Efficient Amortised Bayesian Inference for Hierarchical and Nonlinear Dynamical Systems

Geoffrey Roeder[1], Paul K. Grant, Andrew Phillips, Neil Dalchau[2], Edward Meeds[3]

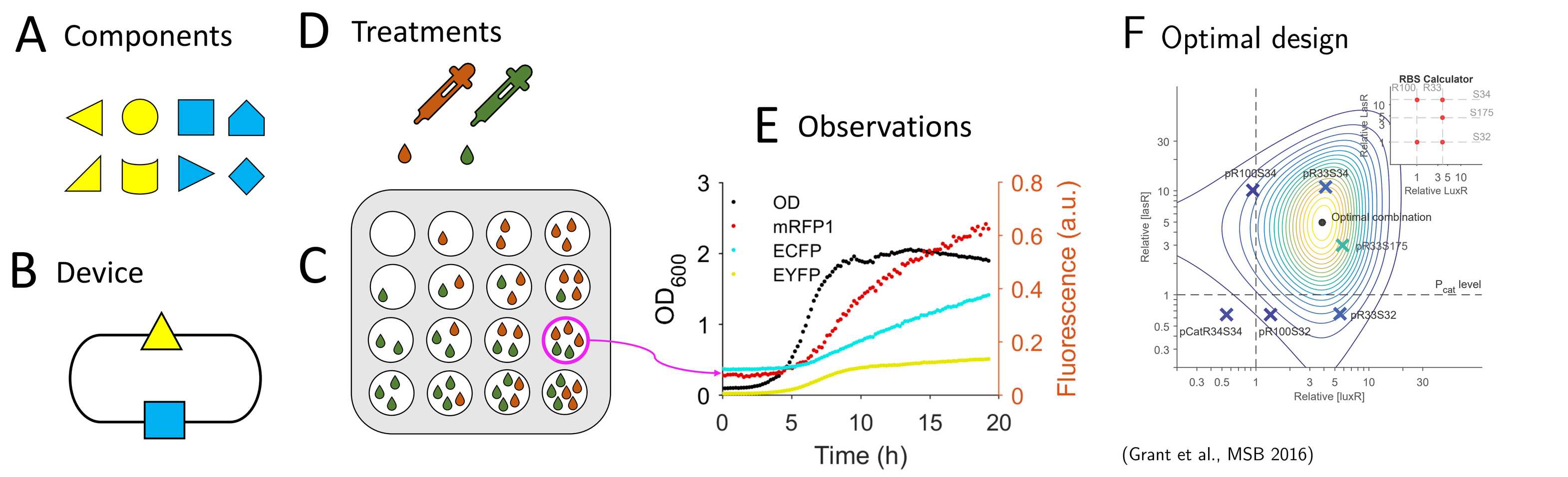
[1]roeder@princeton.edu, [2]ndalchau@microsoft.com, [3]ted.meeds@microsoft.com

Dynamical characterisation of mechanistic models

- Dynamical systems learned from experimental data are widespread in the physical sciences, including fluid dynamics, thermodynamics, and electromagnetism.
- They play a particularly important role in advancing our understanding of biology, typically studied as Ordinary Differential Equations (ODEs).
- The ability to precisely engineer biology could enable substantial breakthroughs in medicine and provide environmentally sustainable processes and products.
- We develop a novel model class made computationally tractable by recent advances in Bayesian Deep Learning

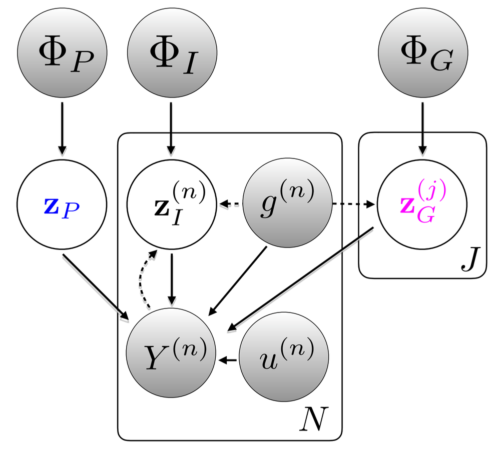
Case study: genetically engineering a biosensor

- We empirically validate our method by predicting the dynamic behaviour of bacteria that were genetically engineered to function as biosensors for two molecular input signals.
- Fluorescence measurements were collected to quantify the behaviour of a range of 2-input biosensors that differ in two of their constituent components (illustrated below as the yellow and blue symbols, panels A & B).
- The goal is to quantify, with uncertainty, the posterior distributions of the parameters of a mechanistic model that describes the interactions between the input signals, the internal components and the ability to produce (fluorescent) outputs.
- This enables the device to be optimised *in silico*, and guides the selection of better genetic components (panel F).



Modelling with nonlinear mixed-effects ODEs

- We propose a deep generative nonlinear mixed-effects (NLME) model, e.g, a generative model of a dynamical system that exhibits hierarchical latent structure. This enables us to combine individual-level (each time-series), group-level (each genotype) and global parameters.
- We cast parameter inference as stochastic optimisation of an end-to-end differentiable, block-conditional variational autoencoder.
- This model class is highly flexible: the ODE right-hand sides can be a mixture of user prescribed or white-box sub-components and neural network or black-box sub-components.



ODEs: interpretable white-box or flexible black-box

White box ODE model

$$\dot{[\text{RFP}]} = r_c - (d_{\text{RFP}} + \gamma).[\text{RFP}]$$
$$\dot{[\text{CFP}]} = a_{\text{CFP}}.r_c.f_{76}(C_6, C_{12}, [R], [S]) - (d_{\text{CFP}} + \gamma).[\text{CFP}]$$
$$\dot{[\text{YFP}]} = a_{\text{YFP}}.r_c.f_{81}(C_6, C_{12}, [R], [S]) - (d_{\text{YFP}} + \gamma).[\text{YFP}]$$
$$\dot{[R]} = a_R.r_c - (d_R + \gamma).[R]$$
$$\dot{[S]} = a_S.r_c - (d_S + \gamma).[S]$$
$$\dot{[F_{480}]} = a_{480}.r_c - \gamma.[F_{480}]$$
$$\dot{[F_{530}]} = a_{530}.r_c - \gamma.[F_{530}]$$

Black box ODE model

$$\dot{\mathbf{v}} = \omega_3^+(\mathbf{v}, \mathbf{x}, \mathbf{\Psi}) - \mathbf{v} \odot \omega_4^+(\mathbf{v}, \mathbf{x}, \mathbf{\Psi})$$

Notation:
- $[s]$ Concentration of s
- $\dot{s}$ ds/dt
- $x.s$ $x \times s$
- f^+ Positive function
- $\odot$ Hadamard product

Observer process

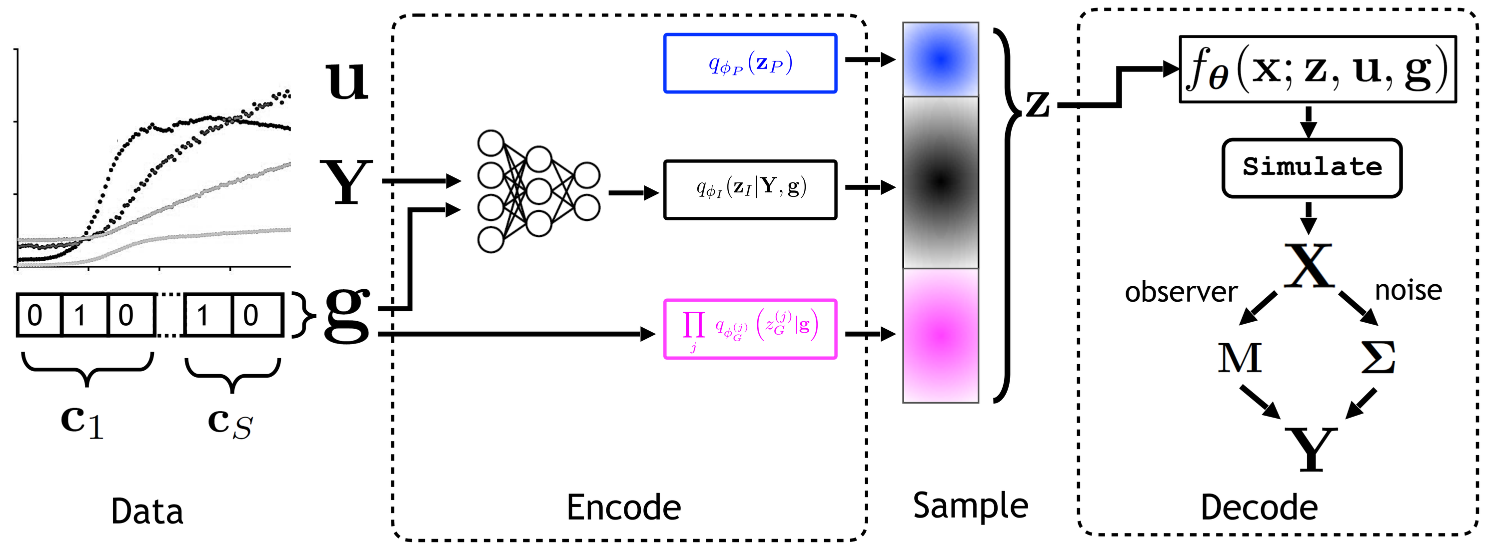
Signal	White-box	Black-box
OD	c	x_0
RFP	$c.[\text{RFP}]$	$x_0.x_1$
YFP	$c.([\text{YFP}] + [F_{530}])$	$x_0.x_2$
CFP	$c.([\text{CFP}] + [F_{480}])$	$x_0.x_3$

Generative process

$$\mathbf{z} \sim p_\theta(\mathbf{z}|\mathbf{g}) \quad (1)$$
$$\dot{\mathbf{x}} = f_\theta(\mathbf{x}; \mathbf{z}, \mathbf{u}, \mathbf{g}) \quad (2)$$
$$\mathbf{X} = \texttt{Simulate}(f_\theta, \mathbf{x}_0) \quad (3)$$
$$\mathbf{M} = \psi(\mathbf{X}), \quad \mathbf{\Sigma} = \rho(\mathbf{X}, \mathbf{z}) \quad (4)$$
$$\mathbf{Y} \sim p(\mathbf{Y}|\mathbf{M}, \mathbf{\Sigma}) \quad (5)$$

Conditional VAEs enable fast, scalable inference

- The computational flow graph for encoding, sampling from the variational posterior, and simulating the dynamical system. Note that the sample and simulate operations are constrained to be differentiable.



- Hence, the variational posterior is

$$\underbrace{q_{\phi_P}(\mathbf{z}_P)}_{\text{Population}} \underbrace{q_{\phi_I}(\mathbf{z}_I|\mathbf{Y}, \mathbf{g})}_{\text{Individual}} \underbrace{\prod_j q_{\phi_G^{(j)}}\left(\mathbf{z}_G^{(j)}|\mathbf{g}\right)}_{\text{Group}}$$

- Previous attempts at learning similar joint distributions have used MCMC
- Conditional VAEs are an order of magnitude faster, although MCMC will converge given enough time

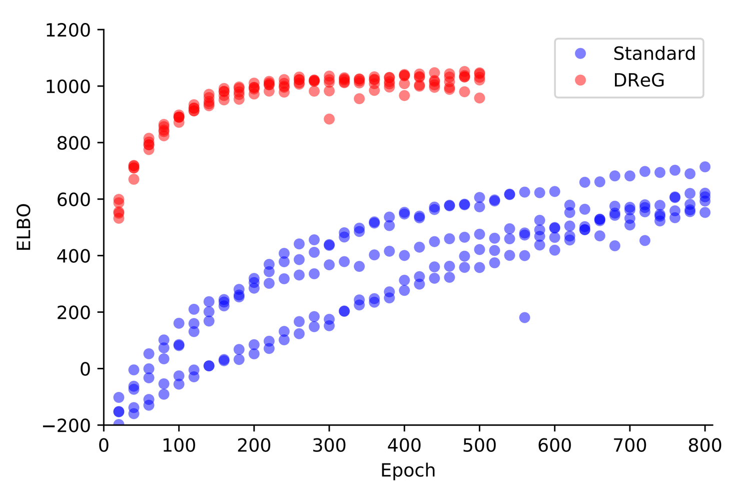


Figure 2: VI Convergence

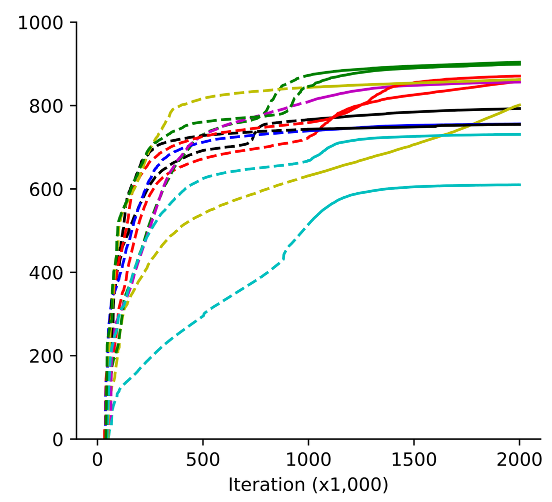
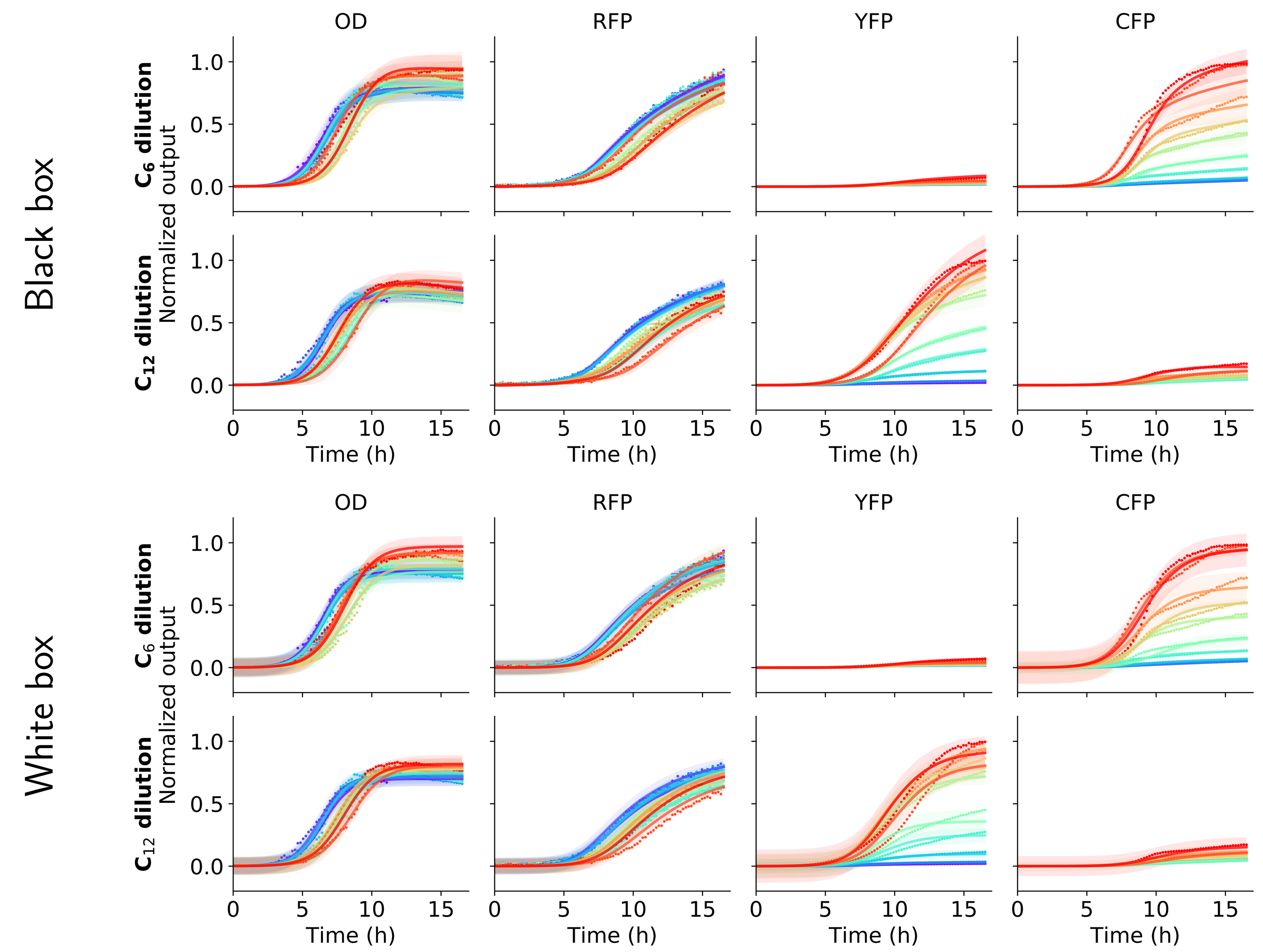


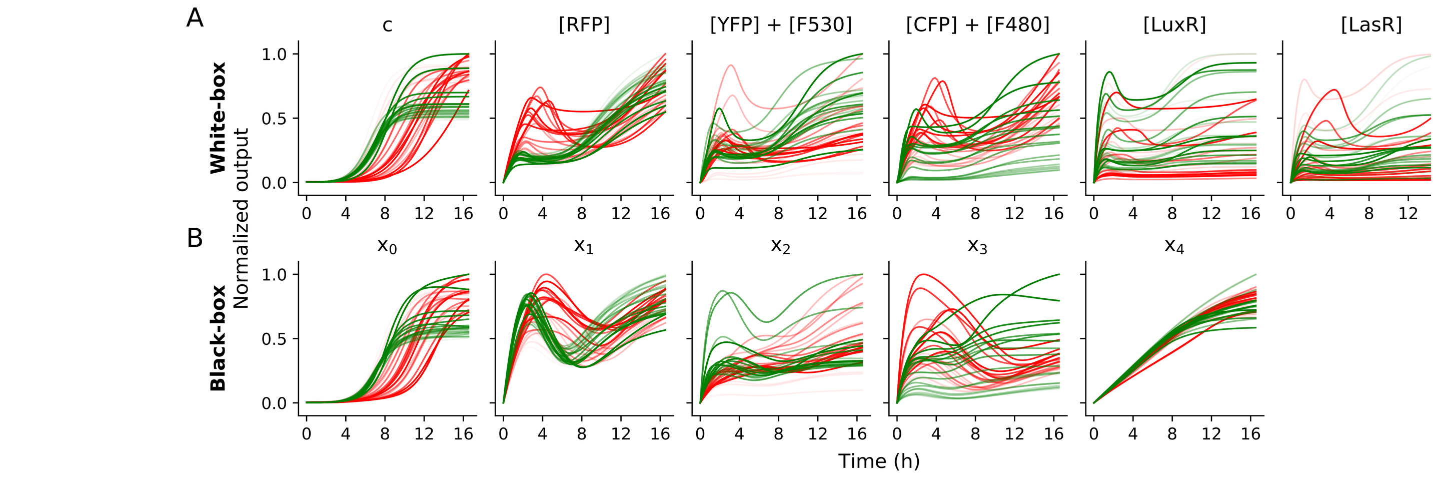
Figure 3: MCMC Convergence

Strong model fit evaluated by simulation

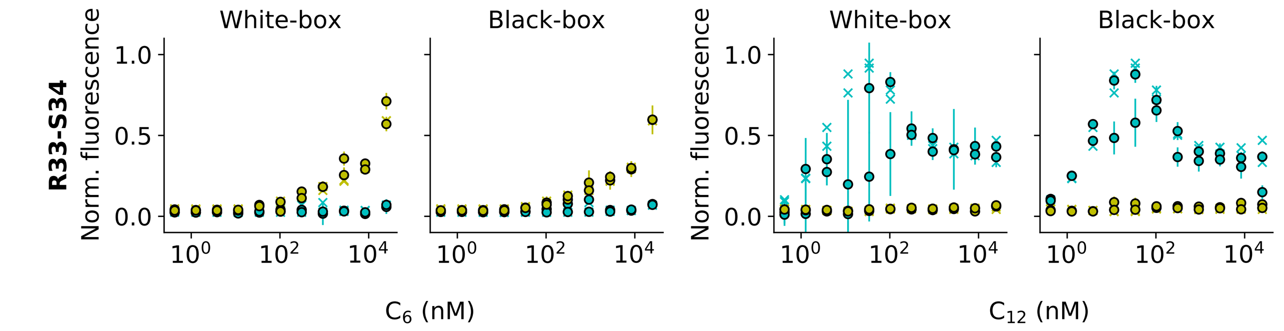


Strong performance on input-output summaries

Black and white box models learn similar dynamics



Zero-shot learning of unseen genotypes



Possible extensions

- "Grey-box" ODE models could use prescribed sub-models for aspects of the system that are well understood (qualitatively) and black-box sub-models for aspects less well understood.
- Extend to stochastic differential equations (replacing equation 2 in the generative process), which is an important model class in biology.
- Active learning, to provide experimenters with suggestions on how to improve models of the data, and potentially optimise against a design objective.