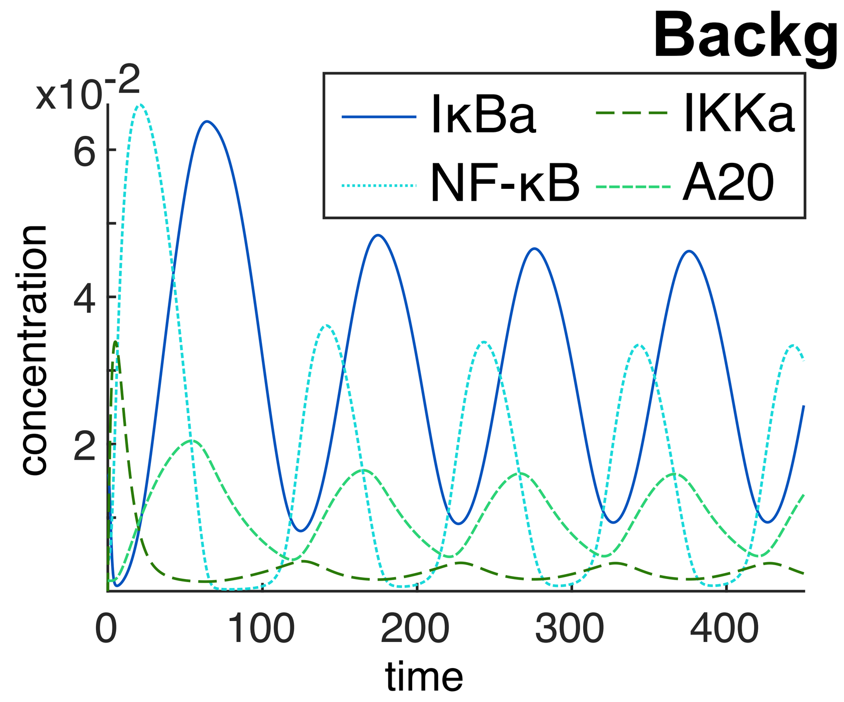
# Fast Bayesian inference for stochastic oscillatory systems using the phase-corrected Linear Noise Approximation

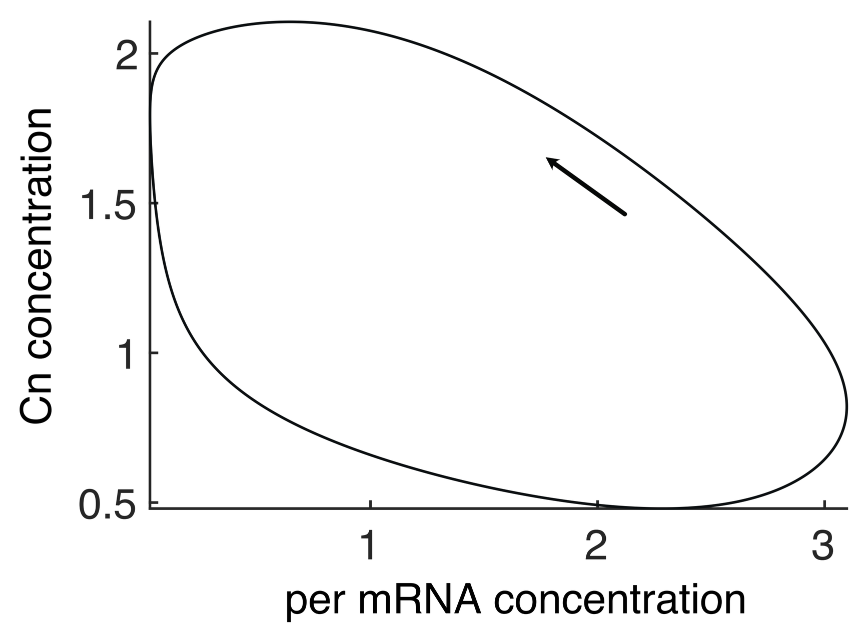
Ben Swallow[1*], David Rand[2] & Giorgos Minas[3]
[1]School of Mathematics and Statistics, University of Glasgow, UK;
[2]Mathematics Institute, University of Warwick, UK;
[3]School of Mathematics and Statistics, University of St Andrews, UK

**Abstract**

We propose a new methodology for inference in stochastic non-linear dynamical systems exhibiting oscillatory behaviour and show the parameters in these models can be realistically estimated from simulated data. We show that parameter sensitivity analysis can predict which parameters are practically identifiable. Several Markov chain Monte Carlo algorithms are compared. Our results suggest parallel tempering consistently gives the best approach for these systems, which are shown to frequently exhibit multi-modal posterior distributions.
Available at: arXiv:2205.05955

## Background



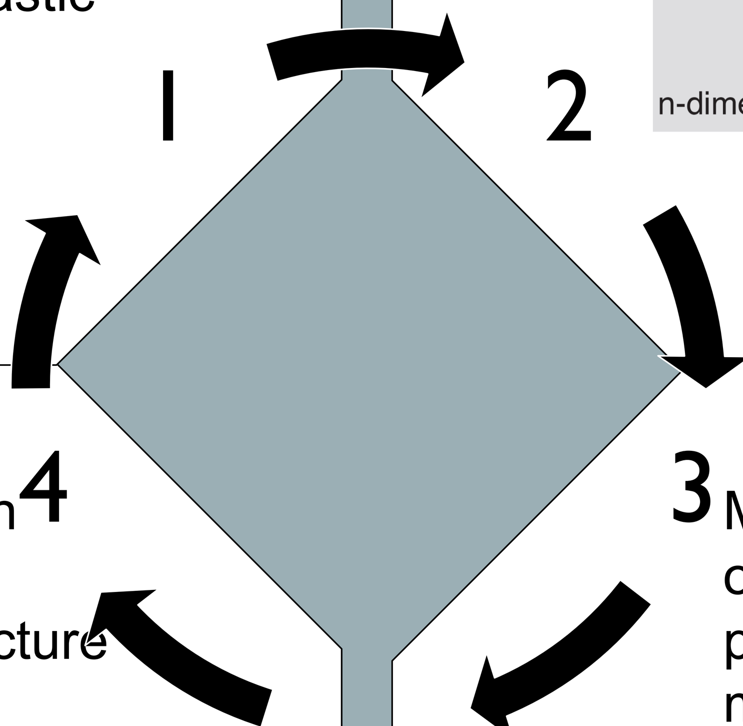

Oscillatory systems based on systems of ODEs with feedback loops are common in systems biology (e.g. NFkB and Drosophila clock above), epidemiology and ecology, amongst others domains.
The system of ODEs represents the deterministic behavior of the system but the biological counterparts are known to be stochastic. Therefore SDEs are required defined as below.

$$\mathbf{X}_t = \phi_t + \Omega^{-1/2}\xi_t$$

where $\mathbf{X}_t$ is the state vector, $\phi_t$ the deterministic solution to the ODEs, $\Omega$ the system size and $\xi_t$ the stochastic component. The stochastic component is the solution to the SDE

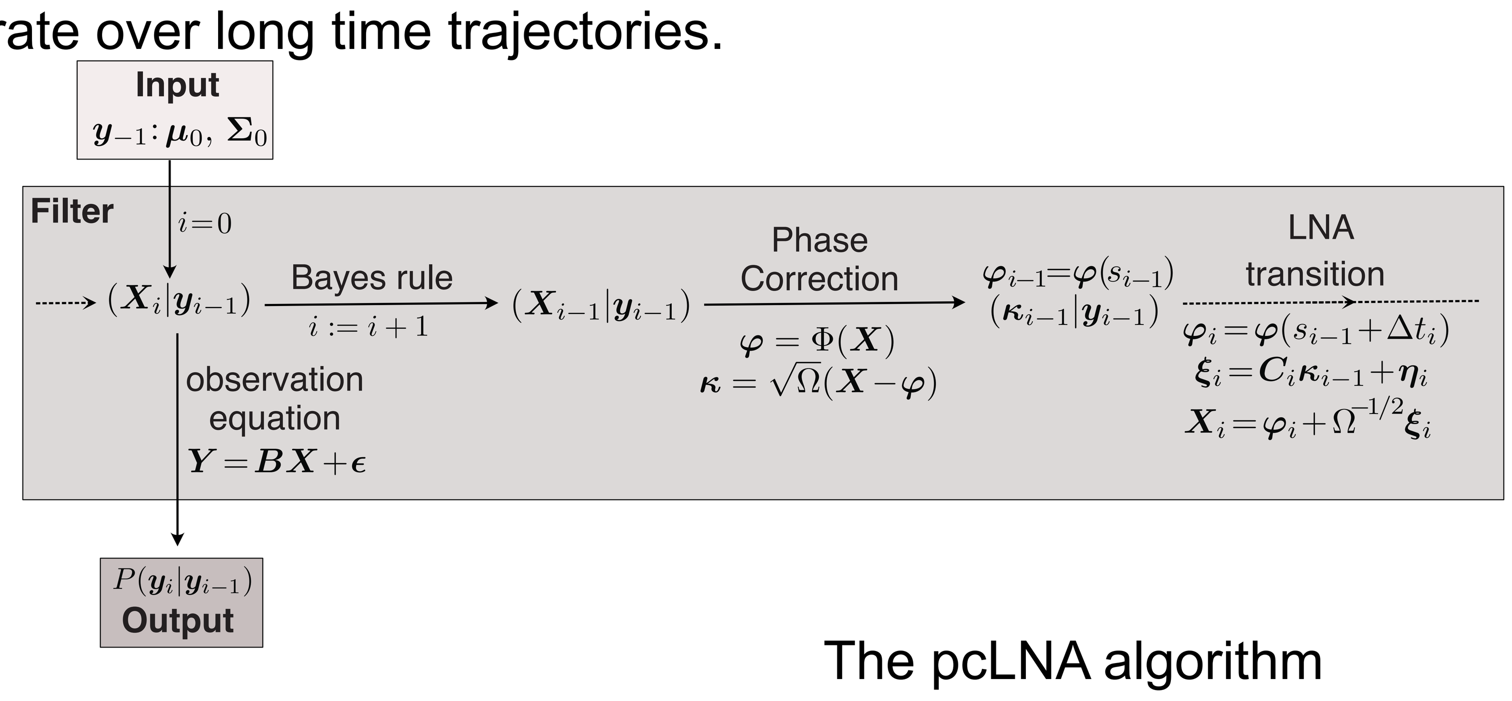
$$d\xi_t = \mathbf{J}\xi_t dt + \mathbf{S}d\mathbf{W}_t$$

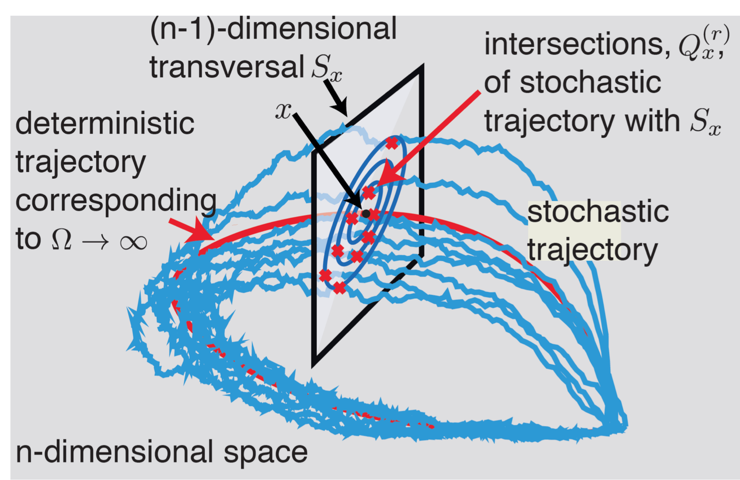
where **J** is the Jacobian, **S** accounts for stoichiometry and the reaction rates and $\mathbf{W}_t$ is a classical Weiner process.




## Simulation

Simulation of the SDE model is conducted according to the underlying Markov jump process theory. The linear noise approximation is an Euler approximation to the jump process, however the approximation becomes inaccurate over long time trajectories.



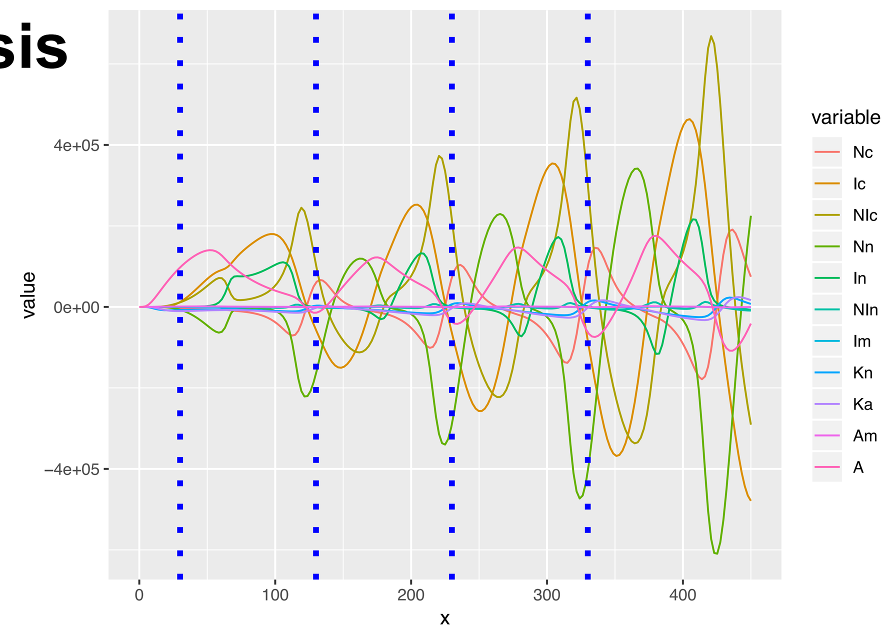



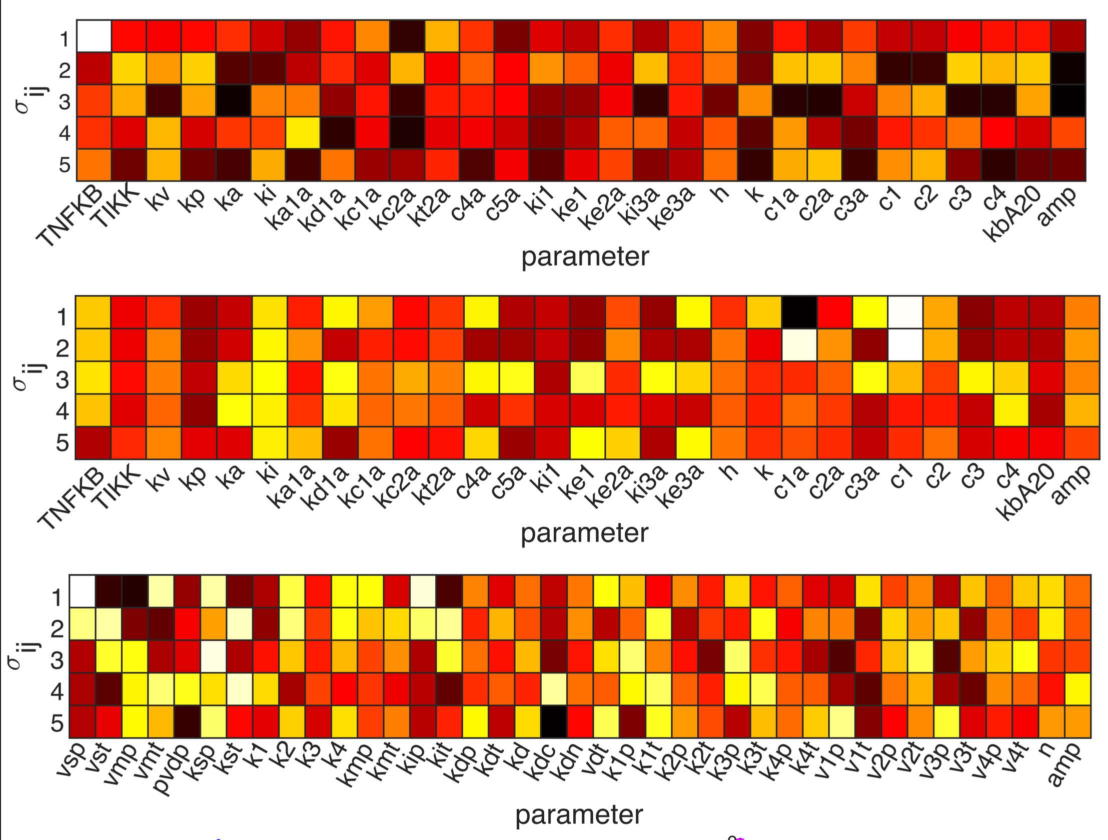

From Minas & Rand (2017)

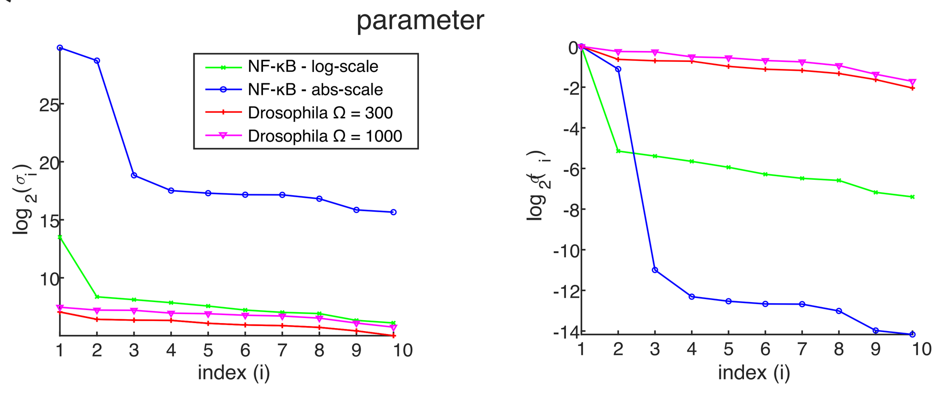
The pcLNA algorithm corrects for the deviations in phase through time, remaining accurate for much longer time periods.
It is analytically tractable requiring only the mean and covariance functions to be updated according to a variant of the LNA Kalman filter (see above).

## Sensitivity analysis

Many variables (order of 10-20) and parameters (30+) makes statistical inference challenging. Sensitivity analysis helps highlight important parameters and timepoints.



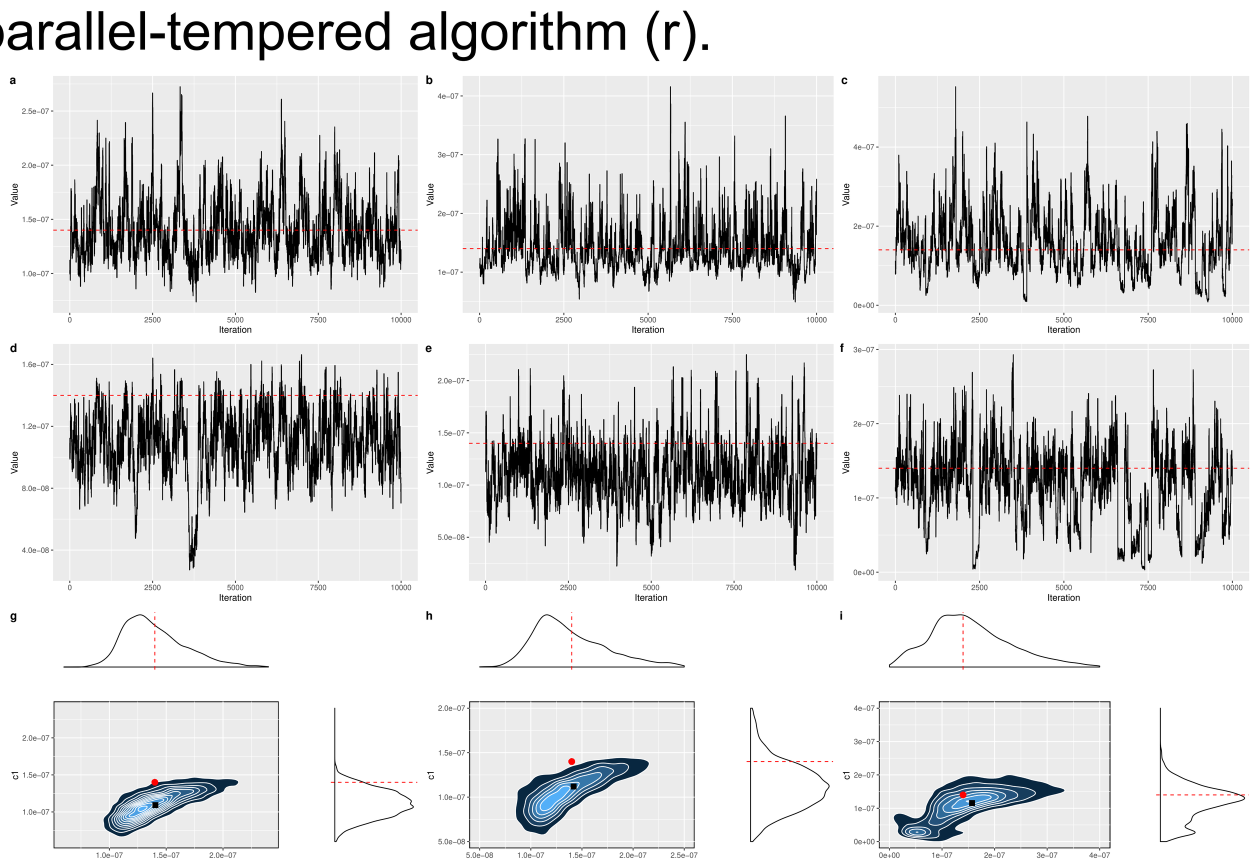



Different systems have different levels of sensitivity to different numbers of parameters and scales.
Using Fisher information, we can determine the numbers of parameters that are likely to be estimable from data. Parameters that do not induce changes in the likelihood function are likely non-identifiable.

## Inference

We use a variety of MCMC algorithms to conduct Bayesian inference. This includes standard random walk Metropolis Hastings (l), a manifold MALA (c) that uses geometric structure and a parallel-tempered algorithm (r).



Combined with the pcLNA algorithm, this provides ability to estimate parameters from data. The parallel tempering algorithm enables improved estimation when there are multimodal distributions, common to models with bifurcations.
In the NFkB system, it highlights a second region of parameter space that is consistent with the discretely observed data. Specifically this combination exhibits no oscillations but matches the peaks of the original oscillations.
We were able to successfully estimate 9 parameters in the Drosophila clock, even with diffuse priors.

University of Glasgow, charity number SC004401

**Funding generously provided by:** Biotechnology and Biological Sciences Research Council; Edinburgh Mathematical Society