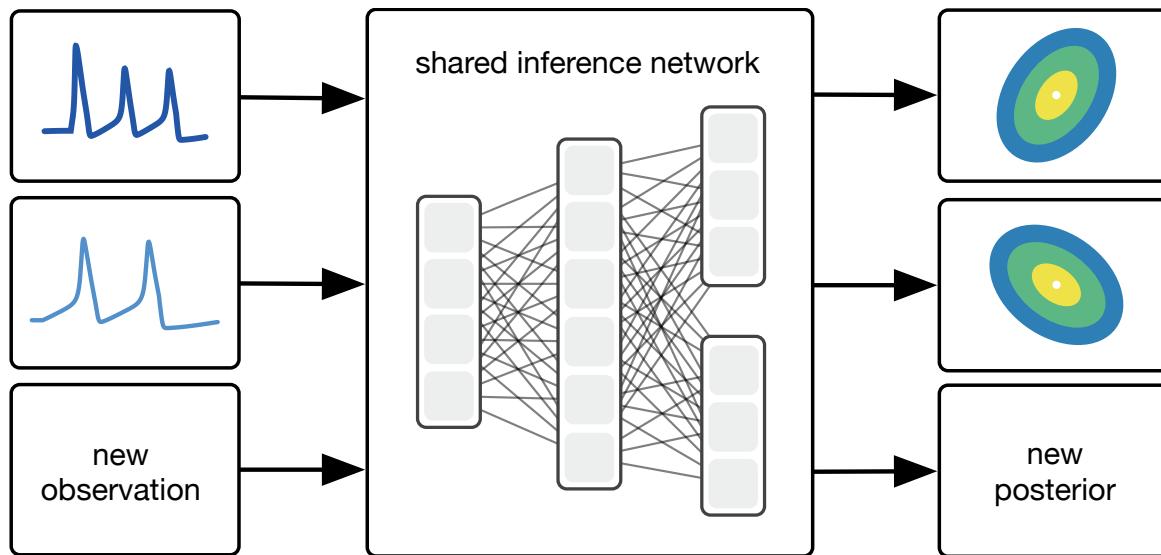
Action Potential

**Goal: Rapidly infer channel
parameters from current responses**

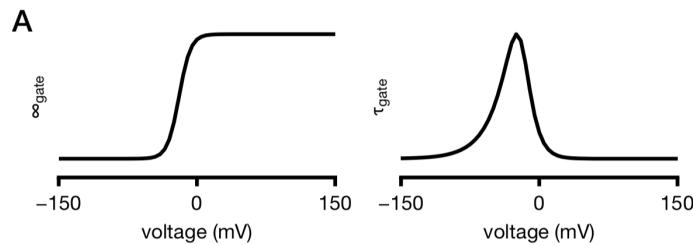
Standard inference scheme



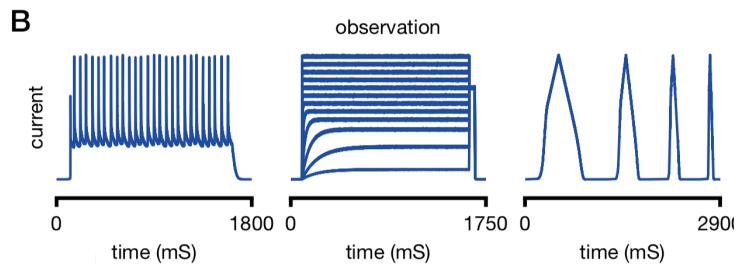
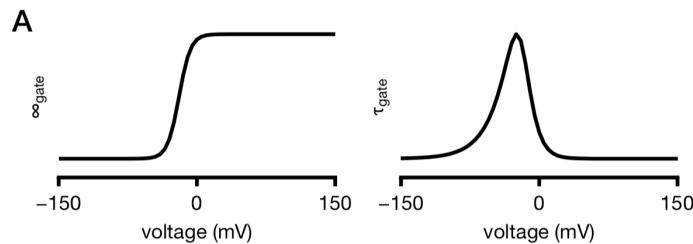
Amortized inference



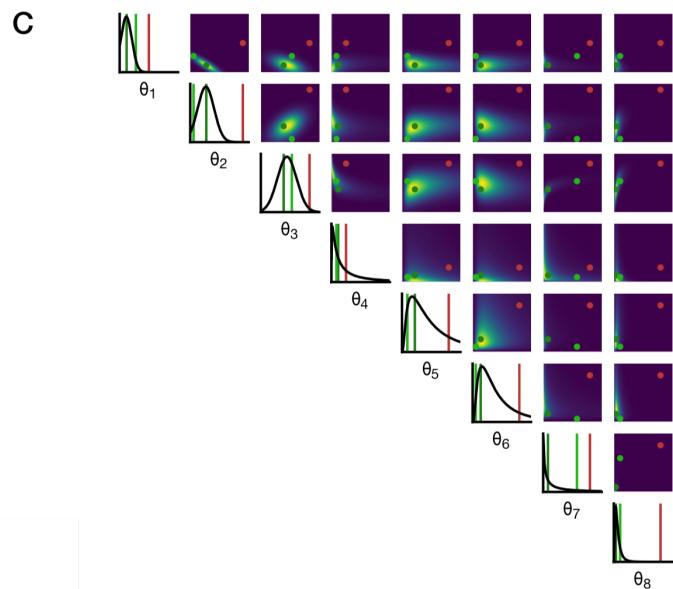
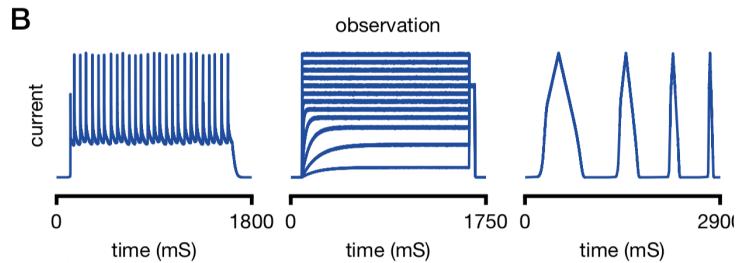
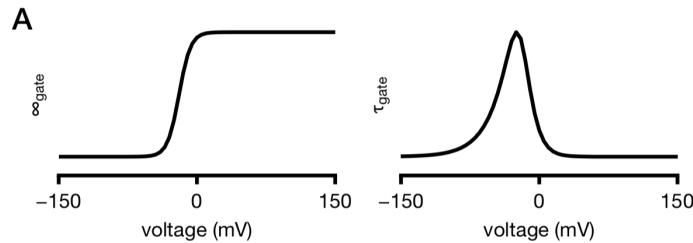
Amortized inference on ion channels



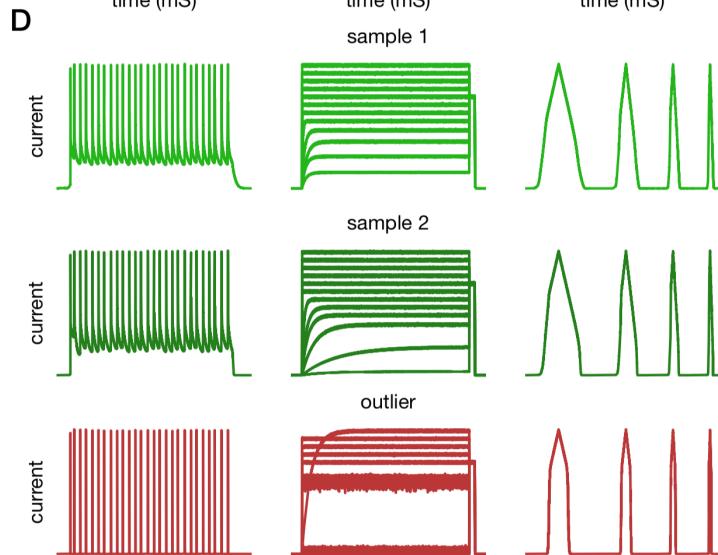
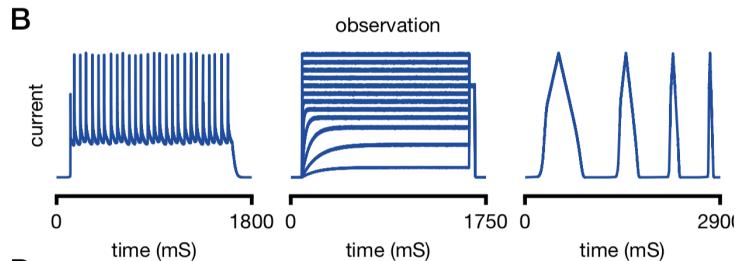
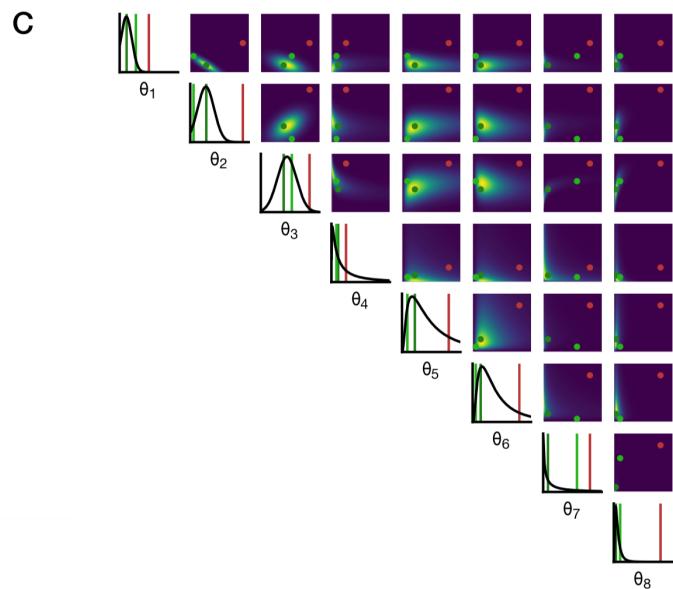
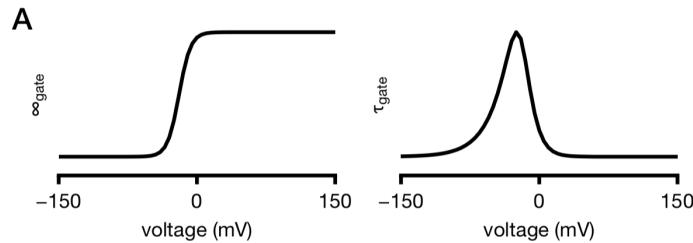
Amortized inference on ion channels



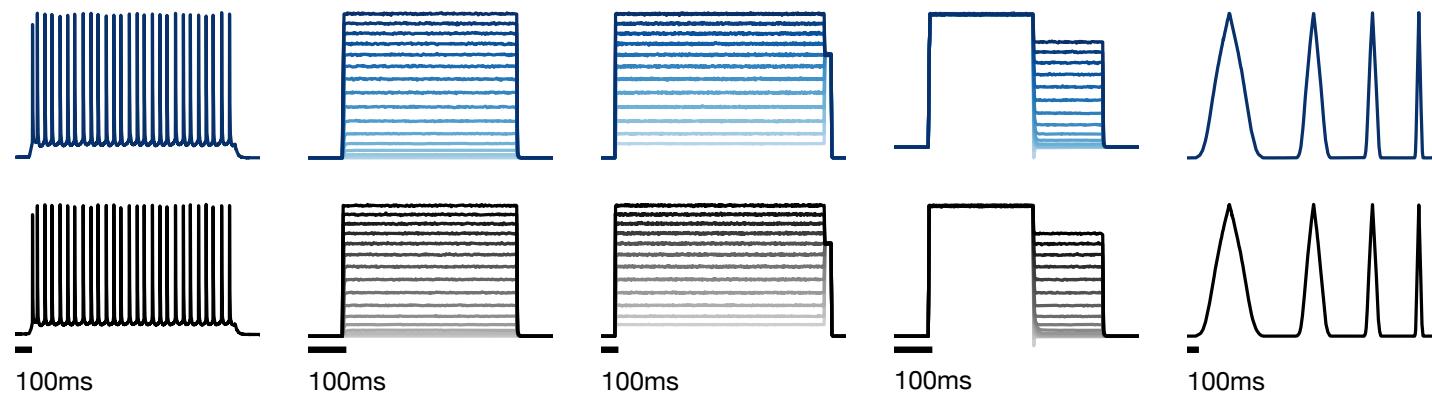
Amortized inference on ion channels



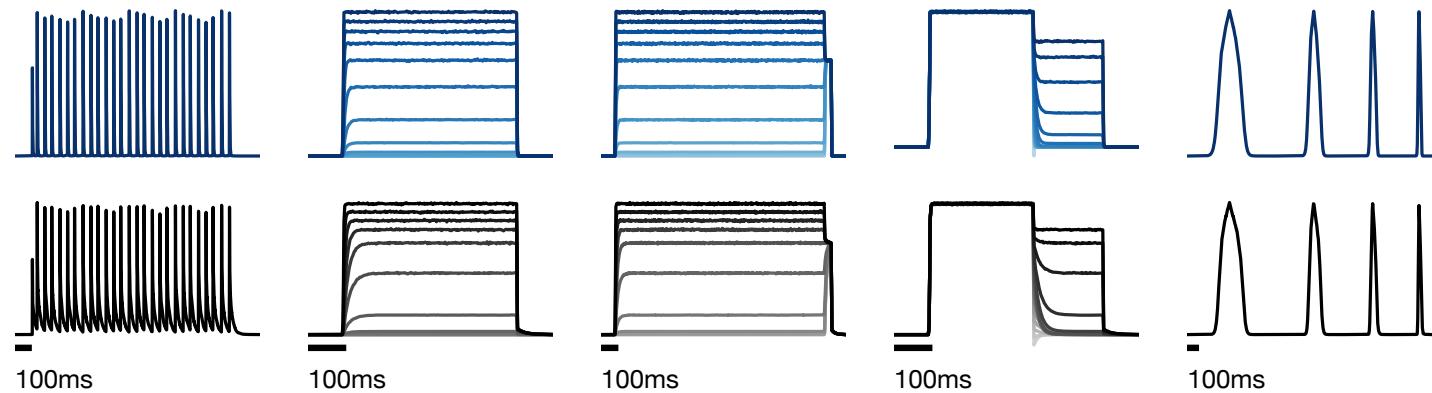
Amortized inference on ion channels



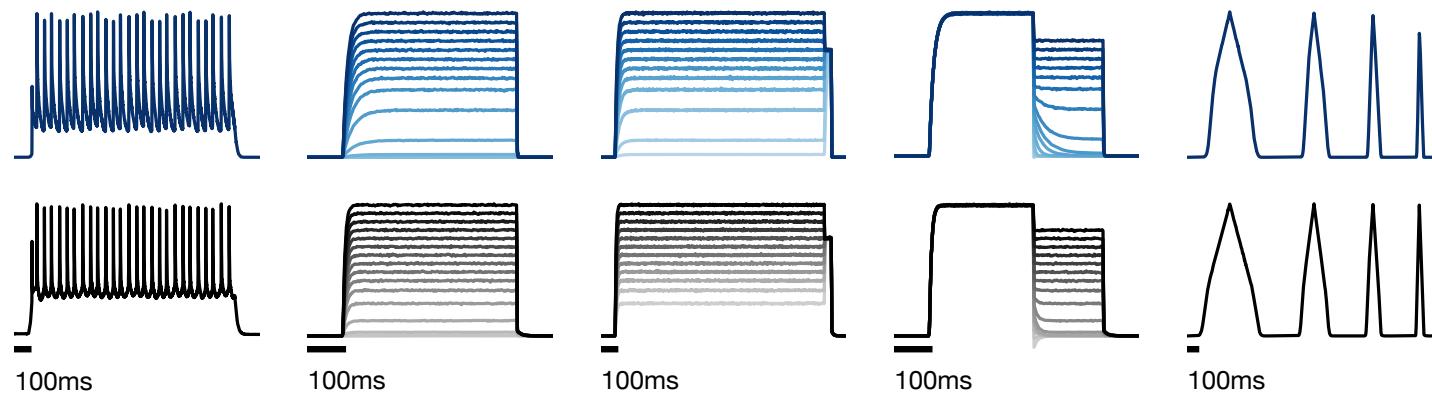
Channel 48332_kpkjslow · L₂-error 0.05



Channel 153280_ch_Kdrp · L₂-error 0.24

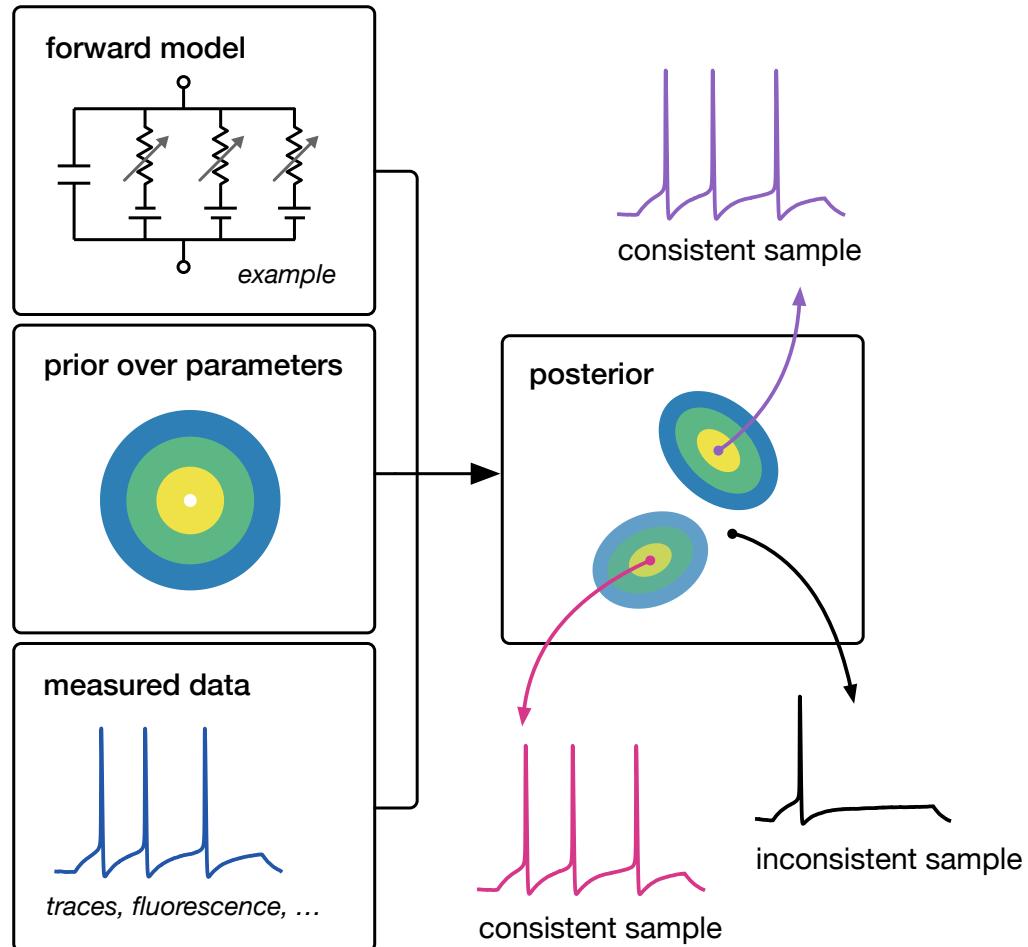


Channel 45539_km · L₂-error 0.49



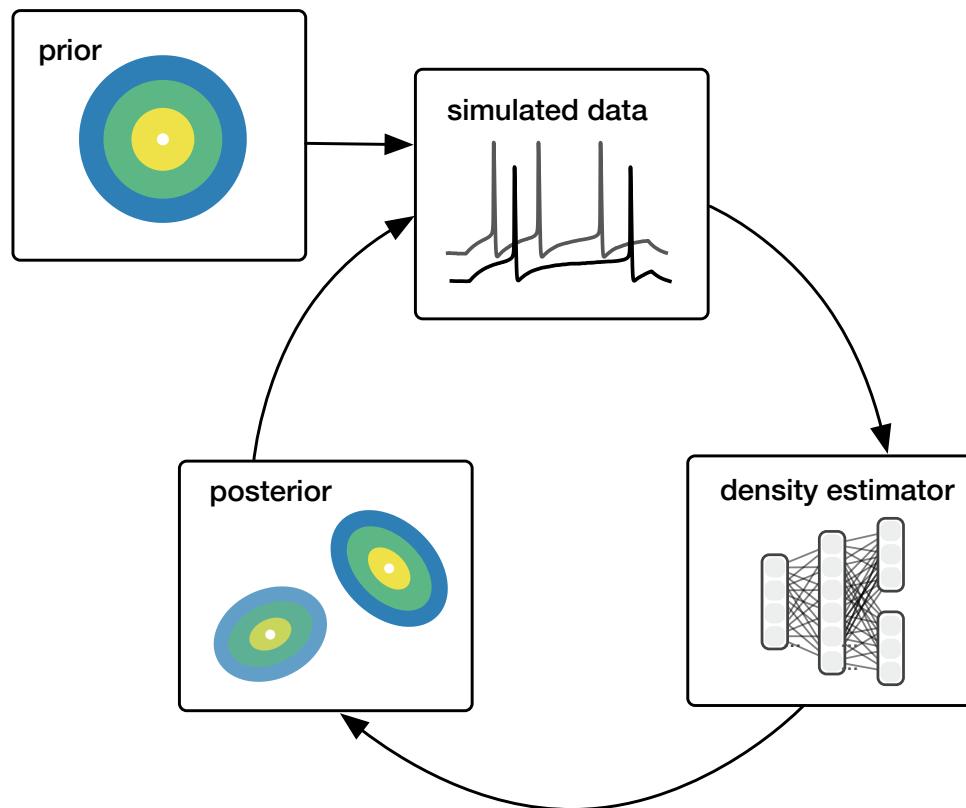
Summary

Goal: Bayesian inference for mechanistic models



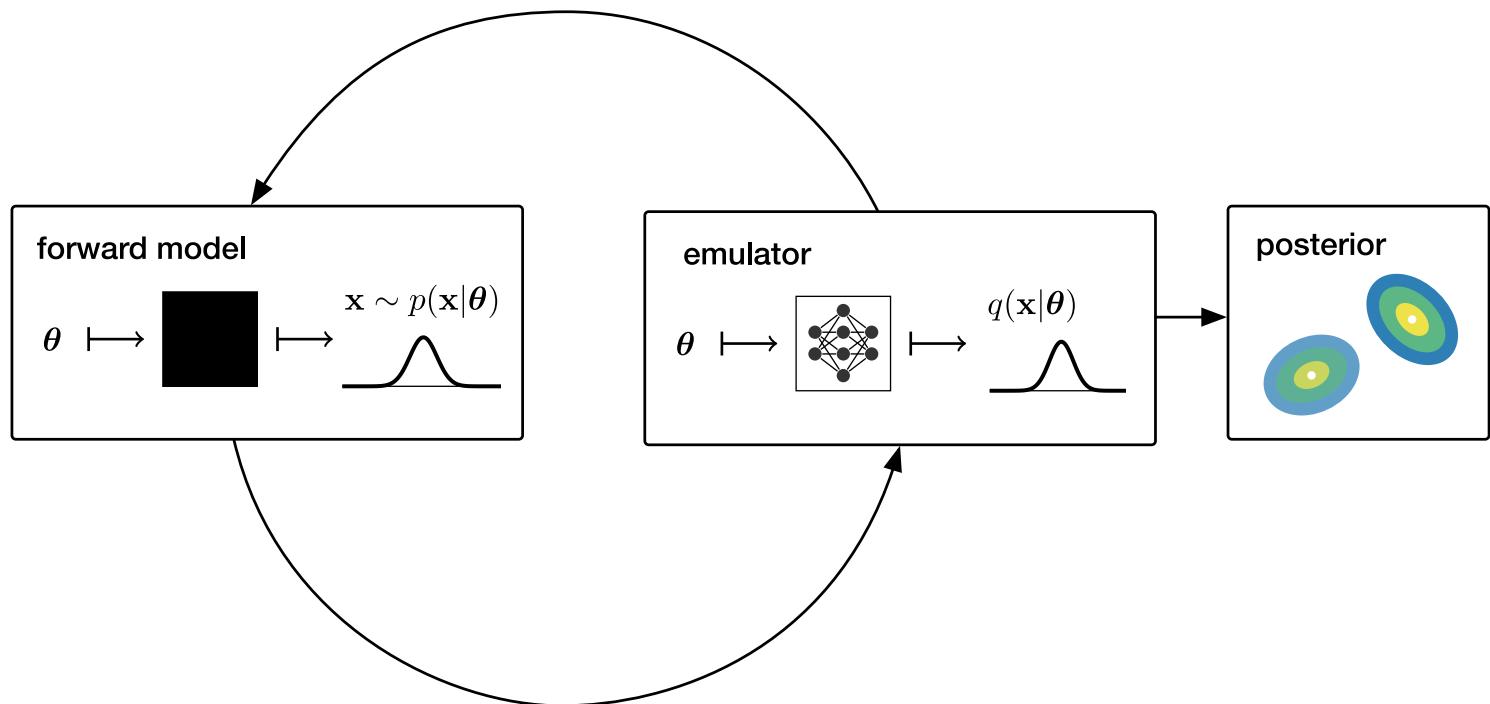
Approach 1: Sequential simulation-based inference

Learn posterior: $x \rightarrow \theta$



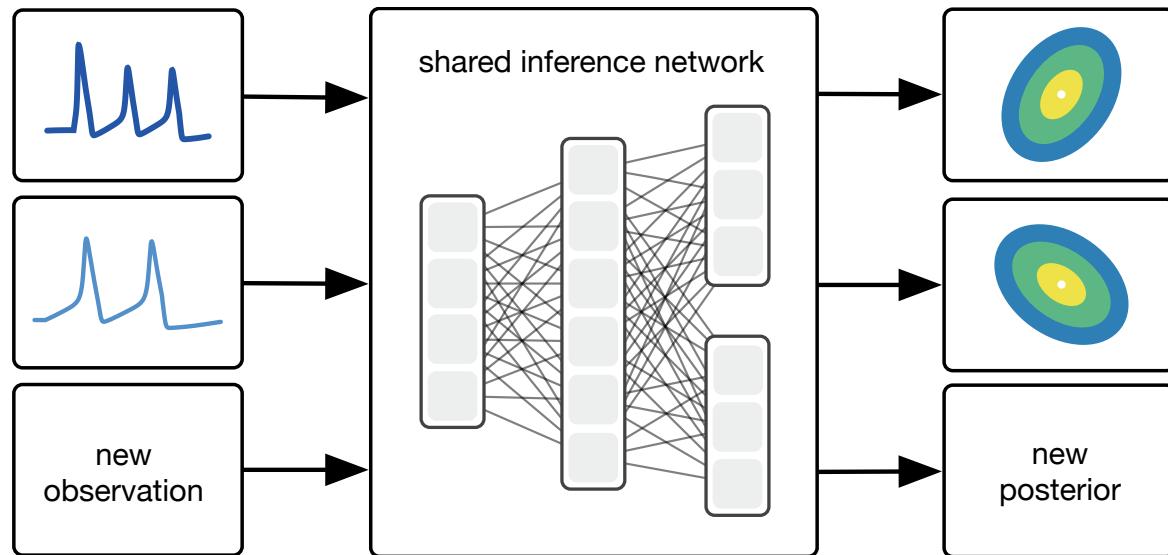
Approach 2: Sequential simulation-based inference

Learn likelihood: $\theta \rightarrow x$



Amortized inference

Applied to channelomics



Thank you!



Jakob Macke



Pedro Gonçalves



Giacomo Bassetto



Marcel Nonnenmacher

Thank you!



Jakob Macke



Pedro Gonçalves



Giacomo Bassetto



Marcel Nonnenmacher



Theofanis Karaletsos



Tim Vogels



William Podlaski



Chaitanya Chintaluri

Thank you!



Jakob Macke



Pedro Gonçalves



Giacomo Bassetto



Marcel Nonnenmacher



Theofanis Karaletsos



Tim Vogels



William Podlaski



Chaitanya Chintaluri

