THE UNIVERSITY OF CHICAGO

QUANTITATIVE ANALYSIS OF DYNAMICAL SYSTEMS MODELS OF BIOLOGICAL SYSTEMS

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF CHEMISTRY

BY
CHRIS CHI

CHICAGO, ILLINOIS
DECEMBER 2026

© 2026 by Chris Chi

ORCID iD: https://orcid.org/0000-0001-9382-0910

TABLE OF CONTENTS

LIST OF FIGURES v

LIST OF TABLES ix

ACKNOWLEDGMENTS xi

ABSTRACT xii

1 INTRODUCTION 1

2 SAMPLING PARAMETERS OF ORDINARY DIFFERENTIAL EQUATIONS WITH CONSTRAINED LANGEVIN DYNAMICS 5
- 2.1 Abstract 5
- 2.2 Introduction 5
- 2.3 Motivating example 8
- 2.4 Theory 11
  - 2.4.1 Numerical continuation in a single parameter 12
  - 2.4.2 Constrained Hamiltonian dynamics 14
  - 2.4.3 Hamiltonian Monte Carlo methods 15
- 2.5 Estimating model parameters for the repressilator 17
  - 2.5.1 Sampling fixed points 18
  - 2.5.2 Sampling Hopf points 19
  - 2.5.3 Sampling limit cycles 21
  - 2.5.4 Comparison with unconstrained MCMC methods 23
- 2.6 Conclusions 27
- 2.7 Methods 29
  - 2.7.1 Numerical integration of Hamilton's equations of motion with constraints 29
  - 2.7.2 Numerical integration of constrained Langevin equations 31
  - 2.7.3 Solving the constraint equations 34
  - 2.7.4 Discretizing periodic orbits 38
  - 2.7.5 Unconstrained Monte Carlo methods 40
  - 2.7.6 Accounting for manifold curvature 43
  - 2.7.7 Assessing convergence 44
  - 2.7.8 Importance sampling 45

3 T432 STABILIZES THE PERIOD OF THE KAI OSCILLATOR BY ENABLING DOUBLY PHOSPHORYLATED KAIC TO ACT AS AN ALTERNATIVE SINK FOR KAIA 51
- 3.1 Abstract 51
- 3.2 Introduction 52
- 3.3 Results 54

3.3.1 A monomeric model of the Kai oscillator fits data over a range of conditions. . . . 54
3.3.2 Model reduction reveals that T432 is dispensable for oscillations but required for robustness to variations in KaiA concentration. . . . 61
3.3.3 A minimal model recasts the oscillator as a competition between KaiA binding at CI and at CII. . . . 71
3.3.4 The phosphorylation and dephosphorylation segments are determined by the relative rates of KaiA sequestration at CI and release at CII. . . . 76
3.3.5 The duration of the phosphorylation segment is sensitive to KaiA and nucleotide concentrations. . . . 78
3.3.6 Doubly-phosphorylated KaiC stabilizes the period by acting as a sink for KaiA. . . . 82
3.3.7 Alternative models . . . 85
3.4 Discussion . . . 86
3.4.1 Relation to previous models of the Kai oscillator . . . 86
3.4.2 Mixed positive/negative feedback and evolvability . . . 88
3.4.3 Broader connections to models of chemical oscillators . . . 89
3.5 Methods . . . 90
3.5.1 Choice of priors . . . 91
3.5.2 Sampling fixed points . . . 93
3.5.3 Sampling Hopf points . . . 94
3.5.4 Surrogate sampling for oscillations that are robust to variations in KaiA and ATP concentrations . . . 96
3.5.5 Sampling limit cycles . . . 97
3.5.6 Experimental methods . . . 100
3.6 Appendix . . . 101
3.6.1 Asymptotic expansions of the single-site model . . . 101
3.6.2 Asymptotic duration of the phosphorylation segment . . . 124
3.6.3 Asymptotic duration of the dephosphorylation segment . . . 126

4 TRANSITION PATH THEORY FOR EXCITABLE DYNAMICS . . . 129
4.1 Abstract . . . 129
4.2 Introduction . . . 129
4.2.1 Expected maximum deviation . . . 132
4.2.2 Expected excursion count . . . 135
4.3 Numerical examples . . . 138
4.4 Discussion . . . 142

5 CONCLUSIONS AND OUTLOOK . . . 143

REFERENCES . . . 146

LIST OF FIGURES

2.1 Schematic of the repressilator model with $2\ell + 1$ species, where ℓ is a positive integer. 9

2.2 Simulated data used in parameter estimation for the three- and seven-species repressilator. We integrate the three-species equations with arbitrarily chosen oscillatory parameters forward in time using Tsitouras's 5/4 explicit Runge-Kutta method [128] to generate the ground truth data shown red. Simulated observations, shown in blue, are generated by adding normally distributed noise with mean zero and variance 2.5×10^{-3} to the concentration of the first species. . . . 10

2.3 Behavior of the derivative vector $(\ \dot{\mathbf{y}}\ \ \dot{\mathbf{k}}\ )$ when approaching a turning point at $(\mathbf{0}, \mathbf{0})$. (left) In the case of natural parameter continuation, the derivative vector of (2.7) increases in norm when $(\mathbf{y}, \mathbf{k})$ approaches $(\mathbf{0}, \mathbf{0})$ due to the constraint $\dot{\mathbf{k}}(t) = 1$ and is undefined at $(\mathbf{0}, \mathbf{0})$. (right) In pseudoarclength continuation (PALC), the derivative vector is constrained to have unit norm in the system of equations in (2.8) and is well defined even at $(\mathbf{y}, \mathbf{k}) = (\mathbf{0}, \mathbf{0})$. 14

2.4 Representative two-dimensional marginals for the three-species represillator model parameter distribution obtained from 10^7 steps of the constrained Langevin sampler with fixed-point constraints (top), Hopf-point constraints (middle), and periodicity constraints (bottom). The log-likelihood when sampling fixed points and Hopf points is constant everywhere, and the log-likelihood when sampling limit cycles is proportional the squared error from the data shown in Figure 2.2 (see Section 2.5.4). 47

2.5 100 best fit solutions obtained after 10^7 steps of the constrained MALA sampler (left) and 2×10^8 steps the ensemble MCMC sampler (right). Results shown are for the three-species repressilator fit to the data shown in Figure 2.2 (see Section 2.5.4). 48

2.6 Representative marginal distributions sampled using constrained MALA, ensemble MCMC, unconstrained MALA, and NUTS after 10^7 steps. Results shown are for the three-species repressilator fit to the data shown in Figure 2.2 (see Section 2.5.4). 49

2.7 The distribution sampled by the unconstrained methods is multimodal despite the marginals along certain dimensions appearing smooth (top left). The multimodal structure is apparent when examining the distribution at different initial values of y_1 and y_2: the unnormalized log-posterior distribution exhibits rapid, non-concave variations with the rest of the parameters fixed at the high probability values indicated by the symbols (q_1, orange; q_2, blue; the linear interpolation $0.8q_1 + 0.2q_2$; pink in the bottom row). The interpolated parameters generate trajectories that are close in period and shape to the trajectories generated by q_1 and q_2 but about half a period out of phase with the data used to compute the log-likelihood (top middle). The apparent noise in the bottom middle plot results from failure to integrate the ODEs for certain initial conditions. Most of the normalized importance weights for reweighting the distribution sampled by constrained Langevin to the distribution sampled by the unconstrained methods are close to one (top right). 50

3.1 A diagram of the states of KaiC and their reactions in our monomeric model of the Kai oscillator. 57

3.2 Experimentally measured phosphorylation dynamics of the Kai oscillator (symbols) and 100 optimized model outputs (thin lines) for different conditions. For the oscillatory conditions shown with total KaiA and ATP fractions of $(0.6, 1.0)$ and $(0.6, 0.25)$, we process the time-dependent phosphorylation data reported in [111] by dividing the experimental data into segments of length equal to the experimentally determined period and averaging over all segments. 60

3.3 Model behavior in response to time-varying conditions, compared to experiment. (Upper left) Entrainment of the model when the solution ATP fraction is alternated between equal intervals of 25% and 100% at driving period τ_d, for τ_d between $0.9\,\tau$ and $1.1\,\tau$, where τ is the intrinsic period; dashed lines indicate the limiting cycle duration at each driving period. (Upper right) Timing of the peak in phosphorylation relative to the ATP pulse for different "day lengths" (fractions of the period at 100% ATP). Different lines show different driving periods, as indicated. (Lower left) Model and experimental [74, 116] phase response curves (PRCs) for a step down in solution ATP from 100% to 25% lasting $0.2\,\tau$. (Lower right) Model and experimental PRCs for a 16% step up in total KaiA concentration. 62

3.4 Simulations of the phase mixing experiment performed in [71]. Each point in the six lower panels represents a state of the system obtained by averaging the state of the system at two points on the curve in the upper panel. Simulations are initiated from each of these mixed conditions, and the time at which the peak in the phosphorylation fraction occurs is recorded. The lower panels show the difference between the peak time between the simulations started from the unmixed conditions, which are plotted with a black marker along the dotted diagonal line, and the mixed conditions. 63

3.5 Marginal posterior densities for the rate constants of the monomeric model of the Kai oscillator, obtained by sampling subject to the Hopf point constraints and likelihood function described in Section 3.5.4. Horizontal axes are logarithmic, and each density is normalized. 66
3.6 Reactions and states that remain after deleting reactions from the full model while requiring that certain properties of the model remain intact. If we only require that the system oscillates at 0.6 μM total KaiA dimer and 100% solution ATP, only the states connected by the reactions represented by solid lines are necessary (the S model). If we require that the system oscillates with a robust period, the states connected by reactions represented by dotted lines are also necessary (the SD model). 67
3.7 Simulated and experimentally measured periods as a function of the total KaiA dimer concentration. The simulated periods are reported relative to the period of the full model at 0.6 μM KaiA dimer concentration, and the experimental periods are repored relative to the wild-type clock at 0.6 μM KaiA dimer concentration. The full model and the SD model oscillate over the entire range of KaiA concentrations at which oscillations have been observed experimentally, whereas the S-only model fails at substantially lower KaiA concentrations. T432V KaiC is experimentally observed to function over a narrower range of KaiA concentrations than wild-type KaiC. 72
3.8 Fluorescence polarization measurements of fluorescently labeled KaiB reconstituted with KaiA and either wild-type KaiC or T432V KaiC. 73
3.9 A minimal model of the single-site oscillator. 75
3.10 Numerical solutions of the single-site model of the Kai oscillator (3.1) when using the parameter values $k_{01} = 10^6, k_{12} = 300, k_{su} = k_{us} = k_{20} = 1$. The solution for $A = 6/35$ has been magnified and plotted on a log scale to emphasize the different segments of the limit cycle. 77
3.11 The dependence on A of duration of the dephosphorylation and phosphorylation segments of the single-site Kai oscillator model when using the parameter values $k_{01} = 10^6, k_{12} = 300, k_{su} = k_{us} = k_{20} = 1$. The duration of the phosphorylation segment has a logarithmic singularity at $A = 1/6 - k_{20}/k_{12} + \sqrt{30k_{20}/k_{12}}/3 \approx 0.2687$. 81
3.12 The duration of the dephosphorylation and phosphorylation segments of the single-site Kai oscillator model when using the parameter values $k_{01} = 10^6, k_{12} = 300, A = 6/35$. The duration of the phosphorylation segment is evaluated with $k_{su} = 1$, and the duration of the dephosphorylation segment is evaluated with $k_{us} = 1$. 81
3.13 A minimal model of the Kai oscillator with the doubly phosphorylated state of KaiC restored. 84

3.14 KaiA dependence of the period of the single-site minimal model and the double-site minimal model. The dynamics are for the specific parameter choices $k_{01} = 650, k_{12} = 600, k_{su} = 1, k_{us} = 40, k_{20} = 5, k_{d01} = 650, k_{d12} = 200, k_{ds} = 1, k_{sd} = 1, k_{d20} = 60$. Oscillations in the single-site minimal model fail at $A \approx 0.2879$ by a fold of limit cycles, whereas the double-site minimal model exhibits robust oscillations throughout the entire physically relevant range of A. The periods in the right panel are reported relative to the periods of each respective model at $A = 6/35$ 84
3.15 The phase portrait of the single-site minimal model (3.1) for the specific parameter choices $k_{01} = 10^6, k_{12} = 300, k_{su} = k_{us} = k_{20} = 1, A = 6/35$ overlaid with asymptotic solutions of the different segments of the limit cycle. The error from the approximation (3.121) causes the asymptotic solution of the dephosphorylation segment to deviate from the numerical solution even though the asymptotic expansion of the dephosphorylation segment is accurate to $\mathcal{O}(\epsilon)$. The segments describing the transition between the dephosphorylation segment and the s_1 depletion segment all occur within a $\mathcal{O}(\delta)$ neighborhood of the point $(a = 0, s_1 = A, s_0 = 0)$ and are not visible on this scale. 103

4.1 The state space of the augmented process when computing the expected maximum deviation for a one-dimensional x. For a fixed $r = \rho$, the augmented process freely moves between all x with $|\mathbf{x}| < \rho$. The history variable r increments whenever a transition is made from x that would cause $|x| > r$. The absorbing set A becomes $A \times [0, \infty)$ in the augmented state space. 133
4.2 A diagram of the nested sets involved in the committor sequence when solving for the expected maximum deviation using (4.17). The probability that a trajectory makes an excursion of at least ρ before returning to A is given by $q_\rho(\mathbf{x})$, the committor to the set B_ρ, which can be interpreted as the complementary cumulative distribution function that a trajectory makes an excursion of at least size ρ. The expected maximum deviation is computed by integrating $q_\rho(\mathbf{x})$ over ρ. 135
4.3 The state space of the augmented process when computing the expected excursion count. S_1 and S_2 are copies of $S \setminus B$ and $S \setminus A$, respectively. A_2 and B_1 are copies of A and B stripped of all internal transitions as to avoid overcounting; states in A_2 and B_1 can only transition to their neighbors in $A_1 \subset S_1$ and $B_2 \subset S_2$. All states in S_1 and S_2 (including those in A_1 and B_2) retain their internal transitions. Transitions to the absorbing state C occur at some state-dependent rate. 137
4.4 A realization of (4.25). The phase portrait is overlaid on the vector field in the case of $z = 0$. 139
4.5 The expected maximum excursion size computed for (4.25) using the set definitions (4.27) and the metric (4.28). 140
4.6 The expectation and standard deviation of the number of $A \to B$ excursions for the system (4.25) with the set definitions (4.27). 141

LIST OF TABLES

2.1 Comparison of the performance of the constrained Metropolis-adjusted Langevin algorithm (CMALA), constrained unadjusted Langevin algorithm (CULA), unconstrained Metropolis-adjusted Langevin algorithm (MALA), no-U-turn sampler (NUTS), and ensemble MCMC (EMCEE) when estimating parameters for the three- and seven-species repressilator models using data that are consistent with those shown in Figure 2.2. The acceptance rate for CULA dynamics is below one because steps that fail to satisfy the constraints are rejected. MALA was not tested for the seven-species model owing to its failure to converge for the three-species model. 24

3.1 A list of all reactions between KaiC species in the model. Reactants and products are indexed with a binary string notation in which a KaiC monomer is specified by six binary variables, listed in the key at the top of the table: S431 phosphorylated (pS431), T432 phosphorylated (pT432), KaiA bound to CII (CII-KaiA), ADP bound to CII (CII-ADP), KaiA bound to CI (CI-KaiA), and ADP bound to CI (CI-ADP). Each variable is equal to 1 if the corresponding condition is satisfied and 0 otherwise; when ADP is not bound, ATP is bound (i.e., we do not allow for unoccupied nucleotide binding sites). A $\cdot$ is a wildcard that accommodates either 0 or 1, so long as the full six-index combination is among the allowed species. The rate constants k_a distinguish the nucleotide exchange reactions, which can only occur if KaiA is bound on CII, and the CII ATP hydrolysis reactions, which can occur regardless of the presence of KaiA. 56

3.2 Cycles of reactions that obey detailed balance. The binary string notation is the same as in Table 3.1. In addition to the reactions shown, there is a pair of reversible reactions between the first and last species in each sequence to close the cycle. Each cycle introduces a constraint that the product of the reaction constants in the forward direction must equal the product of the reaction constants in the reverse direction. 56

3.3 Conditions used to constrain the model. KaiC is at 3.5 μM throughout, and KaiB, where present, at 3.5 μM. A check in the Fit column indicates that the measured phosphoform kinetics enter the least-squares term of the log posterior; a check in the Oscillatory column indicates that the system is oscillatory and that the oscillation period enters a variance penalty on the period. Nucleotide exchange requires KaiA, so %ATP does not enter the model for KaiC alone (dash). Sustained oscillations at the higher KaiA concentrations were reported by [26]. . 59

3.4 Behavior of the marginal posterior densities for each rate constant. Reactants and products follow the same binary string notation as in Table 3.1. Group labels marked *(cont.)* continue from the previous page. 68

3.5 Dynamical variables of the minimal single-site model (3.1). Concentrations are expressed in units of the total KaiC monomer concentration. No state carries KaiA at CII and a hydrolyzed CI simultaneously, because release of KaiA from CI is fast in the doubly bound complex and that complex is contracted out of (3.1); a complex with KaiA at CII therefore reaches s_0 only after KaiA leaves CII. 75

3.6 Reactions of the single-site minimal model (3.1) (Figure 3.9). KaiC concentrations are expressed as fractions of the total KaiC monomer concentration, and a denotes a free KaiA dimer; because one KaiA dimer binds the CII domain of a KaiC hexamer, KaiA binding at CII consumes $1/6$ unit of a per KaiC monomer, whereas sequestration at CI consumes one unit of a per monomer. The k_{12} reaction consumes $1/6$ unit of a at CII and liberates the dimer held at CI, for a net gain of $5/6$ unit of a. 76

3.7 Reactions added to the single-site minimal model to obtain the double-site minimal model (Figure 3.13). The reactions of Table 3.6 are retained unchanged. The last three reactions parallel the corresponding reactions of the S phosphoform, with rate constants distinguished by the subscript d; the two T432 phosphotransfer reactions join the S and D branches. Conventions for concentrations and KaiA stoichiometry are as in Table 3.6. 83

4.1 Parameters for (4.25) and (4.26). 139

ACKNOWLEDGMENTS

I would like to thank my advisor Aaron Dinner for his guidance throughout my Doctoral studies. My research projects have been riddled with false starts, and his patience is what made these results possible. I would also like to thank professors Jonathan Weare, Michael Rust, Elizabeth Jerison, Monika Scholz, and the members of their research groups for the many projects we performed in close collaboration. In particular, Yujia Liu provided fluorescence polarization measurements of T432V KaiC and the KaiA phase response curve in Chapter 3.

ABSTRACT

Mathematical models are powerful tools for understanding physical systems. If the dynamics of the system are of interest, these models are commonly parameterized systems of ordinary differential equations (ODEs). For the model to be useful, it should reproduce known properties of the experimental system and ideally make meaningful predictions. In general, an arbitrary choice of parameters will not result in a model that is consistent with the physical system. As a first step, a suitable set of parameters must be determined such that the model can reproduce properties of the physical system, and this can represent a significant undertaking in itself.

The process of determining a set of parameters that can reproduce the known properties of the system is an example of an inverse problem. In general, inverse problems do not admit unique solutions; there can be no solutions or multiple solutions that are equally valid. Due to this indeterminacy, Monte Carlo Markov Chain (MCMC) methods are commonly used to estimate a probability distribution over the solution space to provide both a set of optimal solutions and a measure of uncertainty. The probability distributions that arise from inverse problems involving parameterized ODE models tend to be poorly conditioned in ways that degrade the performance of general MCMC methods. We first present a specialized method for sampling the distributions that arise from inverse problems involving parameterized ODEs and demonstrate its performance gains over conventional MCMC methods.

Our specialized MCMC method enables the treatment of ODE models that are difficult for conventional MCMC methods to handle without extensive tuning of hyperparameters. We apply our method to a model of the Kai oscillator. The sampling results make an unconventional prediction that the T432 phosphorylation site of KaiC is not necessary for oscillatory dynamics, but instead keeps the period of the oscillator robust over a wider range of KaiA concentrations. We confirm this prediction experimentally. To understand the mechanism by which robustness arises in the Kai oscillator, we reduce the optimized

model to a simple minimal model that we analyze using the method of matched asymptotic expansions. The asymptotic analysis reveals that the robustness in the Kai oscillator is governed by the balance between the rate of autocatalytic KaiA release and the rate of KaiA sequestration, and that the presence of T432 allows the balance to persist across a wider range of KaiA concentrations.

Lastly, we consider how to analyze noisy systems modeled with stochastic differential equations (SDEs). The presence of noise requires us to compute averages of quantities of interest instead of the deterministic quantities in the case of ODEs. We extend the framework of transition path theory, which has traditionally been used to analyze molecular simulations, to compute statistics regarding the trajectories of the system.

CHAPTER 1

INTRODUCTION

An important role of theoretical models is to make predictions regarding a system of interest. In biological systems, the high number of interacting degrees of freedom makes it intractable to describe everything in atomistic detail. Consequently, models that describe biological systems tend to be more phenomenological as the size of the system in question increases. Despite the relative imprecision of models of biological systems compared to those that may arise in more traditional fields of physics, these models can still be useful, but often have parameters that must be inferred from experimental data before they can reliably reproduce the known properties of the system, let alone make meaningful predictions. This is an inverse problem: we seek to determine inputs, in this case a set of model parameters, that will produce a desired output, in this case some experimental measurements.

Since objective functions associated with inverse problems are generally nonconvex and admit multiple solutions, Bayesian methods are commonly used to find optimal parameters and quantify their uncertainty simultaneously. Under the Bayesian framework, known information about the system, such as bounds on parameter values, is incorporated into a prior distribution, and deviation from experimental measurements is incorporated into a likelihood function. The posterior distribution quantifies the quality of a particular set of model parameters, and is proportional to the product between the prior and the likelihood. For models of realistic systems, direct evaluation of the posterior is computationally intractable and must be estimated. Conventionally, this is done using Markov Chain Monte Carlo (MCMC) methods, where trajectories are drawn from a Markov chain constructed to have the same stationary distribution as the posterior distribution. In principle, an unbiased estimate of the posterior distribution can be obtained from the empirical distribution of a sufficiently large number of samples.

The behavior of ODE solutions can cause posterior distributions that arise when estimat-

ing ODE parameters to be particularly poorly behaved, which in turn can cause the number of steps required to obtain a sufficiently converged estimate of the posterior distribution to be very large. One step of a simple MCMC algorithm for ODE parameter estimation might involve proposing a set of parameters by perturbing from a known set of parameters, solving an initial value problem (IVP) for the proposed parameters, and then computing the agreement of the ODE solution with the experimental data. This is a shooting method. For systems with even mild instability (e.g. stiff systems or those with a positive local Lyapunov exponent), small changes in the initial conditions or parameters are exponentially amplified over the solution. This manifests as an ill-conditioned posterior distribution with regions of high probability density concentrated along steep, narrow bands, limiting size of steps that the MCMC sampler can make without stalling. Moreover, bifurcations in the ODE generically occur as parameters are varied. This causes qualitative changes in the solutions that can lead to large inconsistencies with experimental data, which again manifests as very sharp variations in the posterior distribution. Even with parameters far from any bifurcations, the numerical conditioning of the ODE can cause IVP solvers to diverge, resulting in inaccurate values of the posterior or stalling of the sampler.

We address the challenges that arise when estimating parameters for ODEs by developing a specialized MCMC method that combines ideas from numerical bifurcation analysis and trajectory optimization (Chapter 2). First, we note that fixed points of ODE systems contains information that can be used as a surrogate to rapidly identify regions of the parameter space that are worth dedicating more computational resources to. For example, if the dynamics of the physical system relax to a steady state, regions of the parameter space that do not produce a fixed point that agrees with the experimentally observed steady state can be quickly eliminated. Computing the fixed point for this preliminary screen only involves finding a root of the vector field of the ODE, which is orders of magnitude cheaper than solving the IVP. When solving for the full dynamics of the ODE, the conditioning of

the solutions and by extension the posterior distribution can be improved by using collocation methods instead of the more obvious approach of single shooting. Instead of solving an IVP, the domain of integration is decomposed into a mesh, the ODE is solved locally on each mesh element, and the local solutions are required to agree at the interface between two mesh elements. Each local solution has its own sensitivity to perturbations in the parameters instead of the globally compounding sensitivity in the case of single shooting. Collocation methods also allow for easier computation of gradients of the log-posterior with respect to parameters. Gradients in the case of single shooting must be computed by integrating the sensitivity equations of the original ODE, usually the adjoint sensitivities if the original system is high-dimensional. The adjoint equations inherit any instability possessed by the original equations. In contrast, collocation methods store the ODE solution as a set of basis coefficients, making gradient computation a linear operation that is simple to evaluate using automatic differentiation tools. We enforce the collocation constraints with Lagrange multipliers and sample the constrained distribution using methods related to the RATTLE integrator used in molecular simulation. These integrators can be viewed as generalizations of the pseudoarclength continuation method used in numerical bifurcation analysis and are inherently able to track ODE solutions through fold bifurcations. Taken together, these optimizations improve the tractability of inference problems that involve large models of realistic systems.

The performance gains afforded by our specialized MCMC sampler allows us to analyze a molecularly detailed model of the Kai oscillator, the core biochemical oscillator of the circadian clock of the cyanobacterium *Synechococcus elongatus*. The Kai oscillator consists of three proteins, KaiA, KaiB, and KaiC, that, when reconstituted in vitro with adenosine triphosphate (ATP), generate free running oscillations with a 22.4 hour period that persists for weeks. The simplicity of the Kai oscillator has lead to several computational studies that have sought to reproduce specific experimental observations, but to date there have been no

attempts to account for the published properties of the system holistically.

The existing literature has resulted in a conventional understanding that the dynamics of the Kai oscillator are due to ordered phosphorylation the serine (S431) and threonine (T432) residues on the C-terminal domain (CII) of KaiC. KaiA acts as a nucleotide exchange factor for the nucleotide binding site on CII. With ATP bound on CII, KaiC phosphorylates itself, first on T432, which promotes further autophosphorylation on S431. Phosphorylation on S431 promotes ATP hydrolysis on the N-terminal domain of KaiC (CI), which induces a conformational switch that allows KaiC to recruit KaiB on CI. KaiC that has recruited a full ring of six KaiB monomers on CI then sequesters KaiA. Sequestration of KaiA prevents nucleotide exchange on CII and by consequence inhibits phosphorylation on CII, forcing KaiC to dephosphorylate and reset the cycle. Our results in Chapter 3 complicate this simple picture.

When the model of the underlying system exhibits stochastic dynamics, the Bayesian approach we have used to investigate the Kai oscillator faces the additional complication that there is now both noise in the experimental observations and noise intrinsic to the dynamics. There are methods such as sequential Monte Carlo methods designed for this class of problems, but it can be informative to first compute summary statistics of the dynamics of the system, analogous to how we first compute simple attractors for deterministic dynamics. Statistics concerning the trajectories of the stochastic system can be computed from the Kolmogorov Backward Equation or Feynman–Kac formula. This approach has been applied in the field of molecular simulation under the name of Transition Path Theory (TPT). In its simplest form, TPT can only compute quantities that do not depend on the history of the process, but we show in Chapter 4 examples of history-dependent statistics that can be computed by appropriately augmenting the state space.

CHAPTER 2

SAMPLING PARAMETERS OF ORDINARY DIFFERENTIAL EQUATIONS WITH CONSTRAINED LANGEVIN DYNAMICS

2.1 Abstract

Fitting models to data to obtain distributions of consistent parameter values is important for uncertainty quantification, model comparison, and prediction. Standard Markov chain Monte Carlo (MCMC) approaches for fitting ordinary differential equations (ODEs) to time-series data involve proposing trial parameter sets, numerically integrating the ODEs forward in time, and accepting or rejecting the trial parameter sets. When the model dynamics depend nonlinearly on the parameters, as is generally the case, trial parameter sets are often rejected, and MCMC approaches become prohibitively computationally costly to converge. Here, we build on methods for numerical continuation and trajectory optimization to introduce an approach in which we use Langevin dynamics in the joint space of variables and parameters to sample models that satisfy constraints on the dynamics. We demonstrate the method by sampling Hopf bifurcations and limit cycles of a model of a biochemical oscillator in a Bayesian framework for parameter estimation, and we attain performance that matches or exceeds the performance of leading MCMC approaches that require numerically integrating the ODEs forward in time. We describe numerical experiments that provide insight into the speedup. The method is general and can be used in any framework for parameter estimation and model selection.

2.2 Introduction

Fitting mathematical models to data provides insight that may be difficult to obtain directly from experiments. One can formulate models to test competing hypotheses and com-

pare them systematically for how readily the models fit the data while accounting for their complexity (for examples from systems biology, which is our application focus here, see [7, 16, 131, 42, 135, 68, 40, 94, 65]). In so doing, one learns the implications of the fits for parameters and variables that are not measured, and one can interpolate the measurements to arbitrary resolution. These predictions (and their uncertainties) can in turn be used to guide the design of additional experiments and manipulations for control (e.g., therapeutic interventions).

Typically, the data are insufficient to constrain models fully, and the parameters have a distribution of values that are consistent with the data. An accurate representation of this distribution is important for both quantitative and qualitative assessment of the associated models. Because a closed form expression for the distribution is rarely available, numerical methods such as Markov chain Monte Carlo (MCMC) are often used to sample the distribution. In MCMC, one generates random sequences of parameter sets in such a way that their distribution converges to the distribution of interest.

Two features of parameter distributions often make MCMC converge prohibitively slowly. First, they can be multimodal, and sequences of low probability parameter sets must be sampled for MCMC to transition between the modes. Second, parameter distributions are generally poorly scaled [59]. That is, the support of the distribution may be much narrower in certain directions than others. Because the narrow (stiff) and wide (soft) directions often involve combinations of multiple parameters, they are not obvious, and the narrowest direction often limits the separations between successive parameter sets and, in turn, the amount of space explored.

Much work has been done to address these issues. Multimodality can be treated by enhancing the sampling of the low-probability sequences of parameter sets that enable transitions [34, 93, 101, 99, 113, 124]. Poorly scaled distributions can be sampled more efficiently by schemes in which proposed changes to the parameters are asymmetric [30], affine invariant

[55, 85], or otherwise depend on the current position in parameter space [52, 137]. Nevertheless, even with enhanced sampling and preconditioning, generally 10^6 to 10^9 parameter sets must be proposed and evaluated to obtain a reasonable sample of a parameter distribution. As a result, even when a single parameter set is a fraction of a second to test, obtaining a good sample of a distribution can be prohibitively computationally costly.

Here, we develop a complementary approach based on the observation that it can be much faster to compute the limiting behavior of a (deterministic) dynamical model—for example, its fixed points and limit cycles—than to integrate it forward in time. Namely, we show how constrained Hamiltonian dynamics can be used to sample fixed points consistent with data directly; the fixed points can be used to seed more computationally expensive calculations to obtain transient behaviors. While related methods have been applied in a variety of contexts [17, 134], to the best of our knowledge, they have not been previously applied to estimating parameters for systems of ordinary differential equations (ODEs). We focus particularly on ODE systems exhibiting oscillatory dynamics [69], which can be especially difficult to sample because freedom in the phase of oscillations causes the parameter distributions of such models to be poorly scaled and multimodal. In Section 2.3, we motivate the approach by discussing the problem of sampling parameters for a model of the repressilator, an oscillatory gene regulatory network that we introduce. In Section 2.4, we provide a conceptual overview of the approach. In Section 2.5, we return to the problem of sampling parameters of the repressilator and show that, for it, we can outperform widely used MCMC methods. We conclude by discussing how the method could be extended and combined with other strategies for accelerating parameter estimation in Section 2.6. Technical details regarding our implementation of the method are given in Section 2.7. This chapter was previously published in [28].

2.3 Motivating example

Suppose selected measurements are made on a biochemical system. The problem that we consider is how to fit the data (or more generally how to satisfy a set of constraints on the dynamics) to evaluate models and/or estimate parameter values. When spatial heterogeneities are not resolved and stochasticity can be neglected, the dynamics of such a system can be modeled through a set of ODEs.

For concreteness, consider the repressilator [38], a cycle of inhibitory gene regulatory interactions that can give rise to oscillations in gene expression (Figures 2.1 and 2.2). A minimal model of this system is the set of ODEs

$$\frac{d\mathbf{y}}{ds} = \mathbf{f}(\mathbf{y}, \mathbf{k}, \mathbf{n}) \tag{2.1a}$$

$$f_j(\mathbf{y}, \mathbf{k}, \mathbf{n}) = \frac{k_{j,0}}{1 + y_{j-1}^{n_{j-1}}} - \frac{k_{j,1}}{k_{0,1}} y_j, \ j \in \{0, 1, \ldots, 2\ell\}, \tag{2.1b}$$

where y_j is the concentration of the jth gene product, $k_{j,0}$ are dimensionless expression rates, $k_{j,1}$ are dimensionless degradation rates, n_j are Hill coefficients, s is dimensionless time, and $\ell > 1$ is a positive integer. The first term on the right hand side of (2.1b) represents the inhibitory effect of the previous species in the cycle on expression and the second term represents degradation of the gene product. For a fixed ℓ, we would like to infer values for $\mathbf{k}$ and $\mathbf{n}$ that are consistent with the (simulated) data in Figure 2.2, for example.

Bayesian statistics provide a framework for accomplishing this goal [91, 131, 40]: let $\mathbf{k}$ be the parameters of the model and $\mathbf{x}$ be experimental observations. Given a prior distribution $\pi(\mathbf{k})$ and likelihood function $\pi(\mathbf{x}|\mathbf{k})$, the posterior distribution $\pi(\mathbf{k}|\mathbf{x})$ can be computed using Bayes' rule:

$$\pi(\mathbf{k}|\mathbf{x}) = \frac{\pi(\mathbf{x}|\mathbf{k})\pi(\mathbf{k})}{\pi(\mathbf{x})} \propto \pi(\mathbf{x}|\mathbf{k})\pi(\mathbf{k}). \tag{2.2}$$

That is, the probability of the model given the data is proportional to the product of our

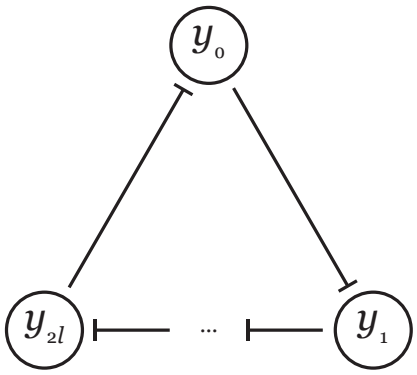


Figure 2.1: Schematic of the repressilator model with $2\ell + 1$ species, where ℓ is a positive integer.

prior belief in the model and the probability of the data given the model. Assuming that the experimental uncertainties of observable i are independent and normally distributed with variance σ_i^2, the likelihood takes the form

$$\pi(\mathbf{x}|\mathbf{k}) \propto \exp\left(-\sum_{i=1}^{N}\sum_{j=1}^{M_i}\frac{1}{2\sigma_i^2}\left(\hat{x}_i - x_i\right)^2\right), \tag{2.3}$$

where $\hat{\mathbf{x}}$ is the model output, N is the number of observables, and M_i is the number of measurements of observable i. As the likelihood can vary over several orders of magnitude and the logarithm is a strictly increasing function, computations are usually performed with the logarithm of the likelihood (log-likelihood). Maximizing the log-likelihood corresponds to minimizing the weighted sum of squared errors:

$$\log \pi(\mathbf{x}|\mathbf{k}) = \sum_{i=1}^{N}\sum_{j=1}^{M_i}\frac{1}{2\sigma_i^2}\left(\hat{x}_i - x_i\right)^2 + C, \tag{2.4}$$

where C is a constant that is independent of $\mathbf{x}$ and $\mathbf{k}$.

In general, closed form expressions for the marginal likelihood $\pi(\mathbf{x})$ are not available and numerical estimates can be difficult to compute when $\mathbf{x}$ is more than a couple of dimensions. Direct computation of $\pi(\mathbf{x})$ can be avoided by using Markov chain Monte Carlo (MCMC)

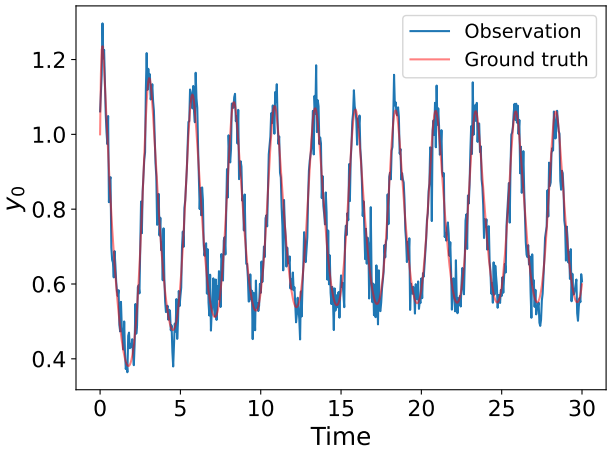


Figure 2.2: Simulated data used in parameter estimation for the three- and seven-species repressilator. We integrate the three-species equations with arbitrarily chosen oscillatory parameters forward in time using Tsitouras's 5/4 explicit Runge-Kutta method [128] to generate the ground truth data shown red. Simulated observations, shown in blue, are generated by adding normally distributed noise with mean zero and variance 2.5×10^{-3} to the concentration of the first species.

methods to sample sequences of parameter sets in such a way that their distribution converges to $\pi(\mathbf{k}|\mathbf{x})$.

Random Walk Metropolis (RWM) is one of the simplest and most widely used forms of MCMC. When applied to an ODE model like that for the repressilator above, it involves many steps of the form in Algorithm 1. In words, a random change to the parameter values is made (drawn from a normal distribution in the example shown), the posterior is evaluated (which involves integrating the ODEs), and the step is accepted with probability

$$\min\left\{\frac{\pi(\mathbf{x}|\mathbf{k}_{i+1})\pi(\mathbf{k}_{i+1})}{\pi(\mathbf{x}|\mathbf{k}_i)\pi(\mathbf{k}_i)}, 1\right\}, \tag{2.5}$$

where $\mathbf{k}_i$ are the current parameter values and $\mathbf{k}_{i+1}$ are the perturbed parameter values.

In principle, RWM provides an unbiased estimate of the target distribution when run for a sufficiently large number of steps; in practice, this number may be intractably large because the distribution may be multimodal and poorly scaled, as discussed above. There are many ways to improve the performance of MCMC, but, in the absence of a closed form expression

Algorithm 1 One step of Random Walk Metropolis

Input: current model parameters $\mathbf{k}_i$, current posterior value π_i, step size Δt
Output: updated model parameters $\mathbf{k}_{i+1}$, updated posterior value π_{i+1}

1: $\eta \sim \mathcal{N}(0, \mathbf{I})$
2: $\mathbf{k}_{i+1} \leftarrow \mathbf{k}_i + \eta\Delta t$
3: $\pi_{i+1} \leftarrow \pi(\mathbf{x}|\mathbf{k}_{i+1})\pi(\mathbf{k}_{i+1})$ ▷ Integrate (2.1b)
4: $r \sim U(0, 1)$
5: **if** $r > \pi_{i+1}/\pi_i$ **then**
6: $\quad \mathbf{k}_{i+1} \leftarrow \mathbf{k}_i$
7: $\quad \pi_{i+1} \leftarrow \pi_i$
8: **end if**
9: **return** $\mathbf{k}_{i+1}, \pi_{i+1}$

for the likelihood, the operations in step 3 of Algorithm 1 are unavoidable and usually constitute the bulk of the computational cost. Therefore, we can improve the performance of any MCMC algorithm by reducing the total amount of effort spent on step 3, through reducing either the cost of evaluating the likelihood or the number of likelihood evaluations required to generate independent samples from $\pi(\mathbf{k}|\mathbf{x})$. Our method accomplishes both of these.

2.4 Theory

The key observation that we make to improve on the efficiency of basic MCMC methods like RWM is that certain properties may be more tractable to compute when expressed as constraints. In the case of a dynamical system, limiting properties can be computed without integrating the differential equations describing the dynamics. For example, if the steady-state behavior of a system described by $\dot{\mathbf{y}}(t) = \mathbf{f}(t)$ is of interest and the system is known to converge to a fixed point at long times, it can be several orders of magnitude cheaper to compute $\mathbf{y}^*$ such that $\mathbf{f}(\mathbf{y}^*) = \mathbf{0}$ with a root-finding algorithm than to integrate $\dot{\mathbf{y}}(t) = \mathbf{f}(t)$ for times of interest. For systems such as the repressilator with complicated limiting dynamics such an approach does not provide a complete description of the system, but it can still provide a way to characterize regions of the parameter space rapidly. To the

extent that a system is known to satisfy certain constraints, it can be better to introduce them directly through equations of the form $\mathbf{f}(\mathbf{y}) = \mathbf{0}$ than indirectly through penalty terms added to the log-likelihood because the latter strategy tends to make the support for the target distribution narrow and limit the step size, as discussed above. Thus our goal is to develop a way to sample parameter values that satisfy constraints of the form $\mathbf{f}(\mathbf{y}) = \mathbf{0}$.

More precisely, suppose that some properties of a system can be expressed as a system of equations $\mathbf{f}(\mathbf{y}, \mathbf{k}) = \mathbf{0}$, where $\mathbf{f} : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}^m$, $\mathbf{y} \in \mathbb{R}^d$ describes the state of the system, and $\mathbf{k} \in \mathbb{R}^n$ is a vector of adjustable parameters. Our goal is to sample points from the solution set $\mathcal{M} \equiv \{(\mathbf{y}, \mathbf{k}) \,|\, \mathbf{f}(\mathbf{y}, \mathbf{k}) = \mathbf{0}\}$ efficiently. We show how numerical continuation can be used for the case $n = 1$ in Section 2.4.1. This procedure is equivalent to running Hamiltonian dynamics on a constraint curve, which suggests a general approach for sampling based on (deterministic) Hamiltonian dynamics for $n \geq 1$. We present the direct extension to multidimensional constraint manifolds in Section 2.4.2 and then explicitly consider its application to the Bayesian framework with a specific (stochastic) numerical integrator in Section 2.4.3.

The resulting method is related to ones used to constrain the lengths and angles of bonds in molecular simulations [2, 6, 82, 84], which enables larger timesteps. More generally, [134] shows that constrained distributions involving up to $\sim 10^3$ variables can be tractably sampled by a symmetric quasi-Newton method similar to the approach to solve for constraints that we describe in Section 2.7.3. Considerations for convergence and bias of constrained Monte Carlo methods are discussed in [86, 87].

2.4.1 Numerical continuation in a single parameter

We first consider the case where $\mathbf{f} : \mathbb{R}^{d+1} \to \mathbb{R}^d$, $\mathbf{y} \in \mathbb{R}^d$, $\mathbf{k} \in \mathbb{R}$. Suppose that $(\mathbf{y}(0), \mathbf{k}(0))$ is a solution to $\mathbf{f}(\mathbf{y}, \mathbf{k}) = \mathbf{0}$ and let $\mathbf{f}_\mathbf{y}$ denote the matrix with elements $\partial f_i / \partial y_j$. If $\mathbf{f}_\mathbf{y}$ is nonsingular at $(\mathbf{y}(0), \mathbf{k}(0))$, then by the implicit function theorem, there is a one-dimensional

curve $\mathcal{M} \subset \mathbb{R}^{d+1}$ in a neighborhood of $(\mathbf{y}(0), \mathbf{k}(0))$ consisting of solutions to $\mathbf{f}(\mathbf{y}, \mathbf{k}) = \mathbf{0}$. To obtain additional solutions to $\mathbf{f}(\mathbf{y}, \mathbf{k}) = \mathbf{0}$, $\mathcal{M}$ can be parameterized as the t-dependent solution to $\mathbf{f}(\mathbf{y}(t), \mathbf{k}(t)) = \mathbf{0}$, and a path of solutions that originate from the initial known solution $(\mathbf{y}(0), \mathbf{k}(0))$ can be extended as long as $\mathbf{f}_\mathbf{y}$ remains nonsingular. A simple way to parameterize $\mathcal{M}$ is

$$\begin{aligned} \mathbf{f}(\mathbf{y}(t), \mathbf{k}(t)) &= \mathbf{0} \\ \dot{\mathbf{k}}(t) &= \mathbf{1}, \end{aligned} \tag{2.6}$$

where $\dot{\mathbf{k}}$ is the derivative of $\mathbf{k}$ with respect to t, and $\mathbf{0}$ and $\mathbf{1}$ are vectors of zeros and ones, respectively. This parameterization is known as natural parameter continuation (NPC). Differentiating $\mathbf{f}(\mathbf{y}(t), \mathbf{k}(t))$ results in the system of ODEs

$$\dot{\mathbf{y}}(t) = -\mathbf{f}_\mathbf{y}(\mathbf{y}(t), \mathbf{k}(t))^{-1} \mathbf{f}_\mathbf{k}(\mathbf{y}(t), \mathbf{k}(t)) \tag{2.7a}$$

$$\dot{\mathbf{k}}(t) = \mathbf{1}. \tag{2.7b}$$

In principle, the curve defined by $\mathbf{f}(\mathbf{y}, \mathbf{k}) = \mathbf{0}$ can be computed by numerically integrating (2.7), but this approach generally leads to issues: solving (2.7) requires inverting $\mathbf{f}_\mathbf{y}(t)$, and the solution is undefined when $\mathbf{f}_\mathbf{y}(t)$ is singular, which generically occurs at turning points in the solution curve of $\mathbf{f}(\mathbf{y}, \mathbf{k}) = \mathbf{0}$ (Figure 2.3(left)). One way to resolve these issues is to parameterize $\mathcal{M}$ instead as the solution to the equations

$$\mathbf{f}(\mathbf{y}(t), \mathbf{k}(t)) = \mathbf{0} \tag{2.8a}$$

$$\|\dot{\mathbf{y}}(t)\|^2 + \|\dot{\mathbf{k}}(t)\|^2 = \mathbf{1}. \tag{2.8b}$$

This parameterization is used to define the pseudoarclength continuation (PALC) procedure in [23]. As previously, we can differentiate $\mathbf{f}(\mathbf{y}(t), \mathbf{k}(t))$ with respect to t to obtain a system of ODEs that can be numerically integrated. Equation (2.8b) has the form of a kinetic energy, so (2.8) can be interpreted as describing the motion of a particle constrained to slide along a one-

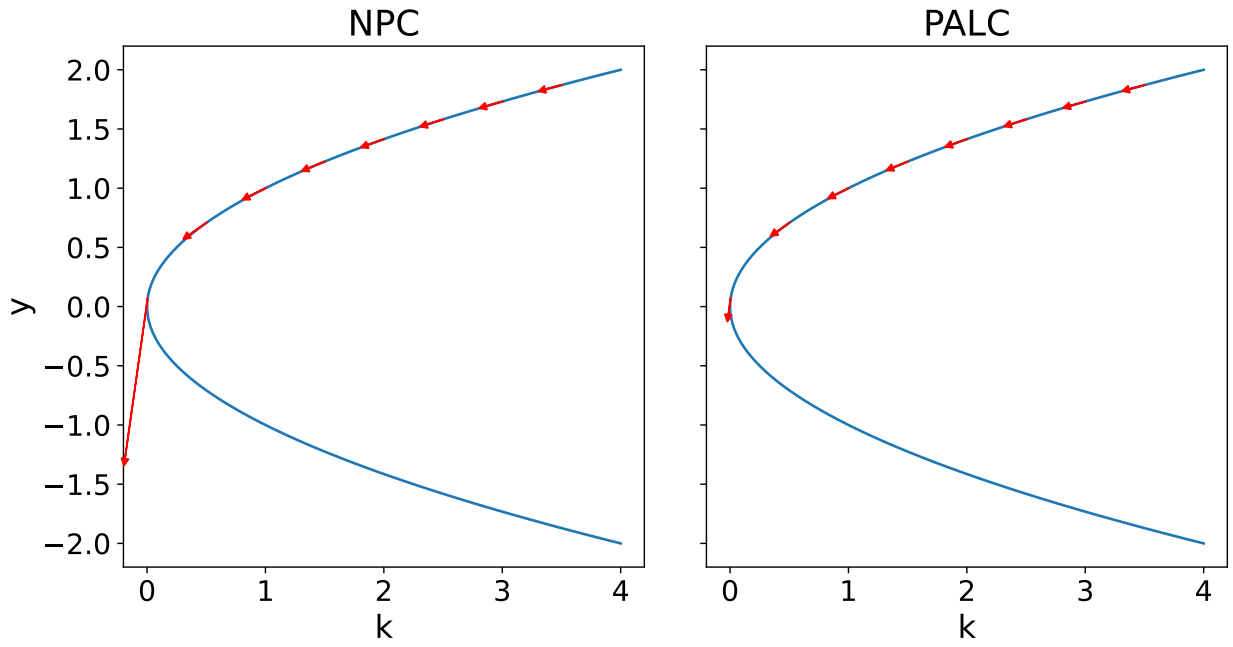


Figure 2.3: Behavior of the derivative vector $(\ \dot{\mathbf{y}}\ \ \dot{\mathbf{k}}\ )$ when approaching a turning point at $(\mathbf{0}, \mathbf{0})$. (left) In the case of natural parameter continuation, the derivative vector of (2.7) increases in norm when $(\mathbf{y}, \mathbf{k})$ approaches $(\mathbf{0}, \mathbf{0})$ due to the constraint $\dot{\mathbf{k}}(t) = 1$ and is undefined at $(\mathbf{0}, \mathbf{0})$. (right) In pseudoarclength continuation (PALC), the derivative vector is constrained to have unit norm in the system of equations in (2.8) and is well defined even at $(\mathbf{y}, \mathbf{k}) = (\mathbf{0}, \mathbf{0})$.

dimensional rail at constant speed. Because the speed is constant, (2.8) is well-behaved even when $\mathbf{f_y}$ is singular (Figure 2.3(right)). This suggests constrained Hamiltonian dynamics as a natural generalization for computing points on constraint manifolds in problems involving more than one parameter.

2.4.2 Constrained Hamiltonian dynamics

For clarity, we now treat the dynamical variables and the parameters together by defining $\mathbf{q} \equiv (\ \mathbf{y}\ \ \mathbf{k}\ )$. Letting $\mathbf{p}$ denote the conjugate momenta, the Hamiltonian of a general mechanical system is

$$\mathcal{H}(\mathbf{q}, \mathbf{p}) = T(\mathbf{p}) + U(\mathbf{q}), \tag{2.9}$$

where $T(\mathbf{p}) = \frac{1}{2}\mathbf{p}^\mathsf{T}\mathbf{M}^{-1}\mathbf{p}$ is the kinetic energy, $U(\mathbf{q})$ is the potential, and $\mathbf{M}$ is a symmetric positive definite mass matrix. The constrained dynamics of the system can be obtained by solving

$$\begin{aligned}\dot{\mathbf{q}} &= \mathcal{H}_{\mathbf{p}}(\mathbf{q},\mathbf{p}) = \mathbf{M}^{-1}\mathbf{p}\\ \dot{\mathbf{p}} &= -\mathcal{H}_{\mathbf{q}}(\mathbf{q},\mathbf{p}) - \boldsymbol{\lambda}^\mathsf{T}\mathbf{c}_{\mathbf{q}}(\mathbf{q}) = -U_{\mathbf{q}}(\mathbf{q}) - \boldsymbol{\lambda}^\mathsf{T}\mathbf{c}_{\mathbf{q}}(\mathbf{q})\\ \mathbf{0} &= \mathbf{c}(\mathbf{q}),\end{aligned} \tag{2.10}$$

where $U_{\mathbf{q}}$ is the gradient of U, and $\mathbf{c}_{\mathbf{q}}$ denotes the matrix with elements $\partial c_i/\partial q_j$, $\boldsymbol{\lambda} \in \mathbb{R}^m$ is a vector of Lagrange multipliers. In general, (2.10) cannot be solved analytically and must be integrated numerically. We use an operator splitting method [84] to integrate (2.10), as we describe in Section 2.7.1.

2.4.3 Hamiltonian Monte Carlo methods

Our goal is to sample a distribution, $\pi(\mathbf{q})$ (e.g., representing (2.2)), subject to constraints (e.g., imposed by (2.8) for an $\mathbf{f}$ such as that in (2.1b)). To this end, we define the potential

$$U(\mathbf{q}) = -\log(\pi(\mathbf{q})). \tag{2.11}$$

Deterministically integrating (2.10) does not generate samples from the whole target distribution owing to energy conservation, and it is necessary to introduce randomness. A few schemes with empirically good performance are jittered Hamiltonian Monte Carlo [63], no-U-turn sampling [64], and underdamped Langevin dynamics [115]. The advantage of these schemes over RWM is that they make use of the gradient of $U(\mathbf{q})$, which directs sampling toward low-energy/high-probability states. For certain classes of target distributions, such as those consisting of n independent and identically distributed components [9] and strongly log-concave distributions that are sufficiently smooth [92], the advantage of gradient based samplers over RWM can be shown to grow with dimension: $\mathcal{O}(n^{1/4})$ steps are required to

generate an independent sample using Hamiltonian Monte Carlo, whereas $\mathcal{O}(n)$ steps are required when using RWM.

For the present study, we use Langevin dynamics. To this end, we modify (2.10) to obtain the system of stochastic differential equations:

$$\dot{\mathbf{q}} = \mathbf{M}^{-1}\mathbf{p} \tag{2.12a}$$

$$\dot{\mathbf{p}} = -U_{\mathbf{q}} - \gamma\mathbf{p} + \sqrt{2T\gamma}\mathbf{M}^{1/2}\boldsymbol{\eta}(t) - \boldsymbol{\lambda}^{\mathsf{T}}\mathbf{c}_{\mathbf{q}}(\mathbf{q}) \tag{2.12b}$$

$$\mathbf{0} = \mathbf{c}(\mathbf{q}), \tag{2.12c}$$

where γ is a friction coefficient, T is the temperature, and $\boldsymbol{\eta}$ is a white-noise vector of independent components, each with unit variance. We integrate these equations numerically by building on the operator splitting scheme mentioned previously [84], as we describe in Section 2.7.2.

Our method for enforcing constraints is related to the SHAKE [118] and RATTLE [2] methods used in molecular simulation. SHAKE with the velocity Verlet integrator can be viewed as an unconstrained half step in $\mathbf{p}$, followed by a step in $\mathbf{q}$ with projection onto the constraint manifold $\mathcal{M}$, and then another unconstrained half step in $\mathbf{p}$. However, even if the position constraints are satisfied, the dynamics with SHAKE do not satisfy the cotangency constraints $\mathbf{c}_{\mathbf{q}}(\mathbf{q})\mathbf{M}^{-1}\mathbf{p} = \mathbf{0}$ in general. To enforce the cotangency constraints, RATTLE with the velocity Verlet integrator adds an orthogonal projection of $\mathbf{p}$ onto the cotangent space after the last half step in $\mathbf{p}$. The operator splitting method that we use (see Section 2.7.2) is similar to RATTLE with the velocity Verlet integrator but also performs an orthogonal projection of $\mathbf{p}$ onto the cotangent space after the first half step in $\mathbf{p}$.

Because the nonlinear projection performed in our sampler is the same as that performed in pseudoarclength continuation, our sampler inherits the properties of pseudoarclength continuation; in particular, the dynamics can move past fold bifurcations if the step size is small

enough. However, other bifurcations may require special treatment. One can monitor and control for local bifurcations by including auxiliary variables and constraint equations that define the bifurcation during sampling, as we do for Hopf bifurcations in Section 2.5.2.

2.5 Estimating model parameters for the repressilator

To demonstrate our method and compare its performance with established Monte Carlo methods for Bayesian parameter estimation, we fit the model for the repressilator in (2.1b) to simulated data. When the number of species in the cycle is odd (integer l in (2.1b) and Figure 2.1), the dynamics tend to either a globally attracting fixed point or a globally attracting limit cycle [121]. These relatively simple dynamics make the repressilator a tractable yet nontrivial test of our methods.

To enforce nonnegativity of both concentrations and rate constants, we make a change of variables to logarithmic coordinates:

$$\begin{aligned} (\ \tilde{\mathbf{y}} \quad \tilde{\mathbf{k}}\ ) &\equiv (\ \log(\mathbf{y}) \quad \log(\mathbf{k})\ ) \\ \frac{d\tilde{\mathbf{y}}}{ds} &= \tilde{\mathbf{f}}(\tilde{\mathbf{y}}, \tilde{\mathbf{k}}, \mathbf{n}) \\ f_j(\tilde{\mathbf{y}}, \tilde{\mathbf{k}}, \mathbf{n}) &= \frac{\exp\left(\tilde{k}_{j,0} - \tilde{y}_j\right)}{1 + \exp\left(n_{j-1}\tilde{y}_{j-1}\right)} - \exp\left(\tilde{k}_{j,1} - \tilde{k}_{0,1}\right),\ j \in \{0, 1, \ldots, 2l\}. \end{aligned} \tag{2.13}$$

For notational convenience, we henceforth omit the tildes on $\tilde{\mathbf{y}}$ and $\tilde{\mathbf{k}}$ where there is no risk of confusion. The repressilator model with $2\ell + 1$ species has a total of $6\ell + 2$ independent parameters after setting the timescale by fixing $k_{0,1} = 0$: the synthesis rates $\{k_{0,i}\}_{i=0}^{2\ell}$, the remaining degradation rates $\{k_{1,i}\}_{i=1}^{2\ell}$, and the Hill coefficients $\{n_i\}_{i=0}^{2\ell}$. We generate simulated data by choosing arbitrary values for the parameters, integrating the equations forward in time using Tsitouras's 5/4 explicit Runge-Kutta method [128], and adding normally distributed noise with mean zero and variance 2.5×10^{-3} to the concentration of the first species (prior to transforming to logarithmic coordinates) (Figure 2.2). That is, we

assume only one of the $2\ell + 1$ species is observable. We then sample from a sequence of constrained distributions, each of which incorporates more information regarding the model and is more computationally demanding than the previous one.

The results are summarized in Figure 2.4, where we project the sampled parameter sets onto three pairs of parameters. The top row shows sampling of all fixed points. The sampling is nearly uniform, indicating that the procedure readily explores the space. The middle row shows sampling of Hopf bifurcations. The constraints of the Jacobian (discussed below) restrict the sampling to a subspace, though it remains quite smooth. The bottom row shows sampling of limit cycles that fit the data. The sampling is most constrained in this case. We discuss how we obtain these results in Sections 2.5.1 to 2.5.3, and we compare our procedure with unconstrained MCMC algorithms in Section 2.5.4. We use a step size of 10^{-1}, friction coefficient $\gamma = 10^{-1}$, and mass matrix $\mathbf{M} = \mathbf{I}$ throughout.

2.5.1 Sampling fixed points

Let $\mathbf{q} \equiv (\ \mathbf{y}\ \ \mathbf{k}\ \ \mathbf{n}\ )$. If the right hand side of (2.13) is specified as the constraint $\mathbf{c}(\mathbf{q}) = \mathbf{0}$ to be satisfied in (2.10), then assuming that $\mathbf{c_q}$ is nonsingular, every $\mathbf{q}$ sampled gives a set of parameters for which (2.13) has a fixed point at $\mathbf{y}$. This can be useful, for instance, if the steady-state behavior of an experimental system is known, allowing the space of parameters to be restricted without performing costly integration of the system of ODEs.

During sampling, we bound the search space by including half quadratic penalty terms in the log-prior:

$$\begin{aligned} &100\,\|(\mathbf{k}-\mathbf{k}_{\max})\odot\mathbb{1}\{\mathbf{k}>\mathbf{k}_{\max}\}\|^2 + 100\,\|(\mathbf{k}-\mathbf{k}_{\min})\odot\mathbb{1}\{\mathbf{k}<\mathbf{k}_{\min}\}\|^2 + \\ &100\,\|(\mathbf{n}-\mathbf{n}_{\max})\odot\mathbb{1}\{\mathbf{n}>\mathbf{n}_{\max}\}\|^2 + 100\,\|(\mathbf{n}-\mathbf{n}_{\min})\odot\mathbb{1}\{\mathbf{n}<\mathbf{n}_{\min}\}\|^2, \end{aligned} \tag{2.14}$$

where $\mathbb{1}\{\mathbf{k} > \mathbf{k}_{\max}\}$ compares $\mathbf{k}$ and $\mathbf{k}_{\max}$ componentwise and returns a vector with components equal to 1 wherever $\mathbf{k} > \mathbf{k}_{\max}$ and 0 otherwise, and $\odot$ denotes the componentwise

product. We choose $\mathbf{k}_{\min}$ to be a length $4\ell+1$ vector with all components equal to -5, $\mathbf{k}_{\max}$ to be a length $4\ell+1$ vector with all components equal to 5, $\mathbf{n}_{\min}$ to be a length $2\ell+1$ vector with all components equal to 0, and $\mathbf{n}_{\max}$ to be a length $2\ell+1$ vector with all components equal to 10. We sample a single chain for 10^7 steps and subsample every 10 steps.

As noted above, the results are shown in the top row of Figure 2.4. Since no additional terms other than the restraints (3.14) are included in the posterior, the sampled points should be uniformly distributed across the solution set of $\mathbf{f}(\mathbf{y}, \mathbf{k}, \mathbf{n}) = \mathbf{0}$. However, we note that the points are not necessarily uniformly distributed throughout the parameter space due to the curvature of the solution manifold. In particular, the top right panel of Figure 2.4 shows points concentrated along $n_2 = 0$. This is a consequence of sampling in logarithmic concentrations. When n_2 is small, combinations of parameters that cause the steady-state concentration of y_1 to be large correspondingly cause the steady-state concentration of y_2 to be close to zero. The logarithm of the concentration of y_2 in this case is uniformly distributed across a large range of negative values, which results in the sampled points concentrating near $n_2 = 0$ when they are projected onto the (n_1, n_2) plane. If desired, one can correct for the curvature of the solution manifold when projecting onto the parameter space by reweighting by the transformed volume element of the solution manifold (see Section 2.7.6).

2.5.2 Sampling Hopf points

For the repressilator, we are interested in estimating model parameters from oscillatory data, so we seek sets of parameter values that give rise to limit cycles. Empirically, we find that these sets represent a small fraction of the space of possible parameter values. To boost efficiency, we can exploit the fact that one way for limit cycles to appear is through Hopf bifurcations and restrict the sampling to a manifold of Hopf bifurcation points. At a Hopf bifurcation, the Jacobian of a system of ODEs evaluated at a fixed point has a complex conjugate pair of imaginary eigenvalues that crosses the imaginary axis, causing the fixed

point to change stability and give rise to a small amplitude limit cycle.

Numerically, the exact point of crossing, at which the Jacobian has a purely imaginary pair of eigenvalues, can be computed as the solution to the system of equations

$$\mathbf{f}(\mathbf{y}, \mathbf{k}, \mathbf{n}) = \mathbf{0} \tag{2.15a}$$

$$(\mathbf{f}_{\mathbf{y}}(\mathbf{y}, \mathbf{k}, \mathbf{n}) - i\omega\mathbf{I})\mathbf{v} = \mathbf{0} \tag{2.15b}$$

$$\mathbf{v}^{\mathsf{T}}\mathbf{v} = 1 \tag{2.15c}$$

$$\mathrm{Im}(v_1) = 0, \tag{2.15d}$$

where $\mathbf{v} \in \mathbb{C}^d$ is the eigenvector of $\mathbf{f}_{\mathbf{y}}(\mathbf{y}, \mathbf{k})$ associated with the eigenvalue $i\omega$, and v_1 denotes the first component of $\mathbf{v}$. We incorporate (2.15) as real valued constraints to be satisfied during sampling after splitting $\mathbf{v}$ into its real and imaginary parts:

$$\mathbf{c}(\mathbf{q}) \equiv \begin{cases} \mathbf{f}(\mathbf{y}, \mathbf{k}, \mathbf{n}) = \mathbf{0} \\ \mathbf{f}_{\mathbf{y}}(\mathbf{y}, \mathbf{k}, \mathbf{n})\,\mathrm{Re}(\mathbf{v}) + \omega\,\mathrm{Im}(\mathbf{v}) = \mathbf{0} \\ \mathbf{f}_{\mathbf{y}}(\mathbf{y}, \mathbf{k}, \mathbf{n})\,\mathrm{Im}(\mathbf{v}) - \omega\,\mathrm{Re}(\mathbf{v}) = \mathbf{0} \\ \mathrm{Re}(\mathbf{v})^{\mathsf{T}}\,\mathrm{Re}(\mathbf{v}) + \mathrm{Im}(\mathbf{v})^{\mathsf{T}}\,\mathrm{Im}(\mathbf{v}) = 1 \\ \mathrm{Im}(v_1) = 0. \end{cases} \tag{2.16}$$

This is the fully extended system for Hopf bifurcations [56]. Sampling $\mathbf{q} \equiv (\ \mathbf{y}\ \ \mathbf{k}\ \ \mathbf{n}\ \ \mathbf{v}\ \ \omega\ )$ subject to (3.22) gives $\mathbf{k}$ and $\mathbf{n}$ for which $\mathbf{f}(\mathbf{y}, \mathbf{k}, \mathbf{n})$ has a Hopf point at $\mathbf{y}$. Solving (3.22) is computationally more demanding than solving for the fixed points of (2.13) but is still orders of magnitude computationally cheaper than numerically integrating (2.13). The period estimate $2\pi/\omega$ can be used to restrict the space of parameters for ODE models of oscillatory systems further and provide reasonable starting points for performing sampling that involves the computationally expensive time integration of (2.13).

The fixed points of (2.13) sampled in Section 2.5.1 cannot be directly used to initiate

Hopf-point sampling since they generally do not have any purely imaginary eigenvalues and thus do not satisfy (3.22). To find starting points consistent with (3.22), we compute the eigenvalues and eigenvectors of the Jacobian matrix at each of the fixed points sampled in Section 2.5.1 and then select the eigenvalue-eigenvector pair whose eigenvalue has the smallest real part relative to its imaginary part:

$$\zeta_i^* = \min_j \left| \frac{\mathrm{Re}(\zeta_{ij})}{\mathrm{Im}(\zeta_{ij})} \right|, \tag{2.17}$$

where ζ_{ij} is the jth eigenvalue of the ith fixed point sampled and $\mathbf{v}_{ij}$ is its corresponding eigenvector. We then select $(\ \mathbf{y}_i\ \ \mathbf{k}_i\ \ \mathbf{n}_i\ )$ for which $|\,\mathrm{Re}(\zeta_i^*)/\,\mathrm{Im}(\zeta_i^*)|$ is small, set $\mathbf{q}_0 \equiv (\ \mathbf{y}_i\ \ \mathbf{k}_i\ \ \mathbf{n}_i\ \ \mathbf{v}_i^*\ \ \zeta_i^*\ )$, and solve (3.22) using Gauss–Newton iteration starting from $\mathbf{q}_0$ as an initial guess. Upon success, we obtain a point $\mathbf{q}$ consistent with the constraints in (3.22) from which we initiate Hopf-point sampling.

During Hopf-point sampling, we again bound the search with the prior in (3.14). We sample a single chain for 10^7 steps and subsample every 10 steps. The results are shown in the middle row of Figure 2.4.

2.5.3 Sampling limit cycles

With candidate parameters obtained from sampling the solution manifold of (3.22), further sampling involving the explicit computation of limit cycle solutions to (2.13) can be initiated from selected points. We formulate the limit cycle computation as a boundary value problem:

$$\begin{aligned} \frac{d\mathbf{y}}{ds} &= \mathbf{f}(\mathbf{y}, \mathbf{k}) \\ \mathbf{y}(0) &= \mathbf{y}(\tau). \end{aligned} \tag{2.18}$$

We discretize (3.29) using fourth-order Gauss–Legendre collocation on 60 mesh intervals (see Section 2.7.4) to obtain a system of $241(2\ell+1)$ equations whose solution $\mathbf{u}(s)$ approximates

$\mathbf{y}(s)$. Incorporating the discretized equations into $\mathbf{c}(\mathbf{q})$ and sampling $\mathbf{q} = (\ \mathbf{u}(s)\ \ \mathbf{k}\ \ \mathbf{n}\ \ \tau\ )$ gives points that allow for periodic solutions to (2.13). While this constraint does not exclude fixed point solutions, which are trivially periodic, such solutions can be discouraged by including a penalty term for solutions with arc lengths below a threshold in the prior. We compute the arc length as

$$L = \int_0^\tau \|\dot{\mathbf{u}}(s)\|\, ds \tag{2.19}$$

and use a penalty of the form

$$\frac{1}{4}\left(\frac{L_0^2}{L^2} - 1\right)^2 \mathbb{1}\{L < L_0\} \tag{2.20}$$

with $L_0 = 0.3$ for the results shown.

In addition to the prior terms in (3.14) and (3.32), we include a least squares log-likelihood term of the form in (2.4). Thus the full log-posterior is

$$\begin{aligned}\pi(\mathbf{q}|\mathbf{x}) = &\sum_{i=1}^{N}\sum_{j=1}^{M_i} \frac{1}{2\sigma_i^2}\left(\hat{x}_i - x_i\right)^2 + \\ &\frac{1}{4}\left(\frac{L_0^2}{L^2} - 1\right)^2 \mathbb{1}\{L < L_0\} + \frac{(\tau - \tau_{\text{data}})^2}{2\sigma^2} + \\ &100\left\|(\mathbf{k} - \mathbf{k}^{\min}) \odot \mathbb{1}\{\mathbf{k} > \mathbf{k}^{\min}\}\right\|^2 + 100\left\|(\mathbf{k} - \mathbf{k}^{\max}) \odot \mathbb{1}\{\mathbf{k} > \mathbf{k}^{\max}\}\right\|^2 + \\ &100\left\|(\mathbf{n} - \mathbf{n}_{\max}) \odot \mathbb{1}\{\mathbf{n} > \mathbf{n}_{\max}\}\right\|^2 + 100\left\|(\mathbf{n} - \mathbf{n}_{\min}) \odot \mathbb{1}\{\mathbf{n} < \mathbf{n}_{\min}\}\right\|^2.\end{aligned} \tag{2.21}$$

We process the simulated data shown in Figure 2.2 to use as $\mathbf{x}$ in the log-likelihood. We estimate the period τ_{data} by dividing the total integration time by the highest amplitude frequency of the fast Fourier transform of the data, excluding the zero-frequency mode. We then segment the data into pieces of length τ_{data}, average them, and scale the time so the period is 1. The model output $\hat{\mathbf{x}}$ in the least squares log-likelihood is $\exp(y_0)$ evaluated at the time associated with each data point. We use $\sigma_i = 0.05$ for all i in the log-likelihood

term (2.4). We reintroduce the period by including a quadratic penalty on the difference between the period of the model output τ and τ_{data} in the log-likelihood: $(\tau - \tau_{\text{data}})^2 / 2\sigma^2$ with $\sigma = 0.05$.

To initiate sampling, we choose $\mathbf{k}$ from the Hopf-point samples that minimize $(2\pi/\omega - \tau_{\text{data}})^2$ and use a fifth-order Radau IIA method to integrate (2.13) forward starting from the initial conditions that were used to generate the simulated data (which we assume are known) for $20\pi/\omega$ time units. The solution over the last $2\pi/\omega$ length interval of time is taken to be $\mathbf{y}(t)$. This $\mathbf{y}(t)$, along with $\mathbf{k}$ and τ are used as an initial guess to solve (3.29) using Gauss–Newton iteration to obtain an initial point for sampling with periodicity constraints. We sample 10 independent chains for 10^7 steps each and subsample every 10 steps. Results are shown in the bottom row of Figure 2.4.

2.5.4 Comparison with unconstrained MCMC methods

To assess the speedup from our method, we compare our results to three widely-used MCMC methods: an affine-invariant ensemble sampler (henceforth, ensemble MCMC) [55, 43], the unconstrained Metropolis-adjusted Langevin algorithm (MALA) [66], and the no-U-turn sampler (NUTS) [64]. Descriptions of each of the methods can be found in Section 2.7.5.

As we do in Section 2.5.3, for the constrained Langevin dynamics, we use the restraint in (3.14) to bound the search space, the arc-length penalty term in (3.32) to discourage constant solutions, and a log-likelihood of the form (2.4). In addition to the model parameters $\mathbf{k}$ and $\mathbf{n}$, we include the initial conditions $\mathbf{y}(0)$ as parameters to be sampled to align the phase of the model output with the phase of the data. The total set of parameters sampled with the unconstrained methods is $(\ \mathbf{k}\ \ \mathbf{n}\ \ \mathbf{y}(0)\ )$. We process the simulated data the same way as described in Section 2.5.3 to obtain $\mathbf{x}$ but instead average over a window of $2\tau_{\text{data}}$ to better control the period of the model output. To obtain the model output $\hat{\mathbf{x}}$ when evaluating the log-likelihood for the unconstrained MCMC methods, we use Tsitouras's 5/4 explicit

Table 2.1: Comparison of the performance of the constrained Metropolis-adjusted Langevin algorithm (CMALA), constrained unadjusted Langevin algorithm (CULA), unconstrained Metropolis-adjusted Langevin algorithm (MALA), no-U-turn sampler (NUTS), and ensemble MCMC (EMCEE) when estimating parameters for the three- and seven-species repressilator models using data that are consistent with those shown in Figure 2.2. The acceptance rate for CULA dynamics is below one because steps that fail to satisfy the constraints are rejected. MALA was not tested for the seven-species model owing to its failure to converge for the three-species model.

Species	Method	Acceptance rate	Steps for $\hat{R} < 1.1$	Average ESS/step	Minimum ESS/step
3	EMCEE	0.0397	2×10^8	9.96×10^{-6}	4.55×10^{-6}
3	CMALA	0.8332	1×10^7	1.87×10^{-4}	8.61×10^{-5}
3	CULA	0.9561	1×10^6	2.77×10^{-3}	1.76×10^{-3}
3	MALA	0.6277	$\gg 1 \times 10^8$	6.95×10^{-8}	6.19×10^{-8}
3	NUTS	0.8537	2.5×10^7	4.12×10^{-5}	1.87×10^{-5}
7	EMCEE	0.0134	$\gg 4 \times 10^8$	5.71×10^{-7}	4.81×10^{-8}
7	CMALA	0.7916	1.4×10^8	9.23×10^{-6}	9.27×10^{-7}
7	CULA	0.9647	1.5×10^7	1.38×10^{-4}	1.28×10^{-5}
7	NUTS	0.9087	5.4×10^7	1.05×10^{-5}	6.83×10^{-6}

Runge-Kutta method [128] to integrate (2.13) forward for 30 time units starting from $\mathbf{y}(0)$. We then take $\hat{\mathbf{x}}$ to be $\exp(y_0)$ on the interval $t \in [30 - 2\tau_{\text{data}}, 30]$.

For all the methods, the bulk of the computational cost comes from numerical integration of (2.13): for the constrained Langevin sampler, we solve a two-point boundary value problem for the constraints at each step, and, for the unconstrained methods, we solve an initial value problem at each step (each attempted walker update in ensemble MCMC and each iteration of the symplectic integrator in the no-U-turn sampler). As such, the cost of a step is similar for all the methods, and we compare their performance in terms of the total number of steps.

One way to quantify the relative performance of the algorithms is the effective sample size (ESS) per step (Table 2.1). For a one-dimensional distribution, the ESS is given by

$$N_{\text{eff}} = \frac{N}{-1 + 2\sum_{t=0}^{\infty} C(t)}, \tag{2.22}$$

where N is the number of samples, and $C(t)$ is the autocorrelation function at time lag t [122]. In practice, the sum must be truncated at some finite time lag. We use the criterion given in [50]. For the example that we consider, there are multiple parameters, and we report the average and minimum ESS over all of them.

To monitor convergence, we compute a version of the potential scale reduction factor $\hat{R}$ described in [48] for $\mathbf{k}$. For a set of M independent chains each with N samples, $\hat{R} = \|\mathbf{\Sigma}_a^{-1}\mathbf{\Sigma}\|_2$, where $\mathbf{\Sigma}_a$ is the average within-chain covariance and $\mathbf{\Sigma}$ is an estimate of the covariance of the stationary distribution, which includes the between-chain covariance (see Section 2.7.7 for formulas). If the sampling is converged, the within-chain covariance $\mathbf{\Sigma}_a$ and stationary covariance $\mathbf{\Sigma}$ matrices should be approximately equal, and the 2-norm (i.e., square root of the largest singular value) of the matrix $\mathbf{\Sigma}_a^{-1}\mathbf{\Sigma}$ should be approximately one.

Table 2.1 gives the ESS and potential scale reduction factors for each of the methods when applied to sampling parameters for the three- and seven-species repressilator models. Best fits and marginal distributions of the parameters for the three-species repressilator model are shown in Figures 2.5 and 2.6, respectively. Both the constrained and unconstrained methods produce fits that are visually reasonable, but some unconstrained methods are less efficient at sampling. The ensemble MCMC sampler has a low acceptance rate due to the high sensitivity to changes in parameter values of initial-value problem solutions to (2.13): small changes in parameter values can cause large changes in the phase, resulting in rapid variations in the posterior distribution (Figure 2.7). Consequently, proposals generated by (2.53) using two points close in Euclidean distance are frequently rejected as they often land in low probability regions. Consequently, the ESS per step of ensemble MCMC is about 20 times lower than that for the constrained Metropolis-adjusted Langevin algorithm (CMALA). Consistent with the ESS, the required number of steps for convergence is about 20 times higher for EMCEE than it is for CMALA.

Similarly, the unconstrained Metropolis-adjusted Langevin algorithm (MALA) mixes ex-

tremely slowly; none of the walkers explore beyond the immediate vicinity of their starting points. The conditioning of the posterior requires MALA to use a step size 100 times smaller than the one that we use for CMALA to achieve a reasonable acceptance rate, making the ESS per step 2000 times lower for MALA than for CMALA and preventing convergence within the time that we allow (10^8 steps). By solving for the periodic solutions directly, the constrained Langevin sampler is agnostic to initial conditions, resulting in a much smoother posterior distribution that allows for a larger step size to be used. This behavior is well known in the context of trajectory optimization: shooting methods often have more difficulty converging than collocation methods due to the highly nonlinear dependence of the objective function on the control variables and initial conditions [10].

In contrast to ensemble MCMC and MALA, the no-U-turn sampler (NUTS) performs slightly worse than CMALA on the three-species repressilator and slightly better than CMALA on the seven-species repressilator. This can be attributed to the use of preconditioning in the no-U-turn sampler: whereas the constrained and unconstrained Langevin samplers use a unit mass matrix, the no-U-turn sampler uses the empirical covariance computed over a warmup period as the inverse mass matrix. We expect that the performance of the constrained Langevin sampler can be further improved if its dynamics are preconditioned as well.

Interestingly, we obtain a further 15 fold speedup when we do not adjust the constrained Langevin dynamics with the Metropolis criterion (the constrained unadjusted Langevin algorithm, CULA, in Table 2.1), at the expense of some bias due to the numerical error of discretizing the Langevin equations. We believe that this is due to the reversal of the momentum when steps are rejected in the MALA dynamics. The lower acceptance rate of MALA over ULA causes the MALA dynamics to retrace previously explored regions more often when the noise in the dynamics is small, as is the case for our calculations.

Because the constrained and unconstrained methods treat the initial conditions differ-

ently, their parameter distributions should not be exactly the same. Nevertheless, they contain similar information: most of the normalized importance weights (see Section 2.7.8) for reweighting the distribution sampled by the constrained Langevin dynamics to the distribution sampled by the unconstrained methods are close to one (and only a tiny fraction are larger than three, as shown in Figure 2.7, representing the unconstrained methods by ensemble MCMC). Given this similarity, the marginal distribution of $\mathbf{k}$ for the constrained Langevin sampler can be reweighted to that for the unconstrained methods or vice versa if desired.

2.6 Conclusions

Here, we introduced a constrained Langevin sampler that enables harvesting features of a dynamical system such as fixed points and limit cycles without integrating the dynamics for long times. The sampler is closely related to ones used previously in molecular simulations [2, 6, 82, 84, 134], but to the best of our knowledge, the present study is the first to apply it to parameter estimation.

To converge the parameter distribution for a model of a biochemical oscillator fit to time-series data, our constrained sampler needed orders of magnitude fewer likelihood evaluations than unconstrained samplers without preconditioning and comparable numbers of likelihood evaluations for appropriately preconditioned unconstrained samplers. This suggests the conditioning of posterior distributions associated with ODE parameter estimation problems can be improved by formulating the ODEs as constraints to be satisfied during sampling. In the specific case of oscillator models, the conditioning of posterior distributions is significantly improved by eliminating the dependence of the phase of the model output on model parameters, a known problem when sampling oscillator model parameters [69]. For simplicity, we did not tune the inverse mass matrix for the constrained dynamics; we note that existing methods for estimating the inverse mass matrix based on Euclidean distances

[85, 46] might not perform well without modification because the constraint manifold is nonlinear. That said, we expect that with appropriate preconditioning the constrained sampler can outperform unconstrained MCMC in general.

The main challenge in applying our method to new problems is in formulating and solving the constraint equations. For example, different defining systems for bifurcations [56], although formally equivalent, may exhibit different numerical behavior. We note that, when incorporating continuous constraints such as those arising from a boundary value problem, the choice of discretization and solution method can greatly affect the efficiency of the sampling procedure. If the mesh is too coarse, the solution to the boundary value problem may fail to converge and result in many rejected steps, whereas, if the mesh is too fine, the constraint equations may be computationally expensive to solve. In our implementation, the number of mesh points is fixed, but their spacing adapts to concentrate them where the solution has the highest curvature. We expect that the performance of the method can be further improved with more sophisticated methods [22, 112].

Our focus here is on ODEs, but the method can be extended to systems of partial differential equations (PDEs) by spatially discretizing the system of PDEs and then incorporating the resulting system of ODEs as constraints. As we have done with ODEs, the limiting behaviors of a system of PDEs can be obtained with much less computational effort than numerically integrating in time as is done in [21, 72]. The choice of spatial discretization, however, may have a large effect on the efficiency of the sampling scheme [30, 18, 106, 19].

Because the method does not rely on a specific form of the likelihood, it can be combined with enhanced sampling methods [34, 93, 101, 99, 113, 124] for further speedups. Recently, machine-learning methods have also been suggested for MCMC for parameter estimation [45, 132, 31]. In principle, the logarithm of a learned parameter distribution could be used as an additional potential term to bias the constrained Langevin dynamics. Thorough sampling of parameter distributions of complex models can open the door to systematic assessments

of models in biology and other fields where they have been limited to date.

2.7 Methods

2.7.1 Numerical integration of Hamilton's equations of motion with constraints

Following [84], the Hamiltonian can be split into terms corresponding to the potential and kinetic energies:

$$\mathcal{H}_A = \frac{1}{2}\mathbf{p}^\mathsf{T}\mathbf{M}^{-1}\mathbf{p} \tag{2.23a}$$

$$\mathcal{H}_B = U(\mathbf{q}), \tag{2.23b}$$

The equations of motion for a system with Hamiltonian $\mathcal{H}_A$ are

$$\begin{aligned} \dot{\mathbf{q}} &= \mathbf{M}^{-1}\mathbf{p} \\ \dot{\mathbf{p}} &= -\mathbf{c}_\mathbf{q}(\mathbf{q})^\mathsf{T}\boldsymbol{\lambda} \\ \mathbf{0} &= \mathbf{c}(\mathbf{q}), \end{aligned} \tag{2.24}$$

and the equations of motion for a system with Hamiltonian $\mathcal{H}_B$ are

$$\begin{aligned} \dot{\mathbf{q}} &= \mathbf{0} \\ \dot{\mathbf{p}} &= -U_\mathbf{q}(\mathbf{q}) - \mathbf{c}_\mathbf{q}(\mathbf{q})^\mathsf{T}\boldsymbol{\mu} \\ \mathbf{0} &= \mathbf{c}_\mathbf{q}(\mathbf{q})\mathbf{M}^{-1}\mathbf{p}, \end{aligned} \tag{2.25}$$

where $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ are the Lagrange multipliers appearing in the kinetic and potential terms, respectively. The dynamics for (2.24) and (2.25) are discretized separately. Discretizing

(2.24) gives

$$\mathbf{q}(t+\Delta t) = \mathbf{q}(t) + \Delta t \mathbf{M}^{-1}\mathbf{p}(t+\Delta t) \tag{2.26a}$$

$$\mathbf{p}(t+\Delta t) = \mathbf{p}(t) + \Delta t \mathbf{c}_{\mathbf{q}}(\mathbf{q}(t))^{\mathsf{T}}\boldsymbol{\lambda}(t+\Delta t) \tag{2.26b}$$

$$\mathbf{0} = \mathbf{c}(\mathbf{q}(t+\Delta t)), \tag{2.26c}$$

and discretizing (2.25) gives

$$\mathbf{q}(t+\Delta t) = \mathbf{q}(t) \tag{2.27a}$$

$$\mathbf{p}(t+\Delta t) = \mathbf{p}(t) - \Delta t \left(U_{\mathbf{q}}(\mathbf{q}(t)) + \mathbf{c}_{\mathbf{q}}(\mathbf{q}(t))^{\mathsf{T}}\boldsymbol{\mu}(t+\Delta t)\right) \tag{2.27b}$$

$$\mathbf{0} = \mathbf{c}_{\mathbf{q}}(\mathbf{q}(t))\mathbf{M}^{-1}\mathbf{p}(t+\Delta t). \tag{2.27c}$$

The momentum update (2.27) is an explicit system of linear equations that can be solved directly, whereas the position update (2.26) is a nonlinear system of equations that must be solved iteratively. The details of solving these equations are discussed in section 2.7.3.

At the start of each step, $\mathbf{q}(t)$ is assumed to satisfy the position constraints $\mathbf{c}(\mathbf{q}(t)) = \mathbf{0}$ and $\mathbf{p}(t)$ is assumed to satisfy the cotangency constraints $\mathbf{c}_{\mathbf{q}}(\mathbf{q}(t))\mathbf{M}^{-1}\mathbf{p}(t) = 0$. For a given timestep Δt, (2.26) and (2.27) define discrete maps $\Phi_{A,\Delta t}$ and $\Phi_{B,\Delta t}$ (Algorithms 2 and 3) that can be composed to approximate the continuous flow map (2.10). Repeatedly applying $\Phi_{B,\Delta t/2}\Phi_{A,\Delta t}\Phi_{B,\Delta t/2}$, for example, results in the WIGGLE integrator [82].

Algorithm 2 `constrained_position_step` ($\Phi_{A,\Delta t}$)

Input: position $\mathbf{q}_i$, momentum $\mathbf{p}_i$, step size Δt, inverse mass matrix $\mathbf{M}^{-1}$, constraint function $\mathbf{c}(\cdot)$

Output: new position $\mathbf{q}_{i+1}$, new momentum $\mathbf{p}_{i+1}$

1: $\boldsymbol{\lambda}_{i+1} \leftarrow$ `nonlinear_solve` $\left[\mathbf{c}\left(\mathbf{q}_i + \Delta t \mathbf{M}^{-1}\left(\mathbf{p}_i + \Delta t \boldsymbol{\lambda}_{i+1}^{\mathsf{T}}\mathbf{c}_{\mathbf{q}}(\mathbf{q}_i)\right)\right) = \mathbf{0}\right]$

2: $\mathbf{p}_{i+1} \leftarrow \mathbf{p}_i + \Delta t \mathbf{c}_{\mathbf{q}}(\mathbf{q}_i)^{\mathsf{T}}\boldsymbol{\lambda}_{i+1}$

3: $\mathbf{q}_{i+1} \leftarrow \mathbf{q}_i + \Delta t \mathbf{M}^{-1}\mathbf{p}_{i+1}$

4: **return** $\mathbf{q}_{i+1}, \mathbf{p}_{i+1}$

Algorithm 3 `constrained_momentum_step` ($\Phi_{B,\Delta t}$)

Input: position $\mathbf{q}_i$, momentum $\mathbf{p}_i$, step size Δt, inverse mass matrix $\mathbf{M}^{-1}$, constraint function $\mathbf{c}(\cdot)$, potential energy function $U(\cdot)$
Output: new position $\mathbf{q}_{i+1}$, new momentum $\mathbf{p}_{i+1}$

1: $\boldsymbol{\mu}_{i+1} \leftarrow \left(\mathbf{c}_\mathbf{q}(\mathbf{q}_i)\mathbf{M}^{-1}\mathbf{c}_\mathbf{q}(\mathbf{q}_i)^\mathsf{T}\right)^{-1}\mathbf{c}_\mathbf{q}(\mathbf{q}_i)\mathbf{M}^{-1}U_\mathbf{q}(\mathbf{q}_i)/\Delta t$
2: $\mathbf{p}_{i+1} \leftarrow \mathbf{p}_i + \Delta t\left(U_\mathbf{q}(\mathbf{q}_i) + \mathbf{c}_\mathbf{q}(\mathbf{q}_i)^\mathsf{T}\boldsymbol{\mu}_{i+1}\right)$
3: **return** $\mathbf{q}_i, \mathbf{p}_{i+1}$

2.7.2 Numerical integration of constrained Langevin equations

Just as we added noise and friction terms in (2.12), we can add a term corresponding to an Ornstein–Uhlenbeck process in $\mathbf{p}$ to (2.24) and (2.25) to make the dynamics stochastic:

$$\begin{aligned}\dot{\mathbf{q}} &= 0\\ \dot{\mathbf{p}} &= -\gamma\mathbf{p} + \sqrt{2T\gamma}\mathbf{M}^{1/2}\boldsymbol{\eta}(t) - \mathbf{c}_\mathbf{q}(\mathbf{q})^\mathsf{T}\boldsymbol{\mu}\\ \mathbf{0} &= \mathbf{c}_\mathbf{q}(\mathbf{q})\mathbf{M}^{-1}\mathbf{p}\end{aligned} \tag{2.28}$$

These equations can be integrated exactly (in the sense of averages) over a time interval Δt:

$$\mathbf{q}(t+\Delta t) = \mathbf{q}(t) \tag{2.29a}$$

$$\mathbf{p}(t+\Delta t) = a\mathbf{p}(t) + b\mathbf{M}^{1/2}\mathbf{r} + \Delta t\mathbf{c}_\mathbf{q}(\mathbf{q}(t))^\mathsf{T}\boldsymbol{\mu}(t+\Delta t) \tag{2.29b}$$

$$\mathbf{0} = \mathbf{c}_\mathbf{q}(\mathbf{q}(t))\mathbf{M}^{-1}\mathbf{p}(t+\Delta t), \tag{2.29c}$$

where $a \equiv \exp(-\gamma\Delta t)$, $b \equiv \sqrt{T(1-a^2)}$, and $\mathbf{r} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. (2.29) defines a stochastic map $\Phi_{O,\Delta t}$ (Algorithm 4).

The $\Phi_{A,\Delta t}$, $\Phi_{B,\Delta t}$, and $\Phi_{O,\Delta t}$ maps are combined in various ways in [84]. Among them, we choose to use the "OBABO" scheme which approximates (2.12) by the map $\Phi_{O,\Delta t/2}\Phi_{B,\Delta t/2}\ \Phi_{A,\Delta t}\Phi_{B,\Delta t/2}\Phi_{O,\Delta t/2}$ because it minimizes the number of applications of

the $\Phi_{A,\Delta t}$ map, which is the most computationally expensive owing to its iterative solution of the constraint equations.

Algorithm 4 `constrained_momentum_noise` ($\Phi_{O,\Delta t}$)

Input: position $\mathbf{q}_i$, momentum $\mathbf{p}_i$, step size Δt, inverse mass matrix $\mathbf{M}^{-1}$, constraint function $\mathbf{c}(\cdot)$, friction coefficient γ, temperature T
Output: new position $\mathbf{q}_{i+1}$, new momentum $\mathbf{p}_{i+1}$

1: $a \leftarrow \exp(-\gamma \Delta t)$ $b \leftarrow \sqrt{T(1-a^2)}$
2: $\mathbf{r} \sim \mathcal{N}(0, \mathbf{I})$
3: $\boldsymbol{\mu}_{i+1} \leftarrow \left(\mathbf{c}_{\mathbf{q}}(\mathbf{q}_i)\mathbf{M}^{-1}\mathbf{c}_{\mathbf{q}}(\mathbf{q}_i)^{\mathsf{T}}\right)^{-1} \mathbf{c}_{\mathbf{q}}(\mathbf{q}_i)\mathbf{M}^{-1/2}\mathbf{r}/\Delta t$
4: $\mathbf{p}_{i+1} \leftarrow a\mathbf{p}_i + b\mathbf{M}^{1/2}\mathbf{r} + \Delta t \mathbf{c}_{\mathbf{q}}(\mathbf{q}_i)\boldsymbol{\mu}_{i+1}$
5: **return** $\mathbf{q}_i$, $\mathbf{p}_{i+1}$

The error in the discrete dynamics introduces bias to the sampled distribution that can be controlled by reducing the timestep Δt. However, it may be preferable to use a large timestep to reduce the amount of computational effort required to generate independent samples. The bias due to discretization errors can be corrected by applying the Metropolis criterion [95]. A particular source of bias that we handle specially is the possibility of non-reversible changes in the position. For large timesteps, the position update (2.26) may not be reversible up to momentum reversal [86]; that is, the point obtained by performing a position update and then performing a second position update after reversing the resulting momentum may differ from the initial point. As the convergence of the samplers that we use relies on the reversibility of the dynamics, such errors introduce bias. We correct for non-reversibility by performing a second position update with reversed momentum after every position update and rejecting any steps for which the difference between the initial point and the endpoint of the reverse step is greater than a specified tolerance. One step of the sampling scheme is given by Algorithm 5.

Algorithm 5 `obabo_step`

Input :position $\mathbf{q}_i$, momentum $\mathbf{p}_i$, step size Δt, inverse mass matrix $\mathbf{M}^{-1}$, constraint function $\mathbf{c}(\cdot)$, potential energy function $U(\cdot)$, friction coefficient γ, temperature T, reversibility tolerance ϵ_{rev}

Output: new position $\mathbf{q}_{i+1}$, new momentum $\mathbf{p}_{i+1}$

$\mathbf{q}_{i+1}, \mathbf{p}_{i+1} \leftarrow$ `constrained_momentum_noise`$(\mathbf{q}_i, \mathbf{p}_i, \Delta t/2, \mathbf{M}^{-1}, \mathbf{c}, \gamma, T)$
$\mathbf{q}_{i+1}, \mathbf{p}_{i+1} \leftarrow$ `constrained_momentum_step`$(\mathbf{q}_i, \mathbf{p}_{i+1}, \Delta t/2, \mathbf{M}^{-1}, \mathbf{c})$
$\mathbf{q}_{i+1}, \mathbf{p}_{i+1} \leftarrow$ `constrained_position_step`$(\mathbf{q}_i, \mathbf{p}_{i+1}, \Delta t, \mathbf{M}^{-1}, \mathbf{c})$
$\tilde{\mathbf{q}}_i, \tilde{\mathbf{p}}_i, \leftarrow$ `constrained_position_step`$(\mathbf{q}_{i+1}, -\mathbf{p}_{i+1}, \Delta t, \mathbf{M}^{-1}, \mathbf{c})$
`accept` $\leftarrow \|\tilde{\mathbf{q}}_i - \mathbf{q}_i\| < \epsilon_{\text{rev}}$ ▷ reversibility check
if `accept` **then**
    $\mathbf{q}_{i+1}, \mathbf{p}_{i+1} \leftarrow$ `constrained_momentum_step`$(\mathbf{q}_{i+1}, \mathbf{p}_{i+1}, \Delta t/2, \mathbf{M}^{-1}, \mathbf{c})$
    $\mathbf{q}_{i+1}, \mathbf{p}_{i+1} \leftarrow$ `constrained_momentum_noise`$(\mathbf{q}_{i+1}, \mathbf{p}_{i+1}, \Delta t/2, \mathbf{M}^{-1}, \mathbf{c}, \gamma, T)$
    $H_i \leftarrow U(\mathbf{q}_i) + \mathbf{p}_i^{\mathsf{T}} \mathbf{M}^{-1} \mathbf{p}_i / 2$
    $H_{i+1} \leftarrow U(\mathbf{q}_{i+1}) + \mathbf{p}_{i+1}^{\mathsf{T}} \mathbf{M}^{-1} \mathbf{p}_{i+1} / 2$
    $r \sim U(0, 1)$
    `accept` $\leftarrow r < \exp(H_i - H_{i+1})$
end if
if not `accept` **then**
    $\mathbf{q}_{i+1}, \mathbf{p}_{i+1} \leftarrow \mathbf{q}_i, -\mathbf{p}_i$
end if
return $\mathbf{q}_{i+1}, \mathbf{p}_{i+1}$

2.7.3 Solving the constraint equations

Dense $\mathbf{c_q}$

In computing the momentum update, substituting the expression for $\mathbf{p}(t+\Delta t)$ from (2.27b) into (2.27c) gives

$$\mathbf{c_q}(\mathbf{q}(t))\mathbf{M}^{-1}\hat{\mathbf{p}}(t+\Delta t) - \Delta t\mathbf{c_q}(\mathbf{q}(t))\mathbf{M}^{-1}\mathbf{c_q}(\mathbf{q}(t))^{\mathsf{T}}\boldsymbol{\mu}(t+\Delta t) = \mathbf{0}, \tag{2.30}$$

where

$$\hat{\mathbf{p}}(t+\Delta t) = \mathbf{p}(t) - \Delta t U_{\mathbf{q}}(\mathbf{q}(t)) \tag{2.31}$$

is the momentum after an unconstrained step. Rearranging (2.30) gives a linear system for $\boldsymbol{\mu}(t+\Delta t)$:

$$\boldsymbol{\mu}(t+\Delta t) = \frac{1}{\Delta t}(\mathbf{c_q}(\mathbf{q}(t))\mathbf{M}^{-1}\mathbf{c_q}(\mathbf{q}(t))^{\mathsf{T}})^{-1}\mathbf{c_q}(\mathbf{q}(t))\mathbf{M}^{-1}\hat{\mathbf{p}}(t+\Delta t). \tag{2.32}$$

If $\mathbf{c_q}(\mathbf{q}(t))$ is dense, an efficient way to solve (2.32) is to compute the LQ factorization of $\mathbf{c_q}(\mathbf{q}(t))\mathbf{M}^{-1/2}$ (i.e., the QR factorization of $\left(\mathbf{M}^{-1/2}\right)^{\mathsf{T}}\mathbf{c_q}(\mathbf{q}(t))^{\mathsf{T}}$), such that $\mathbf{c_q}(\mathbf{q}(t))\mathbf{M}^{-1/2} = \mathbf{LQ}$, where $\mathbf{L}$ is a lower triangular matrix, $\mathbf{Q}$ is a matrix with orthonormal rows, and $\mathbf{M}^{-1/2}$ is a matrix that satisfies $\mathbf{M}^{-1/2}\left(\mathbf{M}^{-1/2}\right)^{\mathsf{T}} = \mathbf{M}^{-1}$ (e.g., the Cholesky factor of $\mathbf{M}^{-1}$). Assuming that $\mathbf{c_q}$ has full row rank everywhere, $\mathbf{p}(t+\Delta t)$ and $\boldsymbol{\mu}(t)$ in terms of $\mathbf{L}$ and $\mathbf{Q}$ are

given by

$$\begin{aligned}
\boldsymbol{\mu}(t+\Delta t) &= \frac{1}{\Delta t}(\mathbf{c}_\mathbf{q}(\mathbf{q}(t))\mathbf{M}^{-1}\mathbf{c}_\mathbf{q}(\mathbf{q}(t))^\mathsf{T})^{-1}\mathbf{c}_\mathbf{q}(\mathbf{q}(t))\mathbf{M}^{-1}\hat{\mathbf{p}}(t+\Delta t) \\
&= \frac{1}{\Delta t}(\mathbf{c}_\mathbf{q}(\mathbf{q}(t))\mathbf{M}^{-1}\mathbf{c}_\mathbf{q}(\mathbf{q}(t))^\mathsf{T})^{-1}\mathbf{c}_\mathbf{q}(\mathbf{q}(t))\mathbf{M}^{-1/2}\left(\mathbf{M}^{-1/2}\right)^\mathsf{T}\hat{\mathbf{p}}(t+\Delta t) \\
&= \frac{1}{\Delta t}\left(\mathbf{L}\mathbf{Q}\mathbf{Q}^\mathsf{T}\mathbf{L}^\mathsf{T}\right)^{-1}\mathbf{L}\mathbf{Q}\left(\mathbf{M}^{-1/2}\right)^\mathsf{T}\hat{\mathbf{p}}(t+\Delta t) \\
&= \frac{1}{\Delta t}\mathbf{L}^{-\mathsf{T}}\mathbf{Q}\left(\mathbf{M}^{-1/2}\right)^\mathsf{T}\hat{\mathbf{p}}(t+\Delta t) \qquad (2.33a) \\
\mathbf{p}(t+\Delta t) &= \hat{\mathbf{p}}(t+\Delta t) - \Delta t\mathbf{c}_\mathbf{q}(\mathbf{q}(t))^\mathsf{T}\boldsymbol{\mu}(t+\Delta t) \\
&= \hat{\mathbf{p}}(t+\Delta t) - \Delta t\left(\mathbf{M}^{1/2}\right)^\mathsf{T}\left(\mathbf{M}^{-1/2}\right)^\mathsf{T}\mathbf{c}_\mathbf{q}(\mathbf{q}(t))^\mathsf{T}\boldsymbol{\mu}(t+\Delta t) \\
&= \hat{\mathbf{p}}(t+\Delta t) + \left(\mathbf{M}^{1/2}\right)^\mathsf{T}\mathbf{Q}^\mathsf{T}\mathbf{L}^\mathsf{T}\mathbf{L}^{-\mathsf{T}}\mathbf{Q}\left(\mathbf{M}^{-1/2}\right)^\mathsf{T}\hat{\mathbf{p}}(t+\Delta t) \\
&= \hat{\mathbf{p}}(t+\Delta t) - \left(\mathbf{M}^{1/2}\right)^\mathsf{T}\mathbf{Q}^\mathsf{T}\mathbf{Q}\left(\mathbf{M}^{-1/2}\right)^\mathsf{T}\hat{\mathbf{p}}(t+\Delta t). \qquad (2.33b)
\end{aligned}$$

For the position update, substituting (2.26a) and (2.26b) into (2.26c) gives the nonlinear system of equations in terms of $\boldsymbol{\lambda}$:

$$\mathbf{c}(\mathbf{q}(t)) = \mathbf{c}(\hat{\mathbf{q}}(t+\Delta t) + \Delta t\mathbf{M}^{-1}\mathbf{c}_\mathbf{q}(\mathbf{q}(t))^\mathsf{T}\boldsymbol{\lambda}(t+\Delta t)) = \mathbf{0}, \qquad (2.34)$$

where

$$\hat{\mathbf{q}}(t+\Delta t) = \mathbf{q}(t) + \Delta t\mathbf{M}^{-1}\mathbf{p}(t) \qquad (2.35)$$

is an unconstrained position update. Differentiating (2.34) with respect to $\boldsymbol{\lambda}$ gives

$$\begin{aligned}
\mathbf{c}_{\boldsymbol{\lambda}}(\mathbf{q}(t+\Delta t)) &= \mathbf{c}_\mathbf{q}(\mathbf{q}(t+\Delta t))\mathbf{q}_{\boldsymbol{\lambda}}(t+\Delta t) \\
&= \Delta t\mathbf{c}_\mathbf{q}(\mathbf{q}(t+\Delta t))\mathbf{M}^{-1}\mathbf{c}_\mathbf{q}(\mathbf{q}(t))^\mathsf{T}.
\end{aligned} \qquad (2.36)$$

$\boldsymbol{\lambda}(t+\Delta t)$ and $\mathbf{q}(t+\Delta t)$ can then be computed using Newton's method. Given approximations $\boldsymbol{\lambda}^{(i)}(t+\Delta t)$ to $\boldsymbol{\lambda}(t+\Delta t)$ and $\mathbf{q}^{(i)}(t+\Delta t)$ to $\mathbf{q}(t+\Delta t)$, improved approximations can be

computed by iteratively solving

$$\begin{aligned}
\delta\boldsymbol{\lambda}^{(i)} &= \frac{1}{\Delta t}(\mathbf{c_q}(\mathbf{q}^{(i)}(t+\Delta t))\mathbf{M}^{-1}\mathbf{c_q}(\mathbf{q}(t))^\mathsf{T})^{-1}\mathbf{c}(\mathbf{q}^{(i)}(t+\Delta t)) \\
\boldsymbol{\lambda}^{(i+1)}(t+\Delta t) &= \boldsymbol{\lambda}^{(i)}(t+\Delta t) + \delta\boldsymbol{\lambda}^{(i)} \\
\mathbf{q}^{(i+1)}(t+\Delta t) &= \hat{\mathbf{q}}(t+\Delta t) + \Delta t^2\mathbf{M}^{-1}\mathbf{c_q}(\mathbf{q}(t))^\mathsf{T}\boldsymbol{\lambda}^{(i+1)}(t+\Delta t) \\
&\equiv \mathbf{q}^{(i)}(t+\Delta t) + \delta\mathbf{q}^{(i)}
\end{aligned} \tag{2.37}$$

However, computing and inverting the matrix $\mathbf{c_q}(\mathbf{q}(t+\Delta t))\mathbf{M}^{-1}\mathbf{c_q}(\mathbf{q}(t))^\mathsf{T}$ can be quite computationally costly, and depending on the conditioning of $\mathbf{c_q}$, may be numerically unstable. Consequently, we instead approximate $\mathbf{c_q}(\mathbf{q}(t+\Delta t))\mathbf{M}^{-1}\mathbf{c_q}(\mathbf{q}(t))^\mathsf{T}$ with $\mathbf{c_q}(\mathbf{q}(t))\mathbf{M}^{-1}\mathbf{c_q}(\mathbf{q}(t))^\mathsf{T}$ for small Δt as proposed in [6]. As $\mathbf{c_q}(\mathbf{q}(t))\mathbf{M}^{-1}\mathbf{c_q}(\mathbf{q}(t))^\mathsf{T}$ is symmetric and assumed to be positive definite, quasi-Newton iterates can be efficiently computed in terms of the LQ factorization of $\mathbf{c_q}(\mathbf{q}(t))\mathbf{M}^{-1/2}$:

$$\begin{aligned}
\delta\boldsymbol{\lambda}^{(i)} &= \frac{1}{\Delta t}\left(\mathbf{c_q}(\mathbf{q}(t))\mathbf{M}^{-1}\mathbf{c_q}(\mathbf{q}(t))^\mathsf{T}\right)^{-1}\mathbf{c}(\mathbf{q}^{(i)}(t+\Delta t)) \\
&= \frac{1}{\Delta t}\left(\mathbf{L}\mathbf{Q}\mathbf{Q}^\mathsf{T}\mathbf{L}^\mathsf{T}\right)^{-1}\mathbf{c}(\mathbf{q}^{(i)}(t+\Delta t)) \\
&= \frac{1}{\Delta t}(\mathbf{L}\mathbf{L}^\mathsf{T})^{-1}\mathbf{c}(\mathbf{q}^{(i)}(t+\Delta t)) \qquad (2.38a) \\
\delta\mathbf{q}^{(i)} &= \Delta t^2\mathbf{M}^{-1}\mathbf{c_q}(\mathbf{q}(t))^\mathsf{T}\delta\boldsymbol{\lambda}^{(i)} \\
&= \Delta t^2\mathbf{M}^{-1/2}\left(\mathbf{M}^{-1/2}\right)^\mathsf{T}\mathbf{c_q}(\mathbf{q}(t))^\mathsf{T}\delta\boldsymbol{\lambda}^{(i)} \\
&= \Delta t\mathbf{M}^{-1/2}\mathbf{Q}^\mathsf{T}\mathbf{L}^\mathsf{T}(\mathbf{L}\mathbf{L}^\mathsf{T})^{-1}\mathbf{c}(\mathbf{q}^{(i)}(t+\Delta t)) \\
&= \Delta t\mathbf{M}^{-1/2}\mathbf{Q}^\mathsf{T}\mathbf{L}^{-1}\mathbf{c}(\mathbf{q}^{(i)}(t+\Delta t)). \qquad (2.38b)
\end{aligned}$$

To improve the rate of convergence, Broyden updates can be applied to the factorization of $\mathbf{c_q}(\mathbf{q}(t))\mathbf{M}^{-1}\mathbf{c_q}(\mathbf{q}(t))^\mathsf{T}$. Our implementation computes Broyden updates using the QNERR algorithm described in [33].

Sparse $\mathbf{c_q}$

If $\mathbf{c}$ contains constraints arising from a boundary value problem discretized using a local basis, then $\mathbf{c_q}$ can be partitioned as

$$\mathbf{c_q} = \begin{pmatrix} \mathbf{c_y} & \mathbf{c_k} \end{pmatrix}, \tag{2.39}$$

where $\mathbf{c_y}$ is a sparse matrix containing derivatives with respect to the variables of the discretized boundary value problem, and $\mathbf{c_k}$ is a dense matrix containing derivatives with respect to parameters. Efficient solution of (2.27) would ideally exploit the sparsity of $\mathbf{c_q}\mathbf{M}^{-1/2}$, where we assume that $\mathbf{M}^{-1}$ is such that $\mathbf{c_q}\mathbf{M}^{-1/2}$ has the same structure as $\mathbf{c_q}$ (e.g., diagonal $\mathbf{M}^{-1}$), so that

$$\mathbf{c_q}\mathbf{M}^{-1/2} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \end{pmatrix} \tag{2.40}$$

with $\mathbf{A}$ sparse and $\mathbf{B}$ dense. In general, direct LQ factorization of $\mathbf{c_q}\mathbf{M}^{-1/2}$ results in a dense $\mathbf{Q}$, and since $\mathbf{c_q}$ contains dense columns $\mathbf{c_k}$, $\mathbf{L}$ is also dense, making the dense solution procedures (2.33) and (2.38) expensive for fine discretizations of the boundary value problem. However, if the LQ factorization of $\mathbf{c_y}$ can be computed efficiently, Algorithms 3 and 4 in [49] can be applied to update a sparse factorization of $\mathbf{c_y}$, allowing for efficient solution of the linear systems that arise in the position and momentum update steps.

If exact Newton iterations are used to compute the position update, the matrices

$$\begin{aligned} \mathbf{c_q}(\mathbf{q}^{(i)}(t+\Delta t))\mathbf{M}^{-1}\mathbf{c_q}(\mathbf{q}(t))^{\mathsf{T}} = \mathbf{A}(\mathbf{q}^{(i)}(t+\Delta t))\mathbf{A}(\mathbf{q}(t))^{\mathsf{T}} \\ + \mathbf{B}(\mathbf{q}^{(i)}(t+\Delta t))\mathbf{B}(\mathbf{q}(t))^{\mathsf{T}} \end{aligned} \tag{2.41}$$

must be inverted. Since $\mathbf{c_k}$ generally has many fewer columns than $\mathbf{c_y}$, the Woodbury matrix identity can in principle be applied, treating $\mathbf{B}(\mathbf{q}(t+\Delta t))\mathbf{B}(\mathbf{q}(t))^{\mathsf{T}}$ as a low rank update to $\mathbf{A}(\mathbf{q}(t+\Delta t))\mathbf{A}(\mathbf{q}(t))^{\mathsf{T}}$. However, $\mathbf{A}(\mathbf{q}(t+\Delta t))\mathbf{A}(\mathbf{q}(t))^{\mathsf{T}}$ tends to be ill-conditioned, which leads to poor convergence of the Newton iterations due to inaccurate solution of the linear

system. Alternatively, $\mathbf{c_q}(\mathbf{q}(t))\mathbf{M}^{-1}\mathbf{c_q}(\mathbf{q}(t))^\mathsf{T}$ can be used as a symmetric approximation to $\mathbf{c_q}(\mathbf{q}(t+\Delta t))\mathbf{M}^{-1}\mathbf{c_q}(\mathbf{q}(t))^\mathsf{T}$ for quasi-Newton solution of the position constraints as in the dense case. We use Algorithm 3 from [49] to invert $\mathbf{c_q}(\mathbf{q}(t))\mathbf{M}^{-1}\mathbf{c_q}(\mathbf{q}(t))^\mathsf{T}$ and use quasi-Newton iterations with Broyden updates to compute position updates as in the dense case. A similar procedure for solving large, sparse constraint equations is proposed in [134] where the Cholesky factorization of $\mathbf{c_q}(\mathbf{q}(t))\mathbf{M}^{-1}\mathbf{c_q}(\mathbf{q}(t))^\mathsf{T}$ is computed instead of the LQ factorization of $\mathbf{c_q}\mathbf{M}^{-1/2}$.

2.7.4 Discretizing periodic orbits

To compute solutions to (3.29) numerically, a suitable discretization must be applied. As only periodic solutions are of interest, a natural choice of discretization is to apply Galerkin's method with Fourier basis functions (i.e., Fourier collocation). This can be very efficient if the system of ODEs has smooth solutions. However, if the solutions to the system of ODEs or its derivatives undergo rapid variation, the number of Fourier basis functions required for an accurate approximation can be prohibitively large. Consequently, we instead discretize the system of ODEs using Gauss-Legendre collocation on piecewise polynomials [77]. Given a mesh

$$\{0 = s_0 < s_1 < \cdots < s_N = \tau\} \tag{2.42}$$

with interval widths

$$\Delta s_i = s_i - s_{i-1}, \qquad 1 \leq i \leq N \tag{2.43}$$

we approximate the solution to (3.29) on the space of piecewise polynomials

$$\mathcal{P}_{m,\Delta s} = \left\{\mathbf{u} \in C[0,\tau] \,\middle|\, \mathbf{u}|_{[s_{i-1},s_i]} \in \mathcal{P}_m\right\}. \tag{2.44}$$

That is, $\mathcal{P}_{m,\Delta s}$ consists of continuous functions $\mathbf{u}$ on the interval $[0, \tau]$ such that on each mesh subinterval $[s_{i-1}, s_i]$, $\mathbf{u}$ is a polynomial of up to order m. We demand that on each subinterval, $\mathbf{u}$ satisfies (2.13) at collocation points $\{z_{i,j}\}$

$$\mathbf{u}'(z_{i,j}) = \mathbf{f}(\mathbf{u}(z_{i,j}), \mathbf{k}), \quad j \in \{0, \ldots, m\}, \quad i \in \{0, \ldots, N-1\} \tag{2.45}$$

where $\mathbf{u}'(s)$ denotes the derivative of $\mathbf{u}$ with respect to s. We choose the collocation points to be the scaled and shifted roots of the m-th Legendre polynomial. Our implementation constructs $\mathbf{u}$ with Newton polynomials on each mesh subinterval.

Accurate solution of boundary value problems using a mesh-based discretization requires the mesh to be sufficiently fine at regions where the solution undergoes rapid variation. If the boundary value problem is discretized on a uniform mesh, a large number of mesh elements may be placed in relatively smooth regions of the solution that only require a few points to resolve. To save computational effort, adaptive mesh refinement procedures can be used to generate meshes that concentrate mesh points at rapidly varying regions of the solution without using a fine mesh globally. For applications such as finite element calculations, mesh refinement is typically performed by dividing mesh elements into smaller pieces in regions where an error estimate is high (h-adaptivity), which changes the total number of mesh points. Changing the size of the mesh, however, complicates reversibility guarantees. As an alternative, we adapt the mesh by redistributing the mesh elements after each step of the sampler while keeping the total number of mesh elements constant (r-adaptivity, moving mesh methods). Central to moving mesh methods is the equidistribution principle: for some strictly positive mesh density function $\rho(s)$ that is proportional to an error estimate of the solution and mesh points $\{s_0, \ldots, s_N\}$, the integral of the mesh density function should be equal on each mesh element of the optimal mesh

$$\int_{s_i}^{s_{i+1}} \rho(s) ds = \frac{1}{N} \int_{s_0}^{s_N} \rho(s) ds = \frac{\sigma}{N}. \tag{2.46}$$

The mesh density function we use is based on the curvature of $\mathbf{u}(t)$ [67]:

$$\rho(s) = \left(1 + \|\mathbf{u}''(s)\|^2\right)^{1/4} \tag{2.47}$$

where $\mathbf{u}''(s)$ is the second derivative of $\mathbf{u}$ with respect to s. We choose to solve the mesh equations (2.46) and the boundary value problem (3.29) simultaneously, leading to the system of equations

$$\begin{aligned} \mathbf{u}'(z_{i,j}) &= \mathbf{f}(\mathbf{u}(z_{i,j}), \mathbf{k}), \quad i \in \{0, \ldots, N-1\} \\ &\qquad\qquad\qquad\qquad\quad j \in \{0, \ldots, m\} \\ \mathbf{u}(0) &= \mathbf{u}(\tau) \\ \int_{s_i}^{s_{i+1}} \rho(s) ds &= \frac{\sigma}{N}, \qquad\qquad i \in \{0, \ldots, N-1\} \end{aligned} \tag{2.48}$$

which we solve using sparse Newton iteration.

2.7.5 Unconstrained Monte Carlo methods

In this section, we briefly describe the unconstrained MCMC methods with which we compare our method and refer readers to their original references for more complete discussions.

Metropolis-adjusted Langevin

One step of the algorithm we refer to as unconstrained Metropolis-adjusted Langevin consists of one timestep of discretized underamped Langevin (i.e., second-order Langevin) dynamics followed by application of the Metropolis criterion with negation of the momenta upon rejection. As in (2.23), the Hamiltonian $\mathcal{H}$ can be split into potential and kinetic parts for which the dynamics can be solved exactly. For an unconstrained system, the equations of

motion for the potential and kinetic parts, respectively, are

$$\begin{aligned} \dot{\mathbf{q}} &= \mathbf{M}^{-1}\mathbf{p} \\ \dot{\mathbf{p}} &= \mathbf{0}, \end{aligned} \tag{2.49}$$

and

$$\begin{aligned} \dot{\mathbf{q}} &= \mathbf{0} \\ \dot{\mathbf{p}} &= -U_{\mathbf{q}}(\mathbf{q}). \end{aligned} \tag{2.50}$$

Stochasticity can be introduced as an unconstrained Ornstein–Uhlenbeck process in the momenta

$$\begin{aligned} \dot{\mathbf{q}} &= 0 \\ \dot{\mathbf{p}} &= -\gamma\mathbf{p} + \sqrt{2T\gamma}\mathbf{M}^{1/2}\boldsymbol{\eta}(t), \end{aligned} \tag{2.51}$$

where γ is a friction coefficient. Equations eqs. (2.49) to (2.51) define unconstrained maps $\Phi_{A,\Delta t}$, $\Phi_{B,\Delta t}$, $\Phi_{O,\Delta t}$ that can be composed to numerically integrate the Langevin equations of motion. We use the composition $\Phi_{B,\Delta t/2}\Phi_{A,\Delta t/2}\Phi_{O,\Delta t}\Phi_{A,\Delta t/2}\Phi_{B,\Delta t/2}$ for the unconstrained Metropolis-adjusted Langevin sampler since it has been observed to produce good statistics for position variables [83]. Note that this method differs from the conventional Metropolis-adjusted Langevin algorithm by the inclusion of momentum variables. This method can be interpreted as generalized Hamiltonian Monte Carlo [66] with the trajectory length set to a single timestep.

The no-U-turn sampler

From a starting point $\mathbf{q}$, Hamiltonian Monte Carlo in its simplest form consists of sampling a normally distributed momentum vector $\mathbf{p}$, integrating the equations of motion for a fixed trajectory length L, and then accepting or rejecting the last point of the trajectory based on the Metropolis criterion. Since a single trajectory is approximately contained within a particular level set of the energy, certain choices of L can result in slow sampling due to

trajectories forming approximately closed orbits. Let $(\mathbf{q}_-, \mathbf{p}_-)$ and $(\mathbf{q}_+, \mathbf{p}_+)$ denote the position and momentum at the two endpoints of a trajectory. The no-U-turn sampler [64] adaptively sets the trajectory length by integrating trajectories until the criterion

$$\begin{aligned} \mathbf{p}_+^\mathsf{T}(\mathbf{q}_+ - \mathbf{q}_-) < 0 \\ \mathbf{p}_-^\mathsf{T}(\mathbf{q}_- - \mathbf{q}_+) < 0 \end{aligned} \tag{2.52}$$

is satisfied. When (2.7.5) is true, the momenta at the endpoints projected onto the line joining the endpoints of the trajectory point towards each other, and further integration is likely to decrease the distance between the endpoints of the trajectory. To preserve time reversibility, the implementation of NUTS [20] that we use iteratively doubles the length of the trajectory until (2.7.5) is satisfied or a maximum trajectory length is exceeded and then returns a point in the trajectory with probability proportional to its energy.

In addition to adaptively tuning L, the step size Δt is tuned using dual averaging [102] to achieve a specified acceptance rate, and the inverse mass matrix is approximated as the empirical covariance matrix computed over a warmup period [123].

Ensemble MCMC

The affine-invariant ensemble sampler that we use efficiently samples poorly scaled distributions by evolving an ensemble of parameter sets (random walkers) and using the spread in the ensemble to guide the directions of moves [55, 43]. Specifically, at each step, two walkers $\mathbf{x}_j$ and $\mathbf{x}_k$ are selected randomly, $\mathbf{x}_j$ is replaced with

$$\mathbf{x}'_j = \mathbf{x}_k + Z[\mathbf{x}_j - \mathbf{x}_k], \tag{2.53}$$

and the move is accepted with probability

$$\min[1, Z^{N-1}\pi(\mathbf{x}'_j)/\pi(\mathbf{x}_j)], \tag{2.54}$$

where N is the dimension of the parameter space, Z is a random variable with density

$$p(z) \propto \begin{cases} z^{-1/2} & \text{if } z \in \left[a^{-1}, a\right] \\ 0 & \text{otherwise,} \end{cases} \tag{2.55}$$

and a is an adjustable step size parameter. We use $a = 1.5$, which we found to give relatively good behavior in preliminary simulations.

2.7.6 Accounting for manifold curvature

The solution manifold to a system of constraints is generally curved, and projecting points sampled on the constraint manifold onto a limited set of variables results concentrates the points where the manifold is close to orthogonal to the coordinate axes. For example, projecting equispaced points on the unit circle onto the interval $[-1, 1]$ results in the Chebyshev nodes, which are concentrated near the endpoints -1 and 1, where the unit circle becomes increasingly orthogonal to the real line. If this concentration of points is undesired, one can correct for it by weighting each point by a factor that accounts for the change of the volume element of the manifold upon projection.

Given an m-dimensional submanifold $\mathcal{M}$ of $\mathbb{R}^n$ and a parameterization $\boldsymbol{\phi} : \mathbb{R}^m \to \mathcal{M}$, the volume element of a point $\boldsymbol{\phi}(\mathbf{x})$ on $\mathcal{M}$ is given by $\det(\boldsymbol{\phi}_\mathbf{x}^\mathsf{T}\boldsymbol{\phi}_\mathbf{x})^{1/2}d\mathbf{x}$, where $\boldsymbol{\phi}_\mathbf{x}$ denotes the matrix of derivatives with elements $\partial\boldsymbol{\phi}_i/\partial x_j$. The columns of $\boldsymbol{\phi}_\mathbf{x}(\mathbf{x})$ are a basis for the tangent space of $\mathcal{M}$ at $\mathbf{x}$. We compute the correction for the effect of manifold curvature by constructing an orthonormal basis $\mathbf{Q} \in \mathbb{R}^{n\times m}$ for the tangent space of $\mathcal{M}$ at each sampled point and then projecting each column of $\mathbf{Q}$ onto k coordinate axes of interest to obtain

$\hat{\mathbf{Q}} \in \mathbb{R}^{k \times m}$, where $k < m$. The factor by which volume is changed by the projection is given by $\det(\hat{\mathbf{Q}}\hat{\mathbf{Q}}^{\mathsf{T}})^{1/2}$, which can be used to weights points when constructing a histogram.

2.7.7 Assessing convergence

We assess the convergence of simulations using a version of the potential scale reduction factor ($\hat{R}$) defined in [48]. For a set of M independent chains each with N samples, we first compute the within-chain mean ($\bar{\mathbf{k}}_m$) and covariance ($\mathbf{\Sigma}_a$):

$$\bar{\mathbf{k}}_m = \frac{1}{N} \sum_{n=1}^{N} \mathbf{k}_{mn} \tag{2.56a}$$

$$\mathbf{\Sigma}_a = \frac{1}{M} \sum_{m=1}^{M} \mathbf{\Sigma}_m \tag{2.56b}$$

$$\mathbf{\Sigma}_m = \frac{1}{N-1} \sum_{n=1}^{N} \left(\mathbf{k}_{mn} - \bar{\mathbf{k}}_m\right) \left(\mathbf{k}_{mn} - \bar{\mathbf{k}}_m\right)^{\mathsf{T}}. \tag{2.56c}$$

Then we compute the between-chain covariance ($\mathbf{\Sigma}_b$):

$$\mathbf{\Sigma}_b = \frac{N}{M-1} \sum_{m=1}^{M} \left(\bar{\mathbf{k}}_m - \bar{\mathbf{k}}\right) \left(\bar{\mathbf{k}}_m - \bar{\mathbf{k}}\right)^{\mathsf{T}} \tag{2.57a}$$

$$\bar{\mathbf{k}} = \frac{1}{M} \sum_{m=1}^{M} \bar{\mathbf{k}}_m. \tag{2.57b}$$

Given the within-chain and between-chain covariances, we estimate the stationary covariance as

$$\mathbf{\Sigma} = \frac{N-1}{N} \mathbf{\Sigma}_a + \frac{1}{N} \mathbf{\Sigma}_b. \tag{2.58}$$

Then,

$$\hat{R} = \|\mathbf{\Sigma}_a^{-1} \mathbf{\Sigma}\|_2. \tag{2.59}$$

The essential idea is that upon convergence, the initial conditions should be forgotten, and the within-chain covariance should equal the stationary covariance; in this case, $\hat{R} = 1$.

2.7.8 Importance sampling

Given two absolutely continuous probability distributions π_1 and π_2 defined over the same sample space, the importance weight

$$w(x) = \frac{\pi_2(x)}{\pi_1(x)} \tag{2.60}$$

allows for integrals with respect to π_2 to be computed using π_1. For any function $f(x)$,

$$\mathbb{E}_{\pi_2}(f(x)) \equiv \int f(x)\pi_2(x)\,dx = \int f(x)\pi_1(x)w(x)\,dx \equiv \mathbb{E}_{\pi_1}(f(x)w(x)), \tag{2.61}$$

where $w(x)$ is the Radon–Nikodym derivative of π_2 with respect to π_1. Given samples $\{x_1, \ldots, x_N\}$ from π_1, an unbiased estimate of (2.61) is

$$\mathbb{E}_{\pi_2}(f(x)) \approx \frac{1}{N}\sum_{i=1}^{N} f(x_i)\frac{\pi_2(x_i)}{\pi_1(x_i)} \tag{2.62}$$

If π_1 and π_2 are only known up to a normalization constant, such that

$$\begin{aligned} \pi_1(x) &= c_1\hat{\pi}_1(x) \\ \pi_2(x) &= c_2\hat{\pi}_2(x), \end{aligned} \tag{2.63}$$

then the expectation with respect to π_2 can instead be computed as

$$\begin{aligned}\mathbb{E}_{\pi_2}(x) &= \int f(x)\pi_1(x)\frac{\pi_2(x)}{\pi_1(x)}\,dx \\ &= \int f(x)\pi_1(x)\frac{\pi_2(x)}{\pi_1(x)}\,dx \Big/ \int \pi_1(x)\frac{\pi_2(x)}{\pi_1(x)}\,dx \\ &= \int f(x)\pi_1(x)\frac{\hat{\pi}_2(x)}{\hat{\pi}_1(x)}\,dx \Big/ \int \pi_1(x)\frac{\hat{\pi}_2(x)}{\hat{\pi}_1(x)}\,dx \\ &\equiv \frac{\mathbb{E}_{\pi_1}(f(x)\hat{w}(x))}{\mathbb{E}_{\pi_1}(\hat{w}(x))}.\end{aligned} \tag{2.64}$$

Accordingly, we refer to

$$\frac{\hat{w}(x)}{\mathbb{E}_{\pi_1}(\hat{w}(x))} \tag{2.65}$$

as the normalized importance weights (cf. Figure 2.7). Given samples $\{x_1, \ldots, x_N\}$ from π_1, a biased estimate for (2.64) is

$$\frac{\mathbb{E}_{\pi_1}(f(x)\hat{w}(x))}{\mathbb{E}_{\pi_1}(\hat{w}(x))} = \left(\sum_{i=1}^{N} f(x_i)\hat{w}(x_i)\right)\left(\sum_{i=1}^{N} \hat{w}(x_i)\right)^{-1}. \tag{2.66}$$

Software

Open-source software implementing the method using the Google JAX package [13] is available at `https://github.com/dinner-group/constrained-Langevin`.

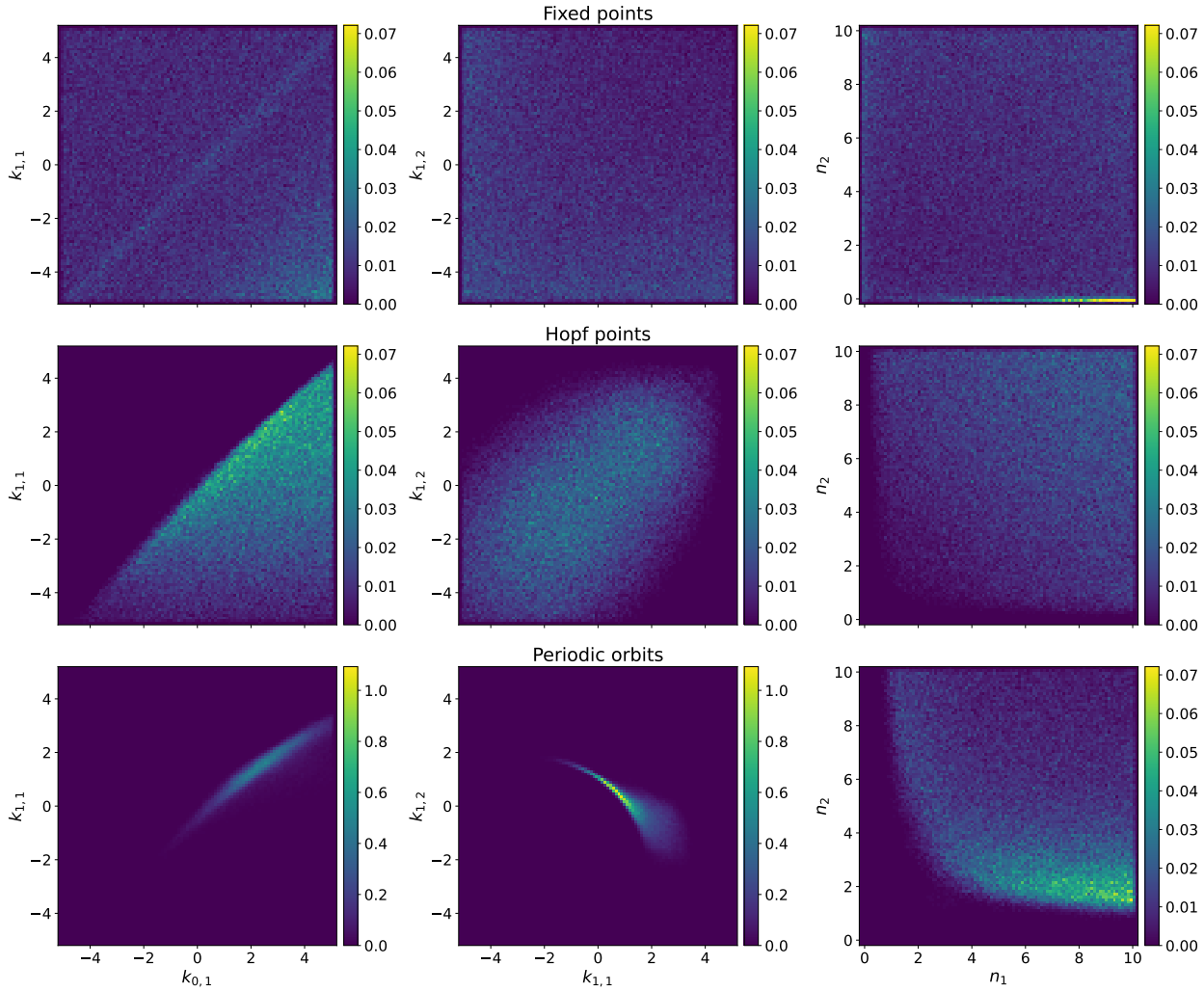


Figure 2.4: Representative two-dimensional marginals for the three-species represillator model parameter distribution obtained from 10^7 steps of the constrained Langevin sampler with fixed-point constraints (top), Hopf-point constraints (middle), and periodicity constraints (bottom). The log-likelihood when sampling fixed points and Hopf points is constant everywhere, and the log-likelihood when sampling limit cycles is proportional the squared error from the data shown in Figure 2.2 (see Section 2.5.4).

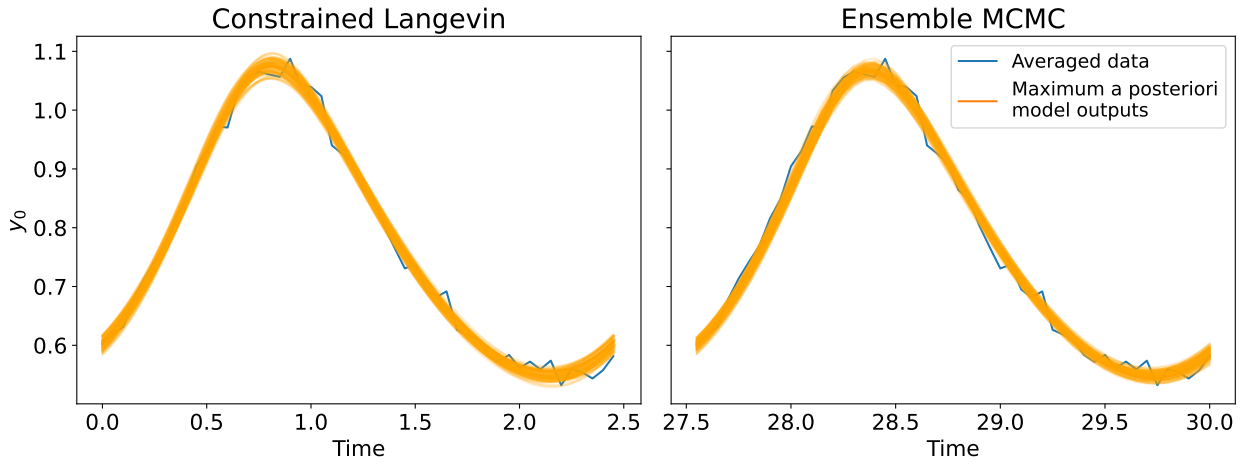


Figure 2.5: 100 best fit solutions obtained after 10^7 steps of the constrained MALA sampler (left) and 2×10^8 steps the ensemble MCMC sampler (right). Results shown are for the three-species repressilator fit to the data shown in Figure 2.2 (see Section 2.5.4).

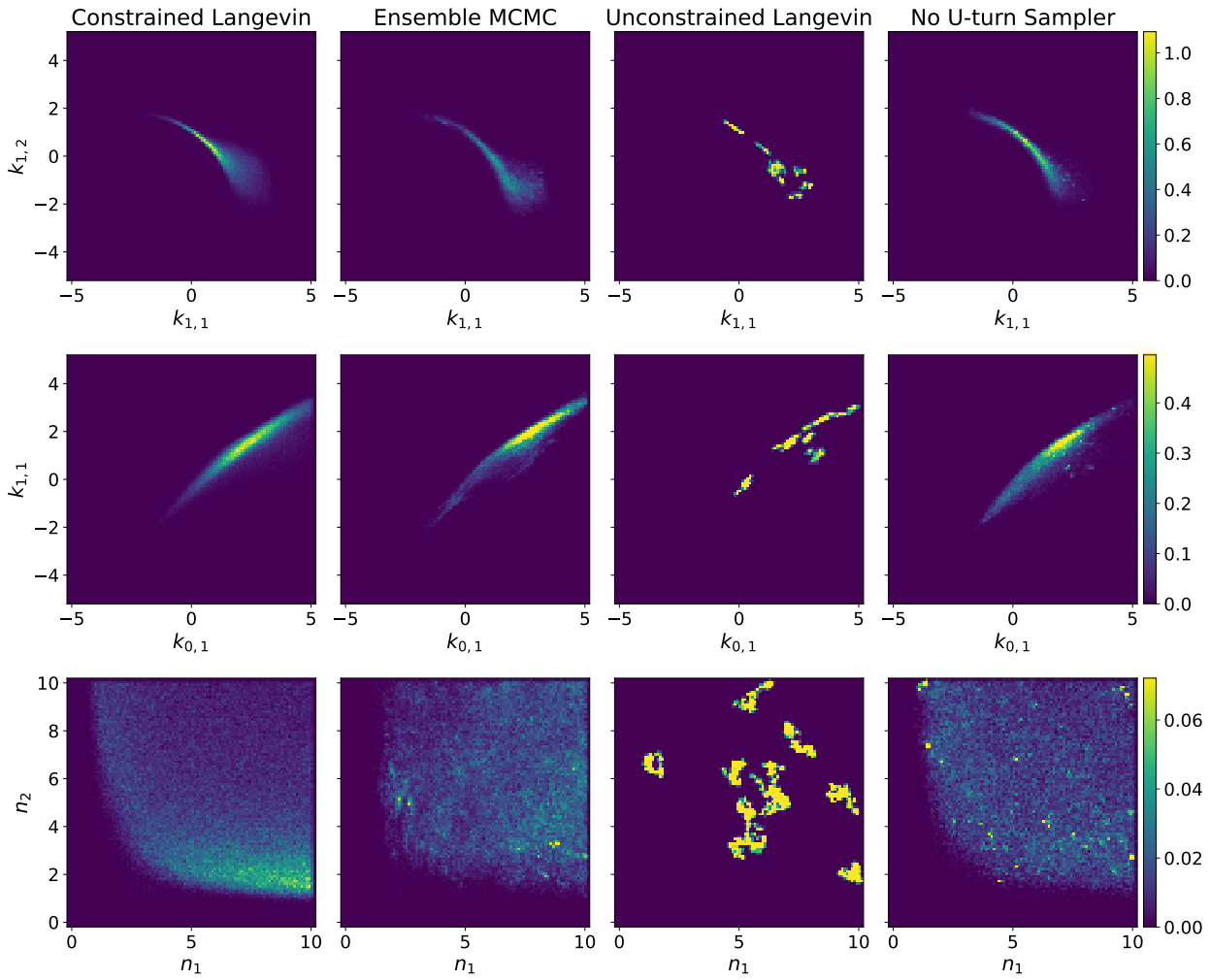


Figure 2.6: Representative marginal distributions sampled using constrained MALA, ensemble MCMC, unconstrained MALA, and NUTS after 10^7 steps. Results shown are for the three-species repressilator fit to the data shown in Figure 2.2 (see Section 2.5.4).

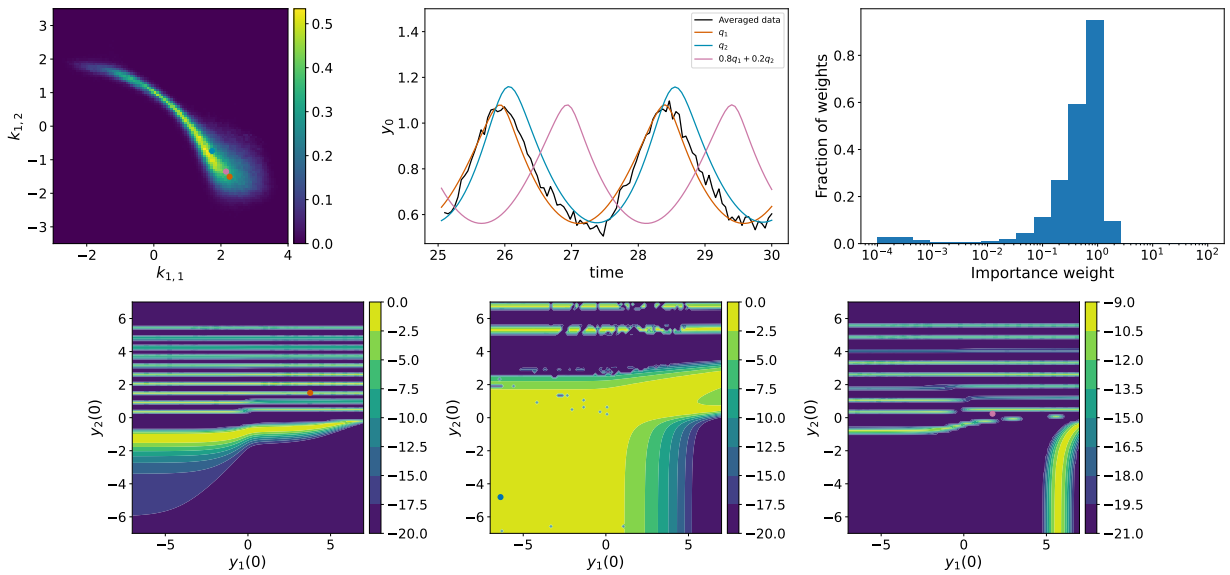


Figure 2.7: The distribution sampled by the unconstrained methods is multimodal despite the marginals along certain dimensions appearing smooth (top left). The multimodal structure is apparent when examining the distribution at different initial values of y_1 and y_2: the unnormalized log-posterior distribution exhibits rapid, non-concave variations with the rest of the parameters fixed at the high probability values indicated by the symbols (q_1, orange; q_2, blue; the linear interpolation $0.8q_1 + 0.2q_2$; pink in the bottom row). The interpolated parameters generate trajectories that are close in period and shape to the trajectories generated by q_1 and q_2 but about half a period out of phase with the data used to compute the log-likelihood (top middle). The apparent noise in the bottom middle plot results from failure to integrate the ODEs for certain initial conditions. Most of the normalized importance weights for reweighting the distribution sampled by constrained Langevin to the distribution sampled by the unconstrained methods are close to one (top right).

CHAPTER 3

T432 STABILIZES THE PERIOD OF THE KAI OSCILLATOR BY ENABLING DOUBLY PHOSPHORYLATED KAIC TO ACT AS AN ALTERNATIVE SINK FOR KAIA

3.1 Abstract

Circadian clocks maintain a nearly constant period despite variation in the concentrations of their components. The core oscillator of the cyanobacterial circadian clock can be reconstituted in vitro from the purified proteins KaiA, KaiB, and KaiC, making it a tractable system for studying this robustness. Models of the Kai system are generally parameterized by hand, so their failure to generalize to other conditions could reflect either the mechanisms that they encode or the parameters chosen for them. Here, we construct a kinetic model that allows for all elementary reactions among the phosphorylation, nucleotide, and KaiA-binding states of KaiC and fit it to phosphorylation kinetics measured across a range of total KaiA concentrations and solution ATP fractions by constrained Langevin Monte Carlo sampling of periodic solutions. Systematic reduction of the sampled models yields networks that oscillate without phosphorylation at T432, contrary to the conventional sequential mechanism but consistent with the oscillation of the T432V mutant. Because these networks oscillate over a narrower range of KaiA concentrations, we predict and confirm experimentally that the T432V clock loses rhythmicity at KaiA concentrations that the wild-type clock tolerates. Asymptotic analysis of a minimal model indicates that the ability of KaiA to inhibit its own sequestration limits the range of KaiA concentrations over which the period is stable, and that doubly phosphorylated KaiC offsets this limitation by acting as an additional sink. More generally, our results show how sampling and systematically reducing models consistent with data from multiple conditions can reveal and dissect biochemical mechanisms.

3.2 Introduction

Circadian clocks are endogenous biochemical oscillators with near-24-h periods that organisms in all kingdoms of life use to time their physiology to the daily solar cycle. The mechanisms underlying these rhythms are generally complex, involving transcription, translation, molecular association, phosphorylation, and degradation. The clock of the cyanobacterium *Synechococcus elongatus* is unusual in that its core oscillator, composed of the proteins KaiA, KaiB, and KaiC, involves only molecular association and phosphorylation of KaiC, and the phosphorylation state of KaiC encodes the time of day. Remarkably, free-running phosphorylation oscillations that persist for over a week with a period of about 22.4 h can be reconstituted in vitro from the purified proteins together with adenosine triphosphate (ATP) [117, 71]. This in vitro system enables quantitative data to be collected under well-controlled conditions and molecularly detailed mathematical models to be formulated, making the Kai oscillator a uniquely tractable system for developing a mechanistic understanding of molecular timekeeping.

KaiC is a hexameric ATPase that autophosphorylates and dephosphorylates at two sites, S431 and T432, in its C-terminal CII domain. Based on the relative phases of site-specific phosphorylation oscillations and mutation studies, the established mechanism is that, starting from the unphosphorylated state, KaiC phosphorylates preferentially on T432, which promotes subsequent phosphorylation on S431 [103, 111]. Phosphorylation of S431 increases the rate of ATP hydrolysis in the N-terminal CI domain, which drives a conformational change that allows KaiB to bind on CI. KaiC in complex with KaiB sequesters KaiA [117]; because KaiA acts as the nucleotide-exchange factor that drives phosphorylation, its sequestration forces S431 and T432 to dephosphorylate, completing the cycle. This KaiB-mediated sequestration constitutes a delayed negative feedback loop that is thought to be the primary driver of oscillation.

The tractability of the Kai system has motivated many mathematical models [80, 130,

111, 60, 89, 107, 108, 65]. Early models were largely phenomenological, constructed to reproduce specific observations such as robustness of the period to ATP levels [111], to protein stoichiometry [89], and to temperature [60]; later models sought to represent the underlying physical interactions more closely [108, 107]. Although existing models offer plausible explanations for the properties they were designed to reproduce, none reproduces behaviors outside the regime for which it was optimized. Whether this failure to generalize reflects structural shortcomings of the models or insufficient optimization of their parameters is unclear, because models parameterized by hand give little indication of the range of parameters consistent with the dynamics.

In principle, these issues can be addressed by constructing a model that allows for all possible mechanisms and fitting it to the data within a Bayesian framework. In this framework, one assigns a prior probability to the parameters and a likelihood for the data given the parameters; the posterior probability of the parameters is then proportional to the product of these probabilities. By then sampling parameters, one can not only maximize the posterior but also learn its distribution to quantify uncertainties. Hong and co-workers previously applied this approach to the KaiA-KaiC subsystem, fitting a molecularly detailed but mechanistically naive kinetic model directly to phosphorylation data by Markov chain Monte Carlo (MCMC) sampling [65]. Even though many microscopic parameters remained loosely constrained, the ensemble of fits made a consistent prediction: an ultrasensitive dependence of KaiC phosphorylation on KaiA concentration, which they confirmed experimentally and attributed to the differential affinity of KaiA for competing nucleotide-bound states of KaiC. This result illustrates how data-driven modeling can reveal a mechanism that was not built into a model a priori.

Extending this approach from the KaiA-KaiC subsystem to the full oscillator is complicated by the oscillatory dynamics themselves. Only a small fraction of parameter space gives rise to sustained oscillations, so conventional Monte Carlo methods expend most of

their effort sampling irrelevant, non-oscillatory regions of the posterior unless the prior is carefully tuned. Moreover, oscillations can appear or disappear through bifurcations, which can make the posterior poorly conditioned and hard to sample [55]. To overcome these difficulties, we adapted constrained Langevin Monte Carlo for sampling oscillatory solutions of ordinary differential equations (ODEs) [28]. This approach enables us to scan parameter space rapidly for regions likely to support oscillations and then to enforce periodicity of the ODE solutions within those regions.

Here, we introduce a parameterized kinetic model of the Kai oscillator (Figure 3.1) and use the constrained Langevin Monte Carlo approach to fit it to KaiC phosphorylation data collected across a range of total KaiA concentrations and solution ATP fractions. Surprisingly, the models that predominate in the sampling reduce to ones that do not require the T432 phosphorylation site to sustain oscillations, consistent with the observation that the T432V mutant oscillates [44]. However, these reduced models are less tolerant of variation in the total KaiA concentration, which leads us to predict that the T432V mutant should be more sensitive to KaiA than wild-type KaiC. We confirm this prediction experimentally. Through an asymptotic analysis of a minimal model obtained by reducing the sampled models, we interpret how the T432 site confers this robustness. We find that the key is the previously known but underappreciated ability of KaiA to inhibit its own sequestration [89].

3.3 Results

3.3.1 A monomeric model of the Kai oscillator fits data over a range of conditions.

We construct a parameterized model of the Kai oscillator by enumerating the states of a KaiC monomer and the reactions connecting them (Figure 3.1). To avoid the combinatorial explosion of states that would result from modeling the full KaiC hexamer, we treat KaiC as

a six independent monomers. For additional computational savings, we assume that KaiB remains at quasi-steady-state and treat it implicitly. We track three properties of each KaiC monomer: its phosphoform, the nucleotide bound at each of its two ATPase domains (CI and CII, both of which we take to be always occupied by ADP or ATP), and the KaiA occupancy of each domain.

Each KaiC monomer carries two phosphorylation sites, S431 and T432, and thus adopts one of four phosphoforms: U (unphosphorylated), T (phosphorylated on T432), S (phosphorylated on S431), and D (doubly phosphorylated). These form the four groups of states in Figure 3.1. Within each group, a superscript gives the state of CI and a subscript the state of CII; in either index, an A in the first position marks KaiA bound to that domain, and a 1 in the second position marks ADP, rather than ATP, at its nucleotide-binding site. The subscript therefore ranges over four CII states, one for each combination of KaiA occupancy and bound nucleotide.

Four reaction types connect the states (Figure 3.1 and Table 3.1): phosphotransfer, KaiA (un)binding, nucleotide exchange, and ATP hydrolysis. We require reversible cycles of these reactions to obey detailed balance (Table 3.2) but otherwise impose only upper and lower bounds on each rate constant to keep the model maximally general. Each time KaiA binds a CII domain, we remove one-sixth of a KaiA dimer from the free pool to account for the hexameric stoichiometry. KaiA bound to CII can catalyze nucleotide exchange. Phosphorylation at S431, present in the S and D phosphoforms, enables ATP hydrolysis at CI; because KaiB is not modeled explicitly, we fold its binding into this hydrolysis step and treat the composite reaction as pseudo-first-order. After hydrolysis, CI can reversibly bind KaiA to sequester it. The full network defines a system of ordinary differential equations (ODEs) for the time-dependent concentrations of all KaiC and KaiA species, coupled through mass-action kinetics.

We fit this model to data using a variant of constrained Langevin Monte Carlo for sam-

Table 3.1: A list of all reactions between KaiC species in the model. Reactants and products are indexed with a binary string notation in which a KaiC monomer is specified by six binary variables, listed in the key at the top of the table: S431 phosphorylated (pS431), T432 phosphorylated (pT432), KaiA bound to CII (CII-KaiA), ADP bound to CII (CII-ADP), KaiA bound to CI (CI-KaiA), and ADP bound to CI (CI-ADP). Each variable is equal to 1 if the corresponding condition is satisfied and 0 otherwise; when ADP is not bound, ATP is bound (i.e., we do not allow for unoccupied nucleotide binding sites). A $\cdot$ is a wildcard that accommodates either 0 or 1, so long as the full six-index combination is among the allowed species. The rate constants k_a distinguish the nucleotide exchange reactions, which can only occur if KaiA is bound on CII, and the CII ATP hydrolysis reactions, which can occur regardless of the presence of KaiA.

Binary string notation key: (pS431, pT432, CII–KaiA, CII–ADP, CI–KaiA, CI–ADP)		
Reaction	Description	Mass action form
$(0\cdot\cdot0\,0\,0) \longrightarrow (1\cdot\cdot1\,0\,0)$	S431 phosphorylation	$ky[0\cdot\cdot0\,0\,0]$
$(1\cdot\cdot1\,0\,0) \longrightarrow (0\cdot\cdot0\,0\,0)$	S431 dephosphorylation	$ky[1\cdot\cdot1\,0\,0]$
$(\cdot\,0\cdot0\cdot\cdot) \longrightarrow (\cdot\,1\cdot1\cdot\cdot)$	T432 phosphorylation	$ky[\cdot\,0\cdot0\cdot\cdot]$
$(\cdot\,1\cdot1\cdot\cdot) \longrightarrow (\cdot\,0\cdot0\cdot\cdot)$	T432 dephosphorylation	$ky[\cdot\,1\cdot1\cdot\cdot]$
$(\cdot\cdot0\cdot\cdot\cdot) \longrightarrow (\cdot\cdot1\cdot\cdot\cdot)$	CII KaiA binding	$kay[\cdot\cdot0\cdot\cdot\cdot]$
$(\cdot\cdot1\cdot\cdot\cdot) \longrightarrow (\cdot\cdot0\cdot\cdot\cdot)$	CII KaiA unbinding	$ky[\cdot\cdot1\cdot\cdot\cdot]$
$(\cdot\cdot1\,0\cdot\cdot) \longrightarrow (\cdot\cdot1\,1\cdot\cdot)$	CII ADP binding	$k_ay[\cdot\cdot1\,0\cdot\cdot]$
$(\cdot\cdot1\,1\cdot\cdot) \longrightarrow (\cdot\cdot1\,0\cdot\cdot)$	CII ATP binding	$k_ay[\cdot\cdot1\,1\cdot\cdot]$
$(\cdot\cdot\cdot0\cdot\cdot) \longrightarrow (\cdot\cdot\cdot1\cdot\cdot)$	CII ATP hydrolysis	$ky[\cdot\cdot\cdot0\cdot\cdot]$
$(1\cdot\cdot\cdot0\,0) \longrightarrow (1\cdot\cdot\cdot0\,1)$	CI ATP hydrolysis	$ky[1\cdot\cdot\cdot0\,0]$
$(1\cdot\cdot\cdot0\,1) \longrightarrow (1\cdot\cdot\cdot0\,0)$	CI ATP binding	$ky[1\cdot\cdot\cdot0\,1]$
$(1\cdot\cdot\cdot0\,1) \longrightarrow (1\cdot\cdot\cdot1\,1)$	CI KaiA sequestration	$kay[1\cdot\cdot\cdot0\,1]$
$(1\cdot\cdot\cdot1\,1) \longrightarrow (1\cdot\cdot\cdot0\,1)$	CI KaiA desequestration	$ky[1\cdot\cdot\cdot1\,1]$

Table 3.2: Cycles of reactions that obey detailed balance. The binary string notation is the same as in Table 3.1. In addition to the reactions shown, there is a pair of reversible reactions between the first and last species in each sequence to close the cycle. Each cycle introduces a constraint that the product of the reaction constants in the forward direction must equal the product of the reaction constants in the reverse direction.

Cycle	Description
$(0\cdot0\,0\,0\,0) \leftrightharpoons (1\cdot0\,1\,0\,0) \leftrightharpoons (1\cdot1\,1\,0\,0) \leftrightharpoons (0\cdot1\,0\,0\,0)$	pS–CII KaiA binding
$(\cdot\,0\,0\,0\cdot\cdot) \leftrightharpoons (\cdot\,1\,0\,1\cdot\cdot) \leftrightharpoons (\cdot\,1\,1\,1\cdot\cdot) \leftrightharpoons (\cdot\,0\,1\,0\cdot\cdot)$	pT–CII KaiA binding
$(1\cdot0\cdot0\,1) \leftrightharpoons (1\cdot1\cdot0\,1) \leftrightharpoons (1\cdot1\cdot1\,1) \leftrightharpoons (1\cdot0\cdot1\,1)$	CI and CII KaiA binding

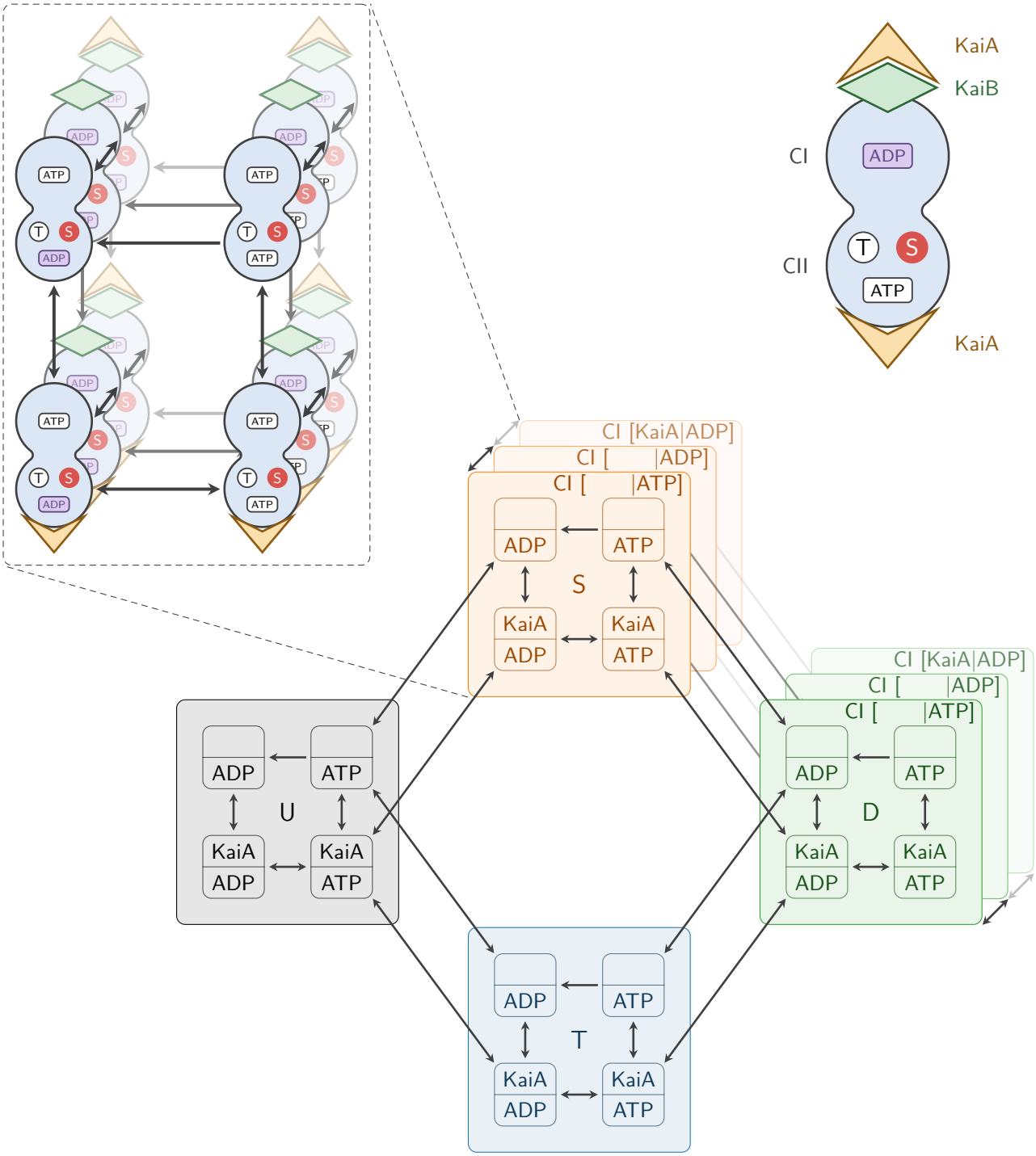

KaiA
KaiB
CI
ADP
T
S
CII
ATP
KaiA
CI [KaiA|ADP]
CI [|ADP]
CI [|ATP]
ADP
ATP
S
KaiA
ADP
KaiA
ATP
U
D
T

Figure 3.1

A diagram of the states of KaiC and their reactions in our monomeric model of the Kai oscillator. Panels labeled with U, T, S, or D represent KaiC monomers with CII domain that is unphosphorylated, phosphorylated on T432 only, phosphorylated on S431 only, or phosphorylated on both sites, respectively. The cells in each panel represent the KaiA and nucleotide binding states of the CII domain; an empty upper half cell denotes CII without KaiA bound and an upper half cell labeled with KaiA denotes KaiA bound on CII. We assume that all nucleotide binding sites are always occupied by either ATP or ADP. KaiC with S431 phosphorylated can hydrolyze ATP on its CI domain and then bind KaiB to enter a sequestration competent state. We model the combined effect of CI ATP hydrolysis and KaiB binding as a pseudo first order reaction. The KaiA and nucleotide binding states of the CI domain are labeled in the upper right corner of the S and D panels. These sequestration competent states are represented by the two panels behind the main S and D panels. The panels in the background contain the same CII reactions as those in the foreground. Each state in the main S and D panels are joined by a pair of reactions to the corresponding state in the panel behind. A cartoon of KaiC phosphorylated on S431 only with KaiA bound on CII and KaiA sequestered on CI is shown on the upper right. The reactions within the entire S block of the diagram is expanded on the upper left with cartoons depicting the states of KaiC involved. KaiA binding reactions follow second order kinetics.

pling parameters of models of oscillatory dynamics [28]. We first sample the parameter space for fixed points of the dynamics with purely imaginary eigenvalues (i.e., Hopf points). This serves as a computationally inexpensive way to identify parts of parameter space where oscillatory dynamics are likely. We then use these Hopf points to initiate sampling of limit cycle solutions to the ODEs. To enforce periodicity of solutions, we discretize and incorporate a periodic boundary value problem (BVP) for the model dynamics as constraints to be solved at every sampling step. We previously found this approach to have better performance for sampling oscillatory dynamics compared to conventional MCMC methods such as Hamiltonian Monte Carlo due to better conditioning of the resulting posterior distributions.

The data consist of phosphorylation kinetics measured under four conditions [111], each chosen to isolate a complementary part of the reaction network (Table 3.3): the full clock reaction at 100% and 25% ATP, the KaiA–KaiC subsystem, and KaiC alone. The two subsystem conditions constrain the phosphorylation and dephosphorylation kinetics in the absence of KaiB-mediated feedback; the two full-clock conditions constrain the complete

oscillator. We fit these data by subtracting a term to from log likelihood for each condition, proportional to the squared error between the predicted and measured phosphoform kinetics. To reduce computational effort, we partition each of the full-clock time series that spans multiple periods into single-period segments, average the phosphoform levels across segments, solve the ODEs over one period, and fit the resulting solution to the period-averaged data.

Sustained oscillations have also been observed at higher KaiA concentrations [26]. Rather than fit these time series directly, we use them to enforce robustness of the period to variations in KaiA concentration and %ATP: we compute the oscillatory solution at each oscillatory condition in Table 3.3 during the sampling and subtract a penalty from the log posterior proportional to the variance of the period across these solutions.

Table 3.3: Conditions used to constrain the model. KaiC is at 3.5 μM throughout, and KaiB, where present, at 3.5 μM. A check in the Fit column indicates that the measured phosphoform kinetics enter the least-squares term of the log posterior; a check in the Oscillatory column indicates that the system is oscillatory and that the oscillation period enters a variance penalty on the period. Nucleotide exchange requires KaiA, so %ATP does not enter the model for KaiC alone (dash). Sustained oscillations at the higher KaiA concentrations were reported by [26].

Proteins	KaiA dimer (μM)	%ATP	Fit	Oscillatory
KaiA, KaiB, KaiC	1.8	100		✓
KaiA, KaiB, KaiC	1.2	100		✓
KaiA, KaiB, KaiC	0.6	100	✓	✓
KaiA, KaiB, KaiC	0.6	25	✓	✓
KaiA, KaiC	0.6	100	✓	
KaiC	0.0	–	✓	

Figure 3.2 compares solutions from the 1000 sampled parameter sets with the highest log posterior values to the period-averaged experimental data (thinned by 10 for clarity). In general, it is easy to find parameters that can reproduce the nonoscillatory autophosphorylation and dephosphorylation dynamics of KaiC. If the model is required to oscillate at only one of the experimental conditions (i.e., at 100% ATP or 25% ATP), then the sampled parameters result in solutions that agree with the data collected at the one experimental condition to high precision. However, demanding oscillatory solutions at both conditions

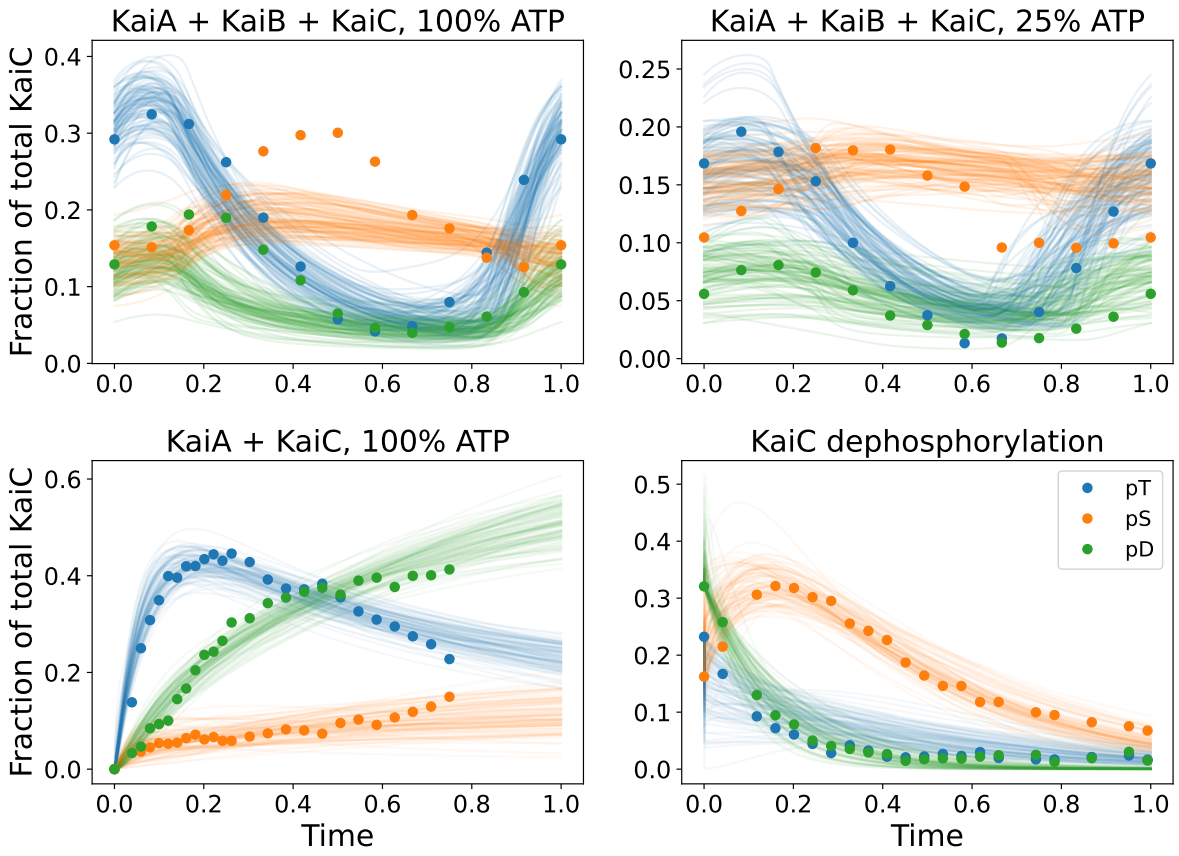


Figure 3.2: Experimentally measured phosphorylation dynamics of the Kai oscillator (symbols) and 100 optimized model outputs (thin lines) for different conditions. For the oscillatory conditions shown with total KaiA and ATP fractions of $(0.6, 1.0)$ and $(0.6, 0.25)$, we process the time-dependent phosphorylation data reported in [111] by dividing the experimental data into segments of length equal to the experimentally determined period and averaging over all segments.

causes the quality of fits to deteriorate. In particular, the amplitude of oscillations in the phosphoserine KaiC fraction is consistently lower than observed experimentally (Figure 3.2). That said, we are not aware of any other model that maintains near stable periods across the conditions considered (Table 3.3).

To further validate the model, we test its ability to reproduce properties of the clock that were not considered when fitting the data (Figure 3.3). The experiments that we consider are ones in which the oscillator is subjected to changes in the ATP/ADP ratio (reported as percentage ATP), which the system senses and, in the cell, correlates with light

intensity [116]. With the parameters in Figure 3.2, the model entrains to periodic variations in the ATP fraction over a range of driving periods within about three cycles (upper left), consistent with experiments [88]. The model reproduces the phase response to a step down in solution ATP from 100% to 25% over a duration of $0.2\,\tau$ nearly quantitatively (lower left) and the phase response to a 16% step up in total KaiA concentration qualitatively (lower right). Finally, we consider experiments in which the fraction of the period at 100% ATP is varied, which mimics seasonal changes in day length. For selected periods, the peak in phosphorylation occurs close to the midpoint of the interval with 100% ATP (midday), consistent with experiments [88], but for other periods, it is closer to the start of the interval with 100% ATP (dawn; upper right). The latter behavior is inconsistent with data for the in vivo oscillator [88]; data for the in vitro oscillator are not available. Finally, we simulate an experiment in which samples with different phases are mixed in a 1:1 ratio [71]; the results of this experiment are also well reproduced by our optimized model (Figure 3.4).

Taken together, these results indicate that the model generalizes reasonably well to conditions beyond the training data, giving us confidence that a closer dissection of its behavior can offer meaningful insight into the underlying mechanism.

3.3.2 Model reduction reveals that T432 is dispensable for oscillations but required for robustness to variations in KaiA concentration.

The data constrain the rate constants of the model to widely varying degrees. Sampling yields an ensemble of parameter sets. Figure 3.5) shows marginal posterior densities obtained by projecting the ensemble obtained from sampling unstable fixed points as a surrogate for oscillatory dynamics (Section 3.5.4) onto one reaction constant at a time, which we take as a measure of how strongly the data support the associated reaction. Nearly all of these densities take one of four shapes: a single peak at an intermediate value; a maximum at the lower end of the sampled range; a maximum at the upper end; or a nearly flat profile

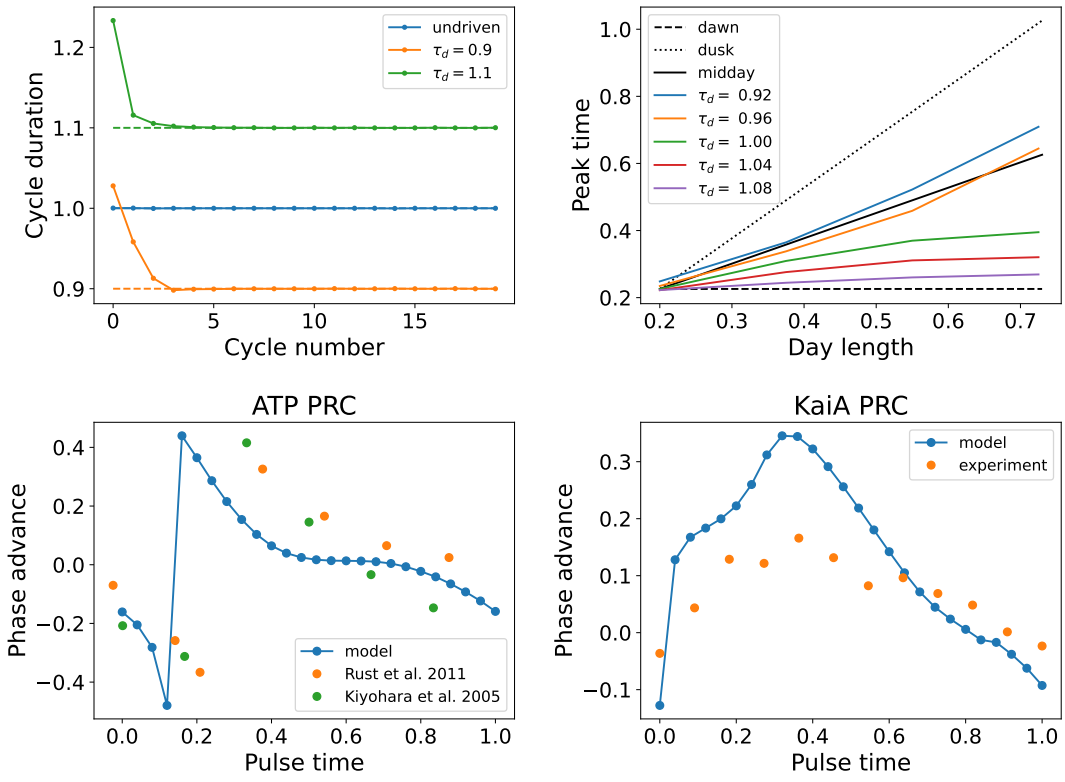


Figure 3.3: Model behavior in response to time-varying conditions, compared to experiment. (Upper left) Entrainment of the model when the solution ATP fraction is alternated between equal intervals of 25% and 100% at driving period τ_d, for τ_d between $0.9\,\tau$ and $1.1\,\tau$, where τ is the intrinsic period; dashed lines indicate the limiting cycle duration at each driving period. (Upper right) Timing of the peak in phosphorylation relative to the ATP pulse for different "day lengths" (fractions of the period at 100% ATP). Different lines show different driving periods, as indicated. (Lower left) Model and experimental [74, 116] phase response curves (PRCs) for a step down in solution ATP from 100% to 25% lasting $0.2\,\tau$. (Lower right) Model and experimental PRCs for a 16% step up in total KaiA concentration.

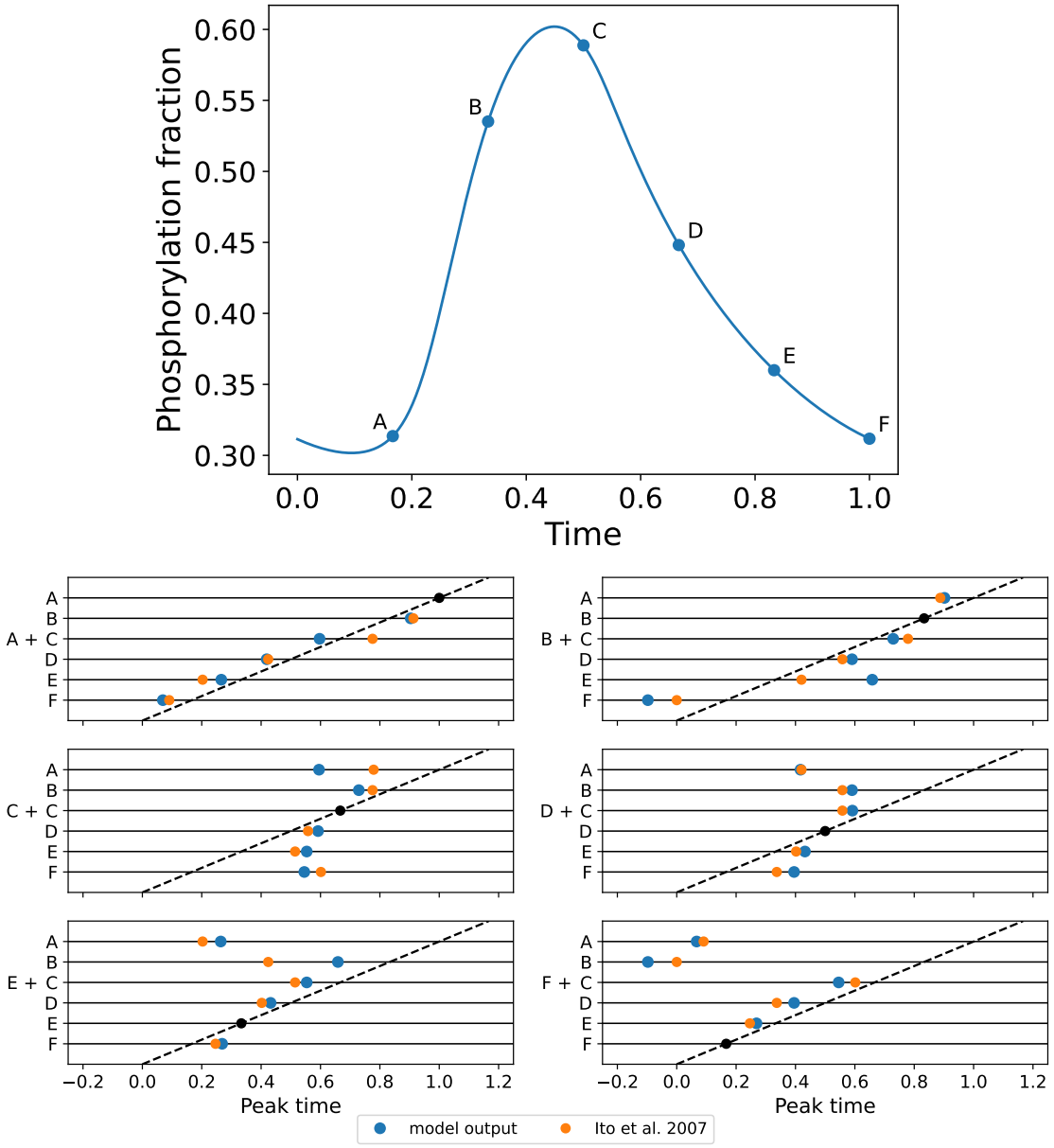


Figure 3.4: Simulations of the phase mixing experiment performed in [71]. Each point in the six lower panels represents a state of the system obtained by averaging the state of the system at two points on the curve in the upper panel. Simulations are initiated from each of these mixed conditions, and the time at which the peak in the phosphorylation fraction occurs is recorded. The lower panels show the difference between the peak time between the simulations started from the unmixed conditions, which are plotted with a black marker along the dotted diagonal line, and the mixed conditions.

spanning the sampled range. Only the first shape indicates a rate constant that is well determined. Reactions are grouped by these behaviors in 3.4.

The remaining three shapes are manifestations of the "sloppiness" typical of multiparameter fits [59] and have different implications for mechanism. Density accumulating at the lower end of the range indicates a reaction that does not help reproduce the data and that degrades the fit if its flux is too high; such reactions are natural candidates for removal from the model. Density accumulating at the upper end instead indicates a reaction that need only be fast relative to the rest of the network, so that increasing its rate constant leaves the dynamics essentially unchanged; such reactions imply a separation of time scales. Finally, there are rate constants that have appreciable density across the sampled range for various reasons. Different regions of parameter space correspond to different mechanisms, and the marginal density superposes the contributions from each; the result can be a peak on a broad background when a reaction is essential to one mechanism and dispensable in another (e.g., the rate constant for CII KaiA binding to unphosphorylated KaiC), or separated maxima when two mechanisms favor different rates. A rate constant can also be unconstrained because its substrate never accumulates, e.g., because that substrate is consumed by a much faster competing reaction, in which case the rate constant can vary over orders of magnitude without appreciably changing the dynamics.

Based on these considerations, we reduce the network by deleting reactions one at a time and asking whether oscillations persist. We start from a parameter set drawn from the high-density region of the posterior, together the reference condition of 0.6 μM KaiA dimer, 100% ATP. We select a reaction and decrease its rate constant in steps, at each step re-solving for the periodic orbit by natural parameter continuation [77], in which the orbit computed at the previous value of the rate constant serves as the initial guess for the next. Continuation allows us to follow a limit cycle whose period and amplitude change by large factors and to locate the point at which the cycle ceases to exist. If oscillations in

the reference condition do not fail at any point during the parameter sweep, we then check whether oscillations persist as we increase the KaiA concentration from 0.6 μM to 1.8 μM while keeping ATP% fixed at 100%, and when we decrease the ATP% from 100% to 25% while keeping the KaiA concentration fixed at 0.6 μM. If nontrivial periodic solutions persist over all of these sweeps, we conclude that the reaction is not required and delete it from the network; if instead the limit cycle collapses onto a fixed point at some finite value of the rate constant, we restore the reaction and move on to the next candidate. We test each reaction in the network in turn and repeat the sweep until no further reaction can be deleted. Any state that can no longer be reached by a remaining reaction is removed as well. We note that, because reactions are deleted sequentially, the resulting network need not be unique. However, we obtain qualitatively consistent results from multiple parameter sets in a given region of parameter space (we discuss alternative mechanisms in Section 3.3.7).

The reactions that survive the reduction depend on which property of the model we insist on preserving at each deletion. Requiring that the model continue to oscillate at all four oscillatory conditions in Table 3.3 with a period that varies by no more than about 10% yields the network shown in Figure 3.6, in which KaiC retains both phosphorylation sites but occupies only the U, S, and D states; we refer to this network as the SD model. Relaxing the robustness requirement, so that the model need only oscillate at the standard conditions of 0.6 μM total KaiA dimer and 100% solution ATP, yields the smaller network shown in Figure 3.6, in which KaiC carries a single phosphorylation site corresponding to S431; we refer to this network as the S-only model.

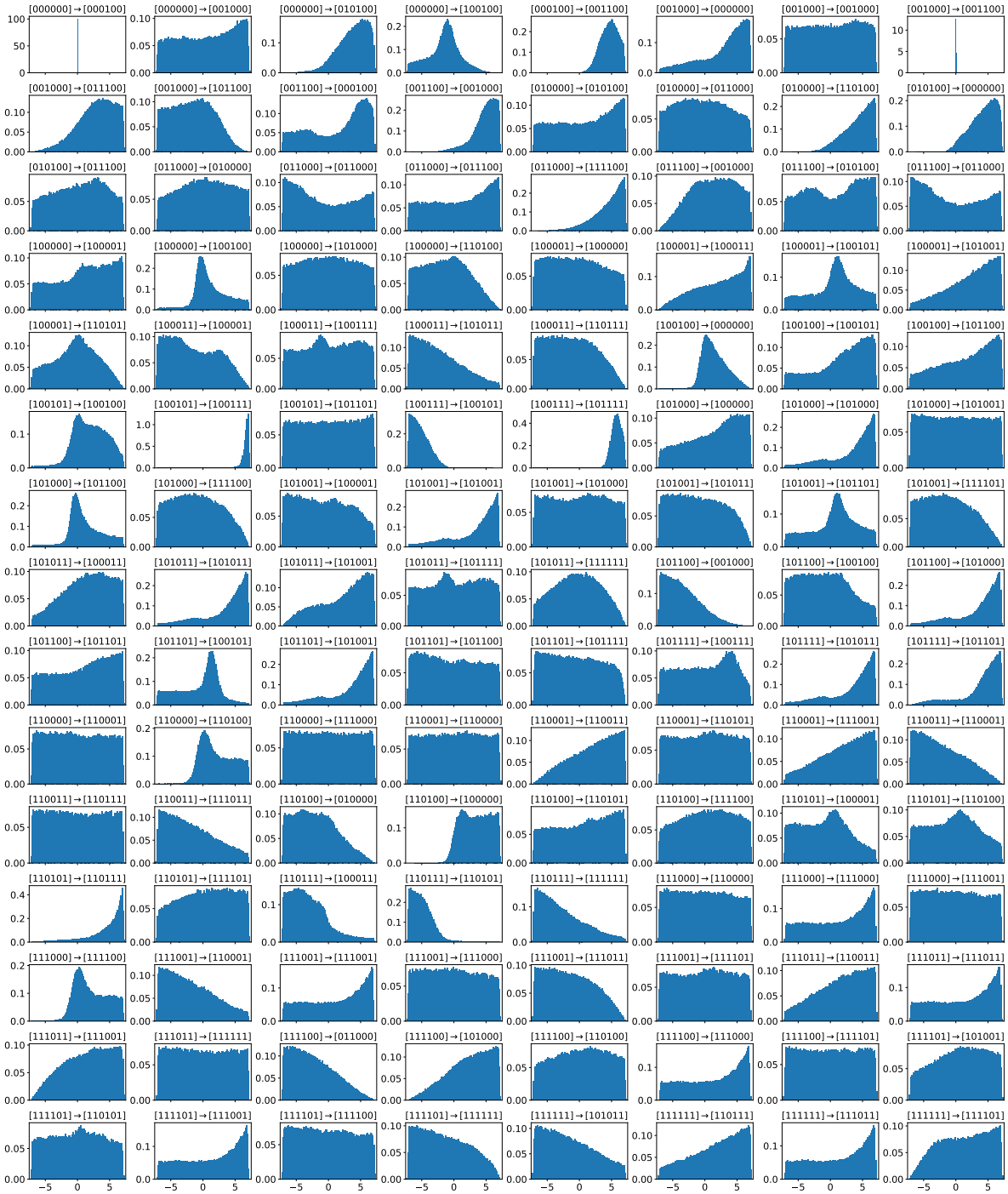


Figure 3.5: Marginal posterior densities for the rate constants of the monomeric model of the Kai oscillator, obtained by sampling subject to the Hopf point constraints and likelihood function described in Section 3.5.4. Horizontal axes are logarithmic, and each density is normalized.

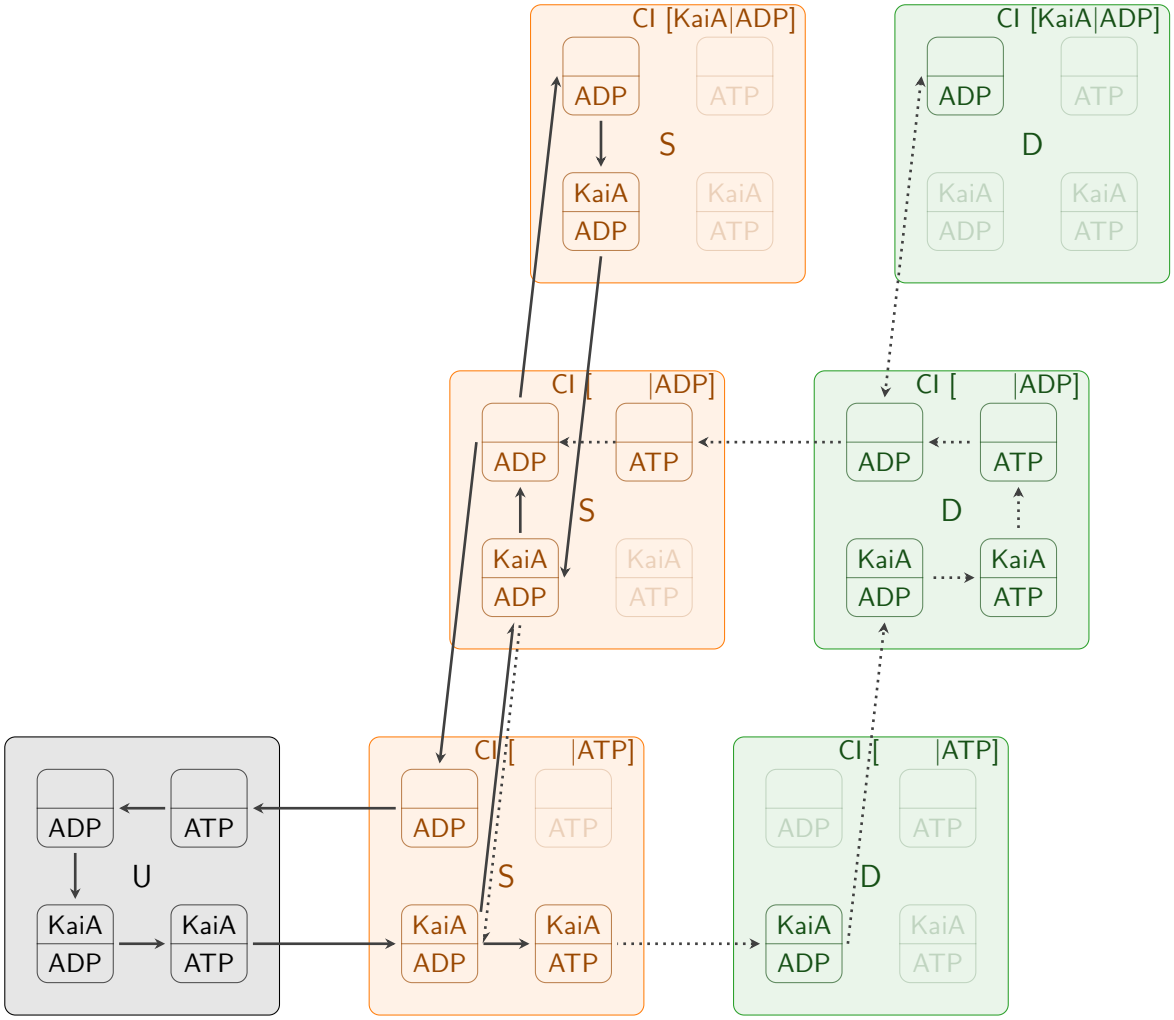


Figure 3.6: Reactions and states that remain after deleting reactions from the full model while requiring that certain properties of the model remain intact. If we only require that the system oscillates at 0.6 μM total KaiA dimer and 100% solution ATP, only the states connected by the reactions represented by solid lines are necessary (the S model). If we require that the system oscillates with a robust period, the states connected by reactions represented by dotted lines are also necessary (the SD model).

Table 3.4: Behavior of the marginal posterior densities for each rate constant. Reactants and products follow the same binary string notation as in Table 3.1. Group labels marked *(cont.)* continue from the previous page.

Posterior behavior	Reaction type	Reaction
Peaked	CI nucleotide exchange	$(1\,0\,0\,1\,0\,1) \longrightarrow (1\,0\,0\,1\,0\,0)$
	CI KaiA binding	$(1\,0\,0\,1\,0\,1) \longrightarrow (1\,0\,0\,1\,1\,1)$
	CII ATP hydrolysis	$(0\,0\,0\,0\,0\,0) \longrightarrow (0\,0\,0\,1\,0\,0)$
		$(0\,0\,1\,0\,0\,0) \longrightarrow (0\,0\,1\,1\,0\,0)$
		$(1\,0\,0\,0\,0\,0) \longrightarrow (1\,0\,0\,1\,0\,0)$
		$(1\,0\,0\,0\,0\,1) \longrightarrow (1\,0\,0\,1\,0\,1)$
		$(1\,0\,1\,0\,0\,0) \longrightarrow (1\,0\,1\,1\,0\,0)$
		$(1\,0\,1\,0\,0\,1) \longrightarrow (1\,0\,1\,1\,0\,1)$
		$(1\,1\,0\,0\,0\,0) \longrightarrow (1\,1\,0\,1\,0\,0)$
		$(1\,1\,1\,0\,0\,0) \longrightarrow (1\,1\,1\,1\,0\,0)$
	CII KaiA binding	$(0\,0\,0\,1\,0\,0) \longrightarrow (0\,0\,1\,1\,0\,0)$
		$(1\,0\,0\,1\,1\,1) \longrightarrow (1\,0\,1\,1\,1\,1)$
	CII KaiA unbinding	$(1\,0\,1\,1\,0\,1) \longrightarrow (1\,0\,0\,1\,0\,1)$
	CII phosphotransfer	$(0\,0\,0\,0\,0\,0) \longrightarrow (1\,0\,0\,1\,0\,0)$
		$(1\,0\,0\,0\,0\,1) \longrightarrow (1\,1\,0\,1\,0\,1)$
		$(1\,0\,0\,1\,0\,0) \longrightarrow (0\,0\,0\,0\,0\,0)$
		$(1\,0\,1\,0\,1\,1) \longrightarrow (1\,1\,1\,1\,1\,1)$
Increasing	CI ATP hydrolysis	$(1\,0\,0\,1\,0\,0) \longrightarrow (1\,0\,0\,1\,0\,1)$
	CI KaiA binding	$(1\,0\,0\,0\,0\,1) \longrightarrow (1\,0\,0\,0\,1\,1)$
		$(1\,1\,0\,0\,0\,1) \longrightarrow (1\,1\,0\,0\,1\,1)$
		$(1\,1\,0\,1\,0\,1) \longrightarrow (1\,1\,0\,1\,1\,1)$
	CI KaiA unbinding	$(1\,0\,1\,0\,1\,1) \longrightarrow (1\,0\,1\,0\,0\,1)$
		$(1\,0\,1\,1\,1\,1) \longrightarrow (1\,0\,1\,1\,0\,1)$
		$(1\,1\,1\,0\,1\,1) \longrightarrow (1\,1\,1\,0\,0\,1)$
		$(1\,1\,1\,1\,1\,1) \longrightarrow (1\,1\,1\,1\,0\,1)$
	CII nucleotide exchange	$(0\,0\,1\,1\,0\,0) \longrightarrow (0\,0\,1\,0\,0\,0)$
		$(1\,0\,1\,0\,0\,0) \longrightarrow (1\,0\,1\,1\,0\,0)$
		$(1\,0\,1\,0\,0\,1) \longrightarrow (1\,0\,1\,1\,0\,1)$
		$(1\,0\,1\,0\,1\,1) \longrightarrow (1\,0\,1\,1\,1\,1)$
		$(1\,0\,1\,1\,0\,0) \longrightarrow (1\,0\,1\,0\,0\,0)$
		$(1\,0\,1\,1\,0\,1) \longrightarrow (1\,0\,1\,0\,0\,1)$
		$(1\,0\,1\,1\,1\,1) \longrightarrow (1\,0\,1\,0\,1\,1)$
		$(1\,1\,1\,0\,0\,0) \longrightarrow (1\,1\,1\,1\,0\,0)$
		$(1\,1\,1\,0\,0\,1) \longrightarrow (1\,1\,1\,1\,0\,1)$
		$(1\,1\,1\,0\,1\,1) \longrightarrow (1\,1\,1\,1\,1\,1)$
		$(1\,1\,1\,1\,0\,0) \longrightarrow (1\,1\,1\,0\,0\,0)$
		$(1\,1\,1\,1\,0\,1) \longrightarrow (1\,1\,1\,0\,0\,1)$
		$(1\,1\,1\,1\,1\,1) \longrightarrow (1\,1\,1\,0\,1\,1)$
	CII KaiA binding	$(1\,0\,0\,0\,0\,1) \longrightarrow (1\,0\,1\,0\,0\,1)$
		$(1\,0\,0\,1\,0\,0) \longrightarrow (1\,0\,1\,1\,0\,0)$
		$(1\,1\,0\,0\,0\,1) \longrightarrow (1\,1\,1\,0\,0\,1)$

continued on next page

Table 3.4 (continued)

Posterior behavior	Reaction type	Reaction
Increasing *(cont.)*	CII KaiA unbinding	$(0\,0\,1\,0\,0\,0) \longrightarrow (0\,0\,0\,0\,0\,0)$
		$(0\,0\,1\,1\,0\,0) \longrightarrow (0\,0\,0\,1\,0\,0)$
		$(1\,0\,1\,0\,1\,1) \longrightarrow (1\,0\,0\,0\,1\,1)$
		$(1\,1\,1\,0\,1\,1) \longrightarrow (1\,1\,0\,0\,1\,1)$
		$(1\,1\,1\,1\,1\,1) \longrightarrow (1\,1\,0\,1\,1\,1)$
	CII phosphotransfer	$(0\,0\,0\,0\,0\,0) \longrightarrow (0\,1\,0\,1\,0\,0)$
		$(0\,0\,1\,0\,0\,0) \longrightarrow (0\,1\,1\,1\,0\,0)$
		$(0\,1\,0\,0\,0\,0) \longrightarrow (1\,1\,0\,1\,0\,0)$
		$(0\,1\,0\,1\,0\,0) \longrightarrow (0\,0\,0\,0\,0\,0)$
		$(0\,1\,1\,0\,0\,0) \longrightarrow (1\,1\,1\,1\,0\,0)$
		$(0\,1\,1\,1\,0\,0) \longrightarrow (0\,0\,1\,0\,0\,0)$
		$(1\,1\,0\,1\,0\,0) \longrightarrow (1\,0\,0\,0\,0\,0)$
		$(1\,1\,1\,1\,0\,0) \longrightarrow (1\,0\,1\,0\,0\,0)$
Decreasing	CI KaiA binding	$(1\,0\,1\,0\,0\,1) \longrightarrow (1\,0\,1\,0\,1\,1)$
		$(1\,0\,1\,1\,0\,1) \longrightarrow (1\,0\,1\,1\,1\,1)$
		$(1\,1\,1\,0\,0\,1) \longrightarrow (1\,1\,1\,0\,1\,1)$
		$(1\,1\,1\,1\,0\,1) \longrightarrow (1\,1\,1\,1\,1\,1)$
	CI KaiA unbinding	$(1\,0\,0\,0\,1\,1) \longrightarrow (1\,0\,0\,0\,0\,1)$
		$(1\,0\,0\,1\,1\,1) \longrightarrow (1\,0\,0\,1\,0\,1)$
		$(1\,1\,0\,0\,1\,1) \longrightarrow (1\,1\,0\,0\,0\,1)$
		$(1\,1\,0\,1\,1\,1) \longrightarrow (1\,1\,0\,1\,0\,1)$
	CII KaiA binding	$(1\,0\,0\,0\,1\,1) \longrightarrow (1\,0\,1\,0\,1\,1)$
		$(1\,1\,0\,0\,1\,1) \longrightarrow (1\,1\,1\,0\,1\,1)$
		$(1\,1\,0\,1\,1\,1) \longrightarrow (1\,1\,1\,1\,1\,1)$
	CII KaiA unbinding	$(1\,1\,1\,0\,0\,1) \longrightarrow (1\,1\,0\,0\,0\,1)$
	CII phosphotransfer	$(0\,0\,1\,0\,0\,0) \longrightarrow (1\,0\,1\,1\,0\,0)$
		$(1\,0\,0\,0\,0\,0) \longrightarrow (1\,1\,0\,1\,0\,0)$
		$(1\,0\,0\,0\,1\,1) \longrightarrow (1\,1\,0\,1\,1\,1)$
		$(1\,0\,1\,0\,0\,0) \longrightarrow (1\,1\,1\,1\,0\,0)$
		$(1\,0\,1\,0\,0\,1) \longrightarrow (1\,1\,1\,1\,0\,1)$
		$(1\,0\,1\,1\,0\,0) \longrightarrow (0\,0\,1\,0\,0\,0)$
		$(1\,1\,0\,1\,0\,0) \longrightarrow (0\,1\,0\,0\,0\,0)$
		$(1\,1\,0\,1\,1\,1) \longrightarrow (1\,0\,0\,0\,1\,1)$
		$(1\,1\,1\,1\,0\,0) \longrightarrow (0\,1\,1\,0\,0\,0)$
		$(1\,1\,1\,1\,1\,1) \longrightarrow (1\,0\,1\,0\,1\,1)$

continued on next page

Table 3.4 (continued)

Posterior behavior	Reaction type	Reaction
Unconstrained	CI ATP hydrolysis	$(1\,0\,0\,0\,0\,0) \longrightarrow (1\,0\,0\,0\,0\,1)$
		$(1\,0\,1\,0\,0\,0) \longrightarrow (1\,0\,1\,0\,0\,1)$
		$(1\,0\,1\,1\,0\,0) \longrightarrow (1\,0\,1\,1\,0\,1)$
		$(1\,1\,0\,0\,0\,0) \longrightarrow (1\,1\,0\,0\,0\,1)$
		$(1\,1\,0\,1\,0\,0) \longrightarrow (1\,1\,0\,1\,0\,1)$
		$(1\,1\,1\,0\,0\,0) \longrightarrow (1\,1\,1\,0\,0\,1)$
		$(1\,1\,1\,1\,0\,0) \longrightarrow (1\,1\,1\,1\,0\,1)$
	CI nucleotide exchange	$(1\,0\,0\,0\,0\,1) \longrightarrow (1\,0\,0\,0\,0\,0)$
		$(1\,0\,1\,0\,0\,1) \longrightarrow (1\,0\,1\,0\,0\,0)$
		$(1\,0\,1\,1\,0\,1) \longrightarrow (1\,0\,1\,1\,0\,0)$
		$(1\,1\,0\,0\,0\,1) \longrightarrow (1\,1\,0\,0\,0\,0)$
		$(1\,1\,0\,1\,0\,1) \longrightarrow (1\,1\,0\,1\,0\,0)$
		$(1\,1\,1\,0\,0\,1) \longrightarrow (1\,1\,1\,0\,0\,0)$
		$(1\,1\,1\,1\,0\,1) \longrightarrow (1\,1\,1\,1\,0\,0)$
	CII ATP hydrolysis	$(0\,1\,0\,0\,0\,0) \longrightarrow (0\,1\,0\,1\,0\,0)$
		$(0\,1\,1\,0\,0\,0) \longrightarrow (0\,1\,1\,1\,0\,0)$
		$(1\,0\,0\,0\,1\,1) \longrightarrow (1\,0\,0\,1\,1\,1)$
		$(1\,0\,1\,0\,1\,1) \longrightarrow (1\,0\,1\,1\,1\,1)$
		$(1\,1\,0\,0\,0\,1) \longrightarrow (1\,1\,0\,1\,0\,1)$
		$(1\,1\,0\,0\,1\,1) \longrightarrow (1\,1\,0\,1\,1\,1)$
		$(1\,1\,1\,0\,0\,1) \longrightarrow (1\,1\,1\,1\,0\,1)$
		$(1\,1\,1\,0\,1\,1) \longrightarrow (1\,1\,1\,1\,1\,1)$
	CII nucleotide exchange	$(0\,0\,1\,0\,0\,0) \longrightarrow (0\,0\,1\,1\,0\,0)$
		$(0\,1\,1\,0\,0\,0) \longrightarrow (0\,1\,1\,1\,0\,0)$
		$(0\,1\,1\,1\,0\,0) \longrightarrow (0\,1\,1\,0\,0\,0)$
	CII KaiA binding	$(0\,0\,0\,0\,0\,0) \longrightarrow (0\,0\,1\,0\,0\,0)$
		$(0\,1\,0\,0\,0\,0) \longrightarrow (0\,1\,1\,0\,0\,0)$
		$(0\,1\,0\,1\,0\,0) \longrightarrow (0\,1\,1\,1\,0\,0)$
		$(1\,0\,0\,0\,0\,0) \longrightarrow (1\,0\,1\,0\,0\,0)$
		$(1\,0\,0\,1\,0\,1) \longrightarrow (1\,0\,1\,1\,0\,1)$
		$(1\,1\,0\,0\,0\,0) \longrightarrow (1\,1\,1\,0\,0\,0)$
		$(1\,1\,0\,1\,0\,0) \longrightarrow (1\,1\,1\,1\,0\,0)$
		$(1\,1\,0\,1\,0\,1) \longrightarrow (1\,1\,1\,1\,0\,1)$
	CII KaiA unbinding	$(0\,1\,1\,0\,0\,0) \longrightarrow (0\,1\,0\,0\,0\,0)$
		$(0\,1\,1\,1\,0\,0) \longrightarrow (0\,1\,0\,1\,0\,0)$
		$(1\,0\,1\,0\,0\,0) \longrightarrow (1\,0\,0\,0\,0\,0)$
		$(1\,0\,1\,0\,0\,1) \longrightarrow (1\,0\,0\,0\,0\,1)$
		$(1\,0\,1\,1\,0\,0) \longrightarrow (1\,0\,0\,1\,0\,0)$
		$(1\,0\,1\,1\,1\,1) \longrightarrow (1\,0\,0\,1\,1\,1)$
		$(1\,1\,1\,0\,0\,0) \longrightarrow (1\,1\,0\,0\,0\,0)$
		$(1\,1\,1\,1\,0\,0) \longrightarrow (1\,1\,0\,1\,0\,0)$
		$(1\,1\,1\,1\,0\,1) \longrightarrow (1\,1\,0\,1\,0\,1)$
	CII phosphotransfer	$(1\,1\,0\,1\,0\,1) \longrightarrow (1\,0\,0\,0\,0\,1)$
		$(1\,1\,1\,1\,0\,1) \longrightarrow (1\,0\,1\,0\,0\,1)$

Notably, neither reduced model retains the T state. At the best fit parameters the full model carries substantial flux through the U→T reactions, yet these reactions can be removed without abolishing oscillations. This result is at odds with the conventional understanding, in which phosphorylation proceeds sequentially with T432 phosphorylated before S431, but it is consistent with the observation that the T432V mutant oscillates with a period approximately twice that of the wild-type clock [44].

Comparing the two reduced models suggests instead that the role of T432 is to confer robustness to variations in KaiA concentration (Figure 3.7, left). The SD model oscillates with a stable period over the tested range of KaiA concentrations, whereas the S-only model oscillates only below 0.8 μM, with the period lengthening sharply as that limit is approached. The computational results thus lead to a prediction: the T432V clock should be less robust to changes in KaiA concentration. To test this prediction, we reconstitute the system in vitro with wild-type or T432V KaiC, fluorescently labeled KaiB, and ATP, over a range of KaiA concentrations, and track oscillations by fluorescence polarization (Figure 3.8). Consistent with the prediction, the T432V clock oscillates over a narrower range of KaiA concentrations than the wild-type clock, with the period lengthening as the concentration increases (Figure 3.7, right).

The sampling and reduction procedures thus reveal an unanticipated function for T432 that the experiments confirm: the site is essential for stabilizing the period rather than generating the oscillation. In the next section, we take advantage of the simplicity of the SD and S-only models to develop analytically tractable approximations that show how the period depends on individual rate constants.

3.3.3 A minimal model recasts the oscillator as a competition between KaiA binding at CI and at CII.

As a first step toward understanding the function of T432 and, more generally, the factors that determine the period, we further reduce the S-only model. The S only model contains two loops of reactions. The first loop describes the phosphorylation dynamics of KaiC and passes through

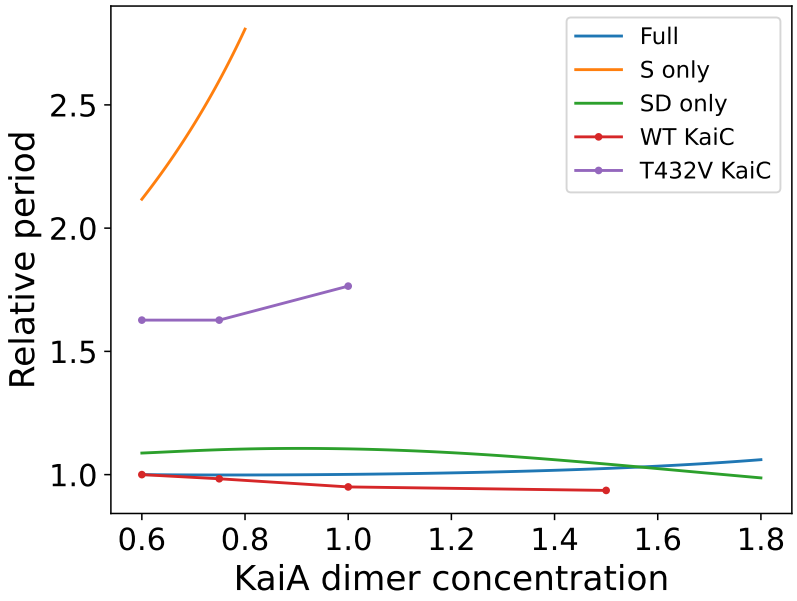


Figure 3.7: Simulated and experimentally measured periods as a function of the total KaiA dimer concentration. The simulated periods are reported relative to the period of the full model at 0.6 μM KaiA dimer concentration, and the experimental periods are repored relative to the wild-type clock at 0.6 μM KaiA dimer concentration. The full model and the SD model oscillate over the entire range of KaiA concentrations at which oscillations have been observed experimentally, whereas the S-only model fails at substantially lower KaiA concentrations. T432V KaiC is experimentally observed to function over a narrower range of KaiA concentrations than wild-type KaiC.

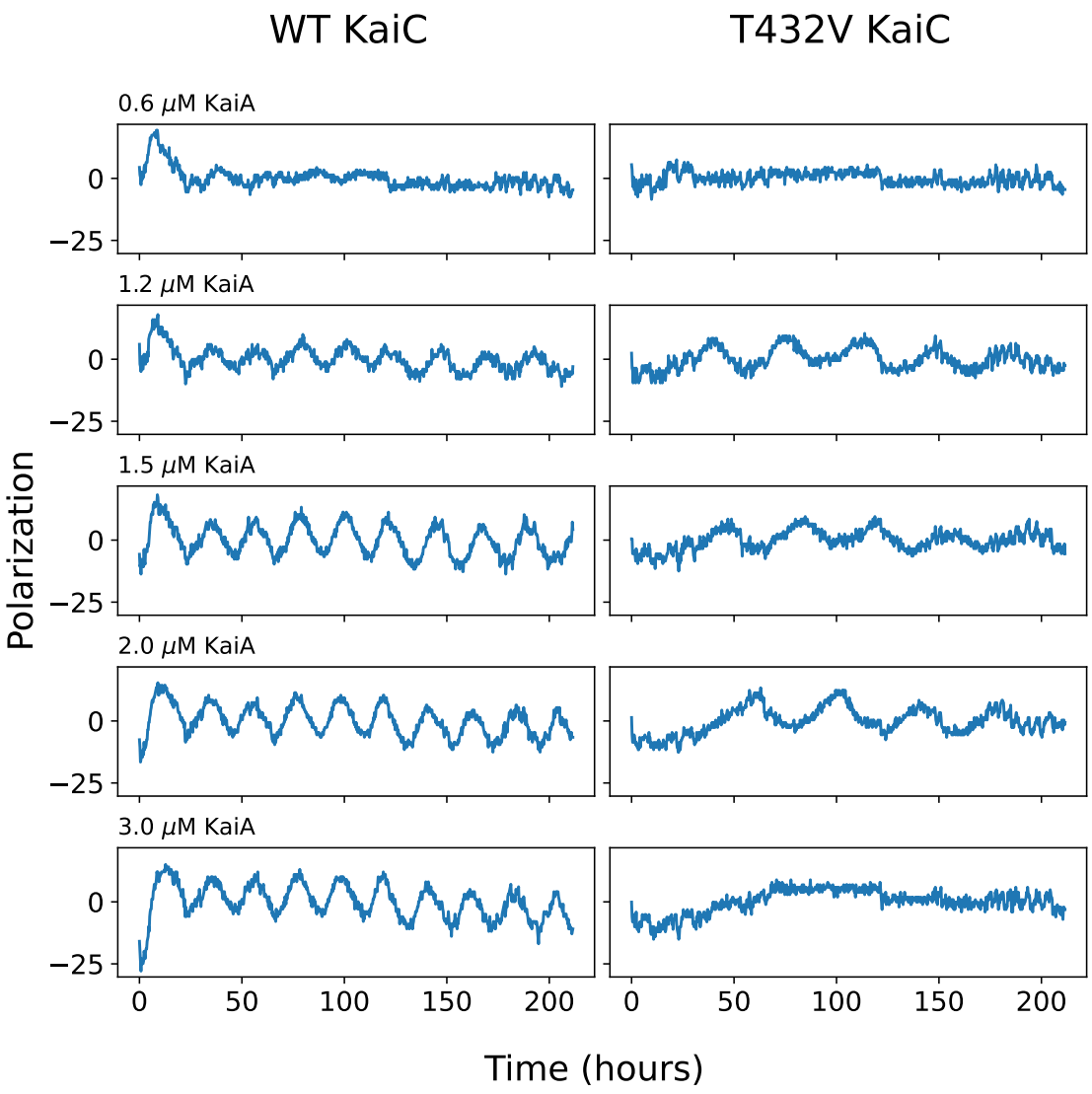


Figure 3.8: Fluorescence polarization measurements of fluorescently labeled KaiB reconstituted with KaiA and either wild-type KaiC or T432V KaiC.

states in the bottom S panel and the states in the U panel in Figure 3.6. This loop imposes a delay such that KaiC cannot immediately resequester KaiA that has been released from CI. We replace this sequence of states with a single state representing unphosphorylated KaiC and the sequence of reactions with a reaction to the [KaiA|ADP] state in the bottom S panel. The rate of this reaction is governed by the slowest reaction in the sequence. The second loop describes the dynamics of KaiA sequestration and joins states in the two upper S panels in Figure 3.6. Since KaiA is rapidly released from CI if KaiA is also bound on CII, we remove the [KaiA|ADP] state in the top S panel in Figure 3.6 and join the [|ADP] state in the top S panel directly to the [KaiA|ADP] state in the middle S panel. The rate of this reaction is determined by the rate of KaiA binding to the [|ADP] state in the top S panel. The resulting minimal model retains five species: free KaiA (a), unphosphorylated KaiC (u), and three forms of KaiC phosphorylated on S431, distinguished by whether KaiA is absent (s_0), sequestered at CI (s_1), or bound at CII (s_2) (Figure 3.9 and Tables 3.5 and 3.6). Their concentrations, expressed relative to the total KaiC monomer concentration, obey

$$\frac{du}{dt} = k_{su}s_0 - k_{us}ua \tag{3.1a}$$
$$\frac{ds_0}{dt} = -k_{su}s_0 - k_{01}s_0a + k_{20}s_2 \tag{3.1b}$$
$$\frac{ds_1}{dt} = k_{01}s_0a - k_{12}s_1a \tag{3.1c}$$
$$\frac{ds_2}{dt} = -k_{20}s_2 + k_{12}s_1a + k_{us}ua \tag{3.1d}$$
$$\frac{da}{dt} = -k_{01}s_0a + \frac{5}{6}k_{12}s_1a + \frac{1}{6}k_{20}s_2 - \frac{1}{6}k_{us}ua. \tag{3.1e}$$

Only three of the five variables are independent because total KaiC and total KaiA are conserved,

$$u + s_0 + s_1 + s_2 = 1 \tag{3.2a}$$
$$a + s_1 + \frac{1}{6}s_2 = A, \tag{3.2b}$$

where A is the total KaiA concentration in the same units. The factors of 1/6 arise because a KaiC hexamer can only bind one KaiA dimer at CII; by contrast, each KaiC monomer can bind KaiA at

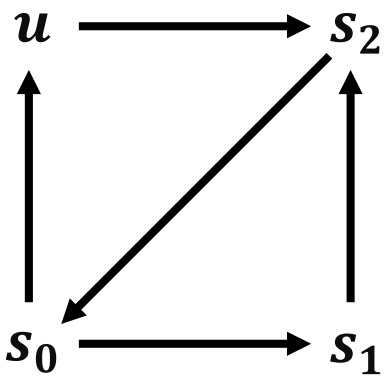


Figure 3.9: A minimal model of the single-site oscillator.

CI, so the net gain of free KaiA in the k_{12} reaction is 5/6 of a dimer.

We take KaiC that is phosphorylated on S431 but is not bound by KaiA to have hydrolyzed ATP and bound KaiB at CI, so that s_0 and s_1 represent KaiB–KaiC and KaiA–KaiB–KaiC complexes, respectively. KaiA binding at CII, nucleotide exchange, and phosphotransfer are treated together and produce s_2 at rate $k_{us}ua$. Dephosphorylation returns s_0 to u at rate $k_{su}s_0$. At CI, KaiA binds s_0 at rate $k_{01}s_0a$; at CII, it unbinds s_2 at rate $k_{20}s_2$. The two domains are coupled by a reaction in which a second dimer binds the CII domain of a sequestered complex and releases the first dimer at rate $k_{12}s_1a$. This reaction, which results directly from the sampling and model reduction, is consistent with the observations that KaiA inhibits KaiB binding at CI and destabilizes existing KaiB–KaiC complexes [89].

Table 3.5: Dynamical variables of the minimal single-site model (3.1). Concentrations are expressed in units of the total KaiC monomer concentration. No state carries KaiA at CII and a hydrolyzed CI simultaneously, because release of KaiA from CI is fast in the doubly bound complex and that complex is contracted out of (3.1); a complex with KaiA at CII therefore reaches s_0 only after KaiA leaves CII.

Variable	Species
u	Unphosphorylated KaiC
s_0	KaiC phosphorylated on S431 and able to sequester KaiA
s_2	KaiC phosphorylated on S431 with KaiA bound at CII
s_1	KaiC phosphorylated on S431 with KaiA sequestered at CI
a	Free KaiA

We analyze the minimal model in the regime $\mathcal{O}(1) = k_{us} \approx k_{su} \approx k_{20} \ll k_{12} \ll k_{01}$. Phosphorylation, dephosphorylation, and KaiA release from CII are comparable; KaiA binding at the CII domain of a sequestered complex is much faster, and sequestration at CI is faster still. The last of

Table 3.6: Reactions of the single-site minimal model (3.1) (Figure 3.9). KaiC concentrations are expressed as fractions of the total KaiC monomer concentration, and a denotes a free KaiA dimer; because one KaiA dimer binds the CII domain of a KaiC hexamer, KaiA binding at CII consumes 1/6 unit of a per KaiC monomer, whereas sequestration at CI consumes one unit of a per monomer. The k_{12} reaction consumes 1/6 unit of a at CII and liberates the dimer held at CI, for a net gain of 5/6 unit of a.

Reaction	Description	Rate
$u + \frac{1}{6}a \longrightarrow s_2$	S431 phosphorylation with KaiA at CII	$k_{us}ua$
$s_0 \longrightarrow u$	S431 dephosphorylation	$k_{su}s_0$
$s_0 + a \longrightarrow s_1$	KaiA sequestration at CI	$k_{01}s_0a$
$s_1 + \frac{1}{6}a \longrightarrow s_2 + a$	KaiA binding at CII and release from CI	$k_{12}s_1a$
$s_2 \longrightarrow s_0 + \frac{1}{6}a$	KaiA unbinding from CII	$k_{20}s_2$

these relations is consistent with the sampling, in which the posterior density for the sequestration rate constant accumulates at the upper end of the sampled range (Figure 3.5). In this parameter regime, the limit cycle can be divided into two segments, corresponding to phosphorylation and dephosphorylation, connected by fast transitions (Figure 3.10). We compute an asymptotic estimate for the limit cycle and its period using the method of matched asymptotic expansions. In each segment, we rescale the variables and time to bring the dominant balance to leading order, substitute into (3.1), and expand the solution in powers of a local small parameter ($\epsilon \equiv 1/k_{01}$, $\delta \equiv k_{12}/k_{01}$, or $\xi \equiv 1/k_{12}$). We then require that the solutions in adjacent segments agree where they overlap. Appendix 3.6.1 gives full mathematical details of the asymptotic analysis; here, we focus on its interpretation.

3.3.4 The phosphorylation and dephosphorylation segments are determined by the relative rates of KaiA sequestration at CI and release at CII.

At the start of the dephosphorylation segment, s_2 is high and begins to decay at rate $k_{20}s_2$. Each unit of s_2 that decays yields one unit of s_0 and 1/6 unit of a, which rapidly converts 1/6 unit of s_0 to s_1 because k_{01} is large. Balancing release of KaiA with sequestration gives

$$a = \frac{k_{20}s_2}{6k_{01}s_0} \sim \epsilon. \tag{3.3}$$

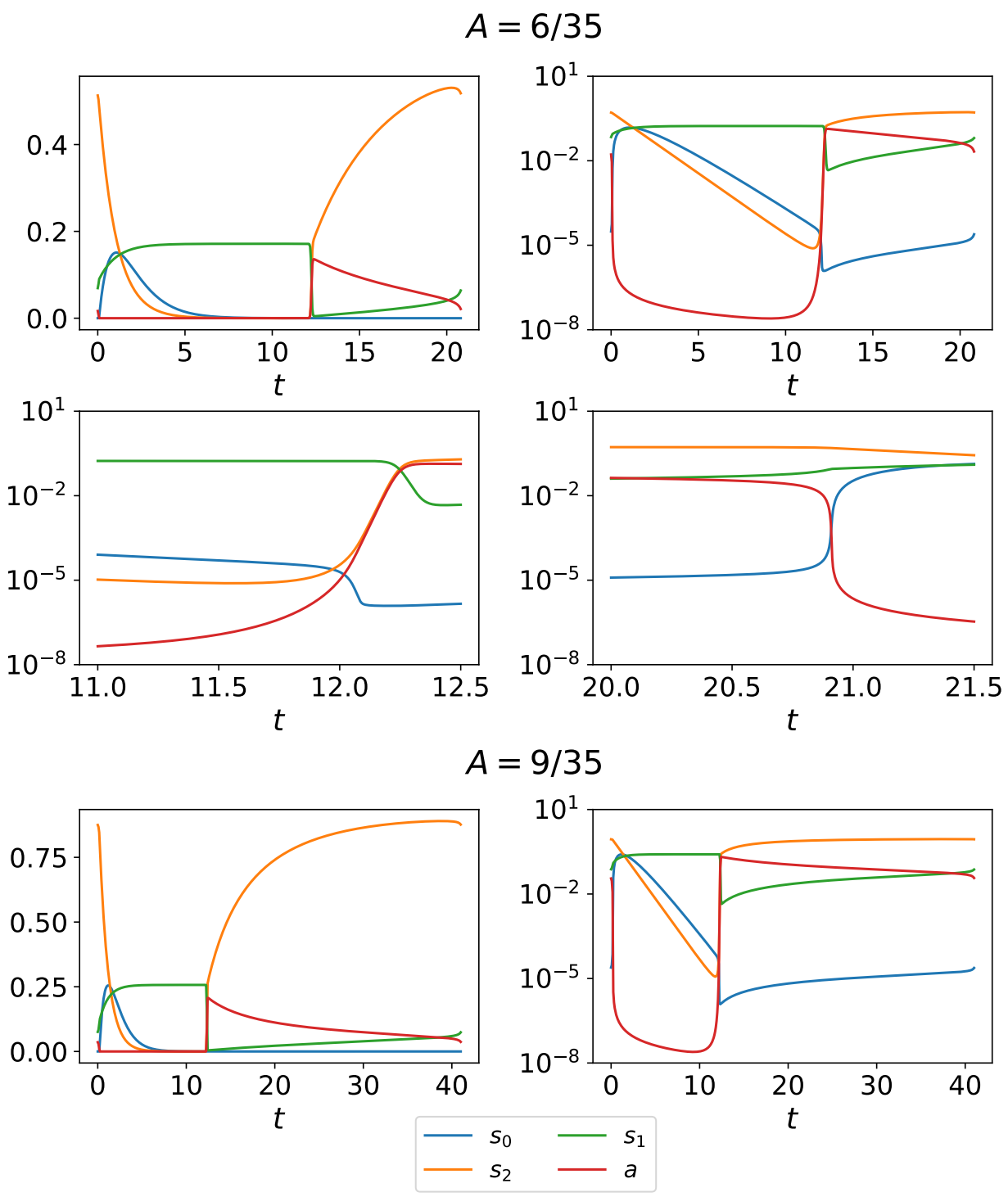


Figure 3.10: Numerical solutions of the single-site model of the Kai oscillator (3.1) when using the parameter values $k_{01} = 10^6, k_{12} = 300, k_{su} = k_{us} = k_{20} = 1$. The solution for $A = 6/35$ has been magnified and plotted on a log scale to emphasize the different segments of the limit cycle.

With $a \sim \epsilon$, no phosphorylation occurs to leading order. Consequently, any s_0 that is not converted to s_1 dephosphorylates to u at rate $k_{su}s_0$. Thus both s_0 and s_2 decay exponentially after an initial transient, as seen from the linear behavior of their logarithms from $t \approx 1$ to $t \approx 10$ in Figure 3.10. Over this interval, $s_1 \to A$.

Eq. (3.3) holds only while sequestration at CI outruns release at CII, that is, while $k_{01}s_0 \gg k_{12}s_1$. The two rates become comparable when $s_0 \sim \delta = k_{12}/k_{01} = \epsilon/\xi$. The threshold that terminates dephosphorylation is thus the ratio of the CII and CI binding rate constants; the segment boundary is set by the antagonism between the two binding sites rather than by the dephosphorylation rate constant alone.

At $t \approx 10$, $s_0 \sim \delta$, so that from $t \approx 10$ to $t \approx 12$ the rate of the $k_{01}s_0a$ and $k_{12}s_1a$ reactions are comparable. In the $k_{12}s_1a$ reaction, $1/6$ unit of a binds to 1 unit of s_1 to form 1 unit s_2 and 1 unit of a for a net gain of $5/6$ unit of a, which can then further deplete s_0. When $s_0 \sim \delta$ the $k_{01}s_0a$ reaction still occurs and consumes s_0 to produces s_1, but once $s_0 = k_{20}s_2/(k_{01}a) \sim \epsilon$ the $k_{12}s_1a$ reaction dominates, consuming s_1 to produce s_2 until $s_1 = k_{20}s_2/k_{12}a \sim \xi$ at $t \approx 12.3$.

At this point, $a \sim 1$ and the system enters the phosphorylation segment. In this segment, the $k_{us}ua$ reaction produces s_2 and consumes a until $a \sim s_1 \sim \sqrt{\xi}$ and $k_{20}s_2 \gg k_{us}ua$. When $a \sim \sqrt{\xi}$, the rate $k_{us}ua$ becomes negligible, and the s_0 formed from the decay of s_2 consumes the remaining a until $a \sim \epsilon$, and the system enters the dephosphorylation segment.

3.3.5 The duration of the phosphorylation segment is sensitive to KaiA and nucleotide concentrations.

In the asymptotic regime, the durations of the dephosphorylation and phosphorylation segments are much longer than the durations of the transitions between them, and contributions of the transitions to the period can be neglected.

The duration of the phosphorylation segment can be shown (Section 3.6.2) to scale as

$$\tau_{\text{phos}} = t_{\text{phos}}(s_{A,\text{fold}}) - t_{\text{phos}}(s_{A,\text{start}}), \tag{3.4}$$

where

$$
\begin{aligned}
t_{\text{phos}}(s_A) = &- \frac{1}{12k_{us}k_{20}(1-A)}\Bigg[2G\arctan\left(\frac{k_{12}(1-s_A)+F}{G}\right) \\
&+ 2(6k_{20}-5k_{12}A)\log\left(-6k_{20}+k_{12}(6A-s_A)+F\right) \\
&- (6k_{20}-k_{12}(1+4A))\log\left(66k_{20}+k_{12}(6A-s_A)+F\right) \\
&- (6k_{20}+k_{12}(1-6A))\log\left(k_{12}(1-s_A)\right)\Bigg] \\
F \equiv &\sqrt{36(k_{20}+k_{12}A)^2 - 12k_{12}(11k_{20}+k_{12}A)s_A + \left(k_{12}s_A\right)^2} \\
G \equiv &\sqrt{-36k_{20}^2 + 12k_{20}k_{12}(11-6A) - k_{12}^2(6A-1)^2} \\
s_{A,\text{start}} = &\, A + \frac{6k_{20}}{5k_{12}} \\
s_{A,\text{fold}} \equiv &\, \frac{6(11k_{20}+k_{12}A-2\sqrt{5k_{20}(6k_{20}+k_{12}A)})}{k_{12}}.
\end{aligned}
\tag{3.5}
$$

Analysis of (3.4) shows oscillations fail when the dephosphorylation segment intersects a saddle node at a total KaiA concentration of

$$A_+ = \frac{1}{6} - \frac{k_{20}}{k_{12}} + \frac{1}{3}\sqrt{\frac{30k_{20}}{k_{12}}}. \tag{3.6}$$

Below A_+, as A increases, more time is required for a to fall to $a \sim \sqrt{\xi}$, and, at $A = A_+$, enough free KaiA is present in solution for the $k_{12}s_1a$ reaction to keep a well above $\sqrt{\xi}$ at all times. For $k_{20}/k_{12} < 5/6$, the regime we considered above, (3.6) shows that the range of A tolerated can be increased by increasing k_{20} and decreasing k_{12}. For example, for $k_{01} = 10^6, k_{12} = 300, k_{su} = k_{us} = k_{20} = 1$, numerical continuation indicates that oscillations fail through a homoclinic bifurcation at $A \approx 0.2687 \approx 9.4/35$; if k_{12} is decreased to 30, oscillations instead fail at $A \approx 0.4666 \approx 16.33/35$. These thresholds agree with the asymptotic result (3.6) quantitatively.

In contrast to the duration of the phosphorylation segment, the duration of the dephosphorylation segment remains bounded for $A \geq 0$. The formula depends on the precise relation between the

rate constants k_{su} and k_{20},

$$\tau_{\text{dephos}} \approx \begin{cases} -\frac{1}{k_{su}} \left(\log\left[\frac{\delta}{K}\right] + \log\left[1 + \left(\frac{\delta}{K}\right)^{\frac{k_{20}}{k_{su}}-1}\right]\right) & \text{if} \quad k_{su} < k_{20} \\ -\frac{1}{\kappa}\left[\log\left(\frac{6\delta}{5s_2(0)}\right) - \log\left(-\log\left[\frac{6\delta}{5s_2(0)}\right]\right)\right] & \text{if} \quad k_{su} = k_{20} \equiv \kappa \\ -\frac{1}{k_{20}} \left(\log\left[-\frac{\delta}{K}\right] - \log\left[1 - \left(-\frac{\delta}{K}\right)^{\frac{k_{su}}{k_{20}}-1}\right]\right) & \text{if} \quad k_{su} > k_{20}, \end{cases} \tag{3.7}$$

where

$$K \equiv \frac{5k_{20}s_2(0)}{6(k_{20} - k_{su})}. \tag{3.8}$$

In all cases, the dependence on A, which enters through the initial condition $s_2(0) \approx 6A - 2\sqrt{5A\xi}$ (from (3.2) and $a \sim s_1 \sim \sqrt{\xi}$), is logarithmic. A comparison of the dependence of the phosphorylation and dephosphorylation segment durations on A is shown in Figure 3.11.

Since the dephosphorylation segment is much less sensitive than the phosphorylation segment to changes in A, the sensitivity of the overall period to A can be decreased by choosing parameters such that the duration of the dephosphorylation segment dominates the overall period. Examining (3.126) shows that the duration of the phosphorylation segment can be shortened by increasing k_{us} and k_{20} and decreasing k_{12}, and (3.137,3.138,3.140) show that the duration of the dephosphorylation segment can be lengthened by decreasing k_{12}, k_{20}, and k_{su} and increasing k_{01}. The behavior of the segment durations as these parameters are varied is shown in Figure 3.12.

Consistent with KaiA's role as a nucleotide exchange factor, the effect of the nucleotide ratio on the period of the oscillator can be controlled similarly. In all of our models, the nucleotide ratio affects the dynamics by modulating the phosphorylation rate, and, for the single-site model, it specifically scales k_{us}. This parameter only appears in the denominator of the multiplicative prefactor in (3.4). Thus, increasing the ATP:ADP ratio shortens the duration of the phosphorylation segment and, in turn, the overall period. In contrast, to leading order in ϵ, k_{us} does not affect the duration of the dephosphorylation segment. Thus, robustness of the period to variations in the

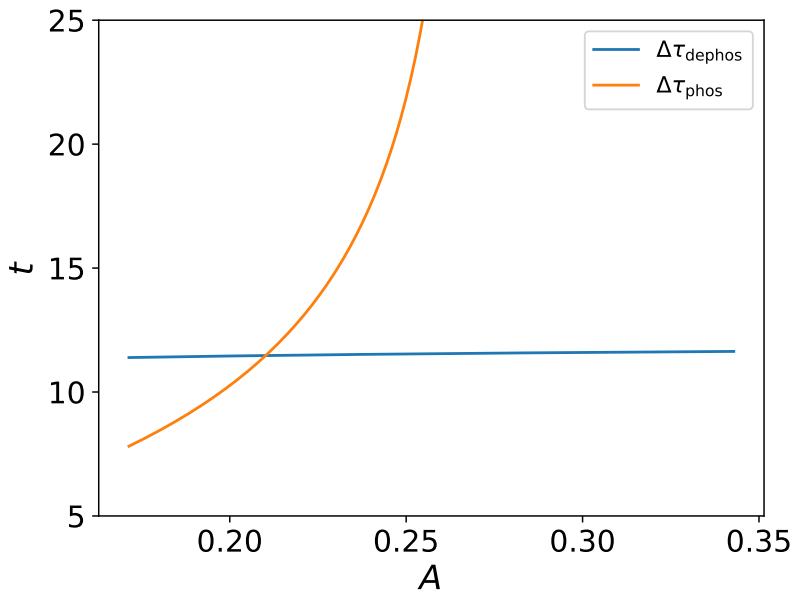


Figure 3.11: The dependence on A of duration of the dephosphorylation and phosphorylation segments of the single-site Kai oscillator model when using the parameter values $k_{01} = 10^6, k_{12} = 300, k_{su} = k_{us} = k_{20} = 1$. The duration of the phosphorylation segment has a logarithmic singularity at $A = 1/6 - k_{20}/k_{12} + \sqrt{30k_{20}/k_{12}}/3 \approx 0.2687$.

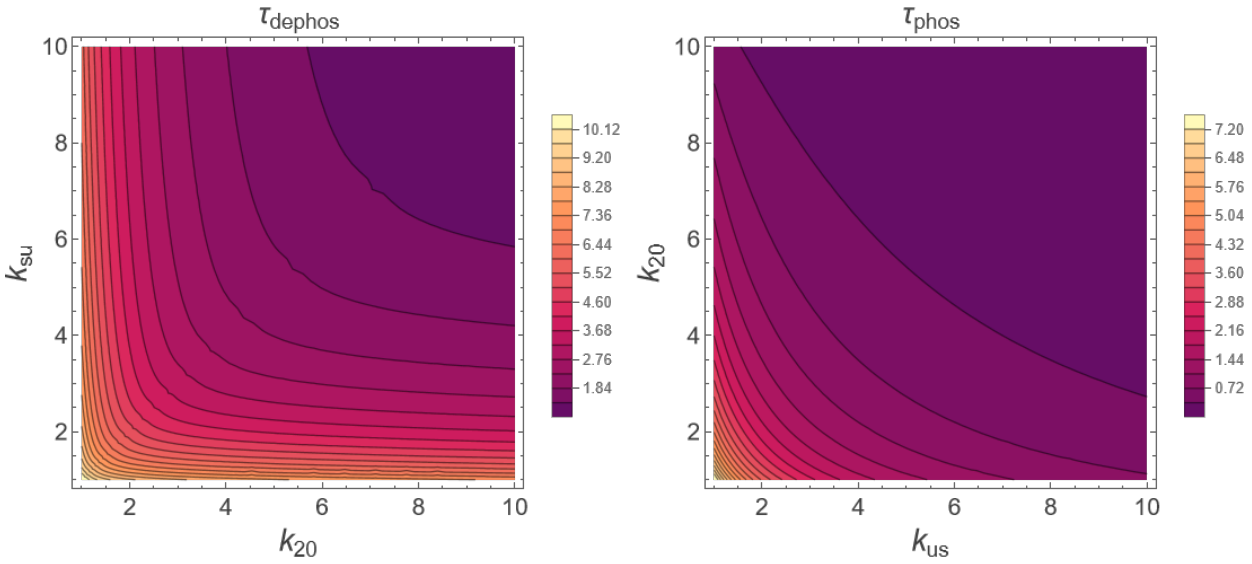


Figure 3.12: The duration of the dephosphorylation and phosphorylation segments of the single-site Kai oscillator model when using the parameter values $k_{01} = 10^6, k_{12} = 300, A = 6/35$. The duration of the phosphorylation segment is evaluated with $k_{su} = 1$, and the duration of the dephosphorylation segment is evaluated with $k_{us} = 1$.

nucleotide ratio can also be achieved by having the duration of the dephosphorylation segment dominate the total duration of the limit cycle.

That said, the hierarchy of scales in the reaction constants required for robust oscillations in the single-site model results in unrealistic solutions: the system spends long stretches of time in a strongly phosphorylated state before rapidly jumping to a short-lived weakly phosphorylated state. Experimentally, the waveform of total KaiC phosphorylation is close to sinusoidal and has no sharp transitions, with the phosphorlyation segment being only slightly shorter than the dephosphorylation segment. This suggests that the T432 phosphorylation site provides additional flexibility that allows the full system to function with a smaller separation of scales than when only a single phosphorylation site is present.

3.3.6 Doubly-phosphorylated KaiC stabilizes the period by acting as a sink for KaiA.

Our simulations, theory, and experiments suggest that the role of T432 is to confer robustness to variations in KaiA (and by extension nucleotide ratio). In studies with KaiC S431E;T432A and KaiC S431E;T432E, which mimic the S and D phosphoforms, respectively, Lin *et al.* [89] observed that KaiA has a weaker inhibitory effect on the formation KaiB:KaiC complexes with the D phosphomimic than the S phosphomimic. More D forms at higher KaiA concentrations. If KaiA has a weaker inhibitory effect on the formation of KaiB-KaiC complexes by D, D could replace S as the primary sink for KaiA at high KaiA concentrations, extending the range of clock function.

To test this hypothesis, we reintroduce the D state to the single-site minimal model to obtain a double-site minimal model (Figure 3.13 and Table 3.7)

$$\frac{du}{dt} = k_{su}s_0 - k_{us}ua \tag{3.9a}$$

$$\frac{ds_0}{dt} = -k_{01}s_0a - k_{su}s_0 + k_{ds}q_0 + k_{20}s_2 \tag{3.9b}$$

$$\frac{ds_1}{dt} = k_{01}s_0a - k_{12}s_1a \tag{3.9c}$$

$$\frac{ds_2}{dt} = k_{12}s_1a - k_{20}s_2 - k_{sd}s_2 + k_{us}ua \tag{3.9d}$$

$$\frac{dq_0}{dt} = -k_{d01}q_0a - k_{ds}q_0 + k_{d20}q_2 \tag{3.9e}$$

$$\frac{dq_1}{dt} = k_{d01}q_0a - k_{d12}q_1a \tag{3.9f}$$

$$\frac{dq_2}{dt} = k_{sd}s_2 + k_{d12}q_1a - k_{d20}q_2 \tag{3.9g}$$

$$\frac{da}{dt} = -\frac{1}{6}k_{us}ua - (k_{01}s_0 + k_{d01}q_0)\,a + \frac{5}{6}(k_{12}s_1 + k_{d12}q_1)\,a + \frac{1}{6}(k_{20}s_2 + k_{d20}q_2) \tag{3.9h}$$

Where we have introduced additional variables for D that is able to sequester KaiA (q_0), D with KaiA bound to CI (q_1), and D with KaiA bound to CII (q_2), paralleling the variables for S in the single-site model. Total KaiC and total KaiA are conserved analogously,

$$u + s_0 + s_1 + s_2 + q_0 + q_1 + q_2 = 1 \tag{3.10a}$$

$$a + s_1 + \frac{1}{6}s_2 + q_1 + \frac{1}{6}q_2 = A. \tag{3.10b}$$

KaiC phosphorylated only on the T432 site is still not allowed to form.

We compare the behavior of the single-site and double-site minimal models in a non-asymptotic parameter regime: $k_{01} = 650, k_{12} = 600, k_{su} = 1, k_{us} = 40, k_{20} = 5, k_{d01} = 650, k_{d12} = 200, k_{ds} = 1, k_{sd} = 1, k_{d20} = 60$ (Figure 3.14). As A approaches 0.2879, the period of the single-site model lengthens significantly before oscillations fail by a fold of limit cycles. In contrast, the double-site model maintains a robust period throughout the physically relevant range of A.

Table 3.7: Reactions added to the single-site minimal model to obtain the double-site minimal model (Figure 3.13). The reactions of Table 3.6 are retained unchanged. The last three reactions parallel the corresponding reactions of the S phosphoform, with rate constants distinguished by the subscript d; the two T432 phosphotransfer reactions join the S and D branches. Conventions for concentrations and KaiA stoichiometry are as in Table 3.6.

Reaction	Description	Rate
$s_2 \longrightarrow q_2$	T432 phosphorylation with KaiA at CII	$k_{sd}s_2$
$q_0 \longrightarrow s_0$	T432 dephosphorylation	$k_{ds}q_0$
$q_0 + a \longrightarrow q_1$	KaiA sequestration at CI	$k_{d01}q_0a$
$q_1 + \frac{1}{6}a \longrightarrow q_2 + a$	KaiA binding at CII and release from CI	$k_{d12}q_1a$
$q_2 \longrightarrow q_0 + \frac{1}{6}a$	KaiA unbinding from CII	$k_{d20}q_2$

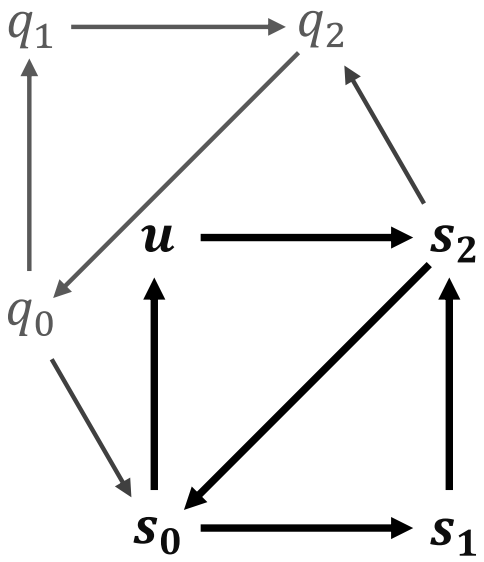


Figure 3.13: A minimal model of the Kai oscillator with the doubly phosphorylated state of KaiC restored.

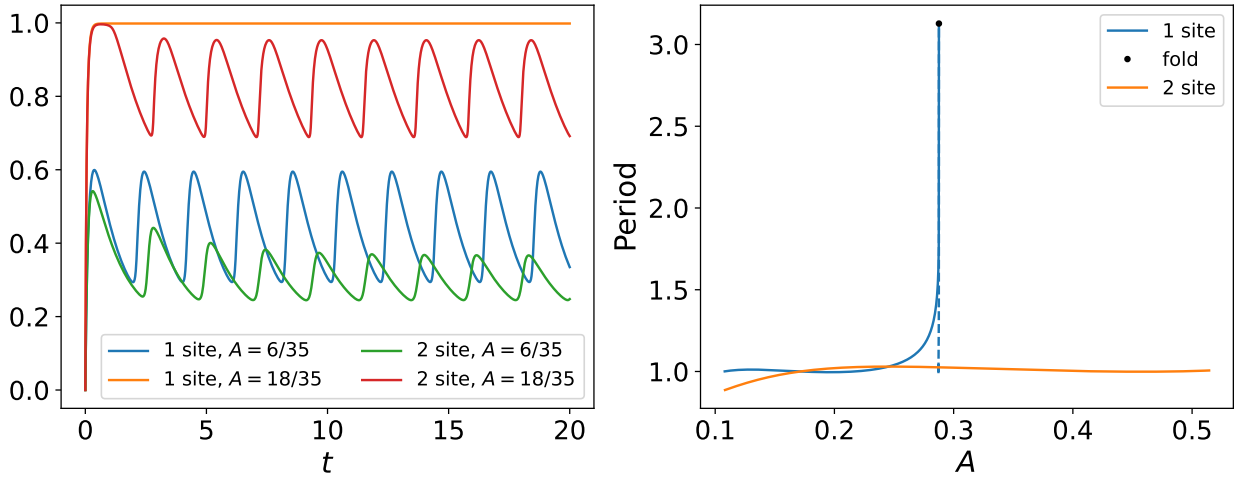


Figure 3.14: KaiA dependence of the period of the single-site minimal model and the double-site minimal model. The dynamics are for the specific parameter choices $k_{01} = 650, k_{12} = 600, k_{su} = 1, k_{us} = 40, k_{20} = 5, k_{d01} = 650, k_{d12} = 200, k_{ds} = 1, k_{sd} = 1, k_{d20} = 60$. Oscillations in the single-site minimal model fail at $A \approx 0.2879$ by a fold of limit cycles, whereas the double-site minimal model exhibits robust oscillations throughout the entire physically relevant range of A. The periods in the right panel are reported relative to the periods of each respective model at $A = 6/35$

3.3.7 Alternative models

The high connectivity of the network shown in Figure 3.1 results in many different mechanisms (i.e., paths in the reaction network) for which oscillations are possible. Most of these alternative mechanisms are not qualitatively different in that they can be converted into each other by a combination of relabeling the species and contracting or subdividing reactions in the network. A specific instance of this arises when reducing the full model to the S only model (Figure 3.6). The sequence of reactions between the bottom S panel and the U panel imposes a delay such that KaiC cannot immediately resequester KaiA after it has been released from CI. This entire loop can be substituted with the single reaction connecting the [|ADP] and [KaiA|ADP] states in the bottom S panel, effectively contracting the sequence of five reactions into a single reaction. Although the precise shape of the orbit varies with the number of reactions in the sequence, which affects the timing of the delay, the resulting models have the same qualitative behavior.

A more significant, but subtle difference arises in how D acts as a sink for KaiA. In the majority of sampled parameters, KaiA very weakly inhibits the formation of KaiB–KaiC complexes by D, the affinity for KaiA by the CI domain of D is high, and D forms at an intermediate rate. Although we have included terms in the prior that encourage the rate of KaiA release from CI to be higher for D when KaiA is bound to CII than when KaiA is not, parameters for which both rates are low are still compatible with the prior. In this regime, D acts less like a parallel sink for KaiA that takes over at KaiA concentrations when S can no longer function, and more like a buffer for KaiA. Higher total KaiA concentrations lead to a higher steady state concentration of D, which then sequesters enough KaiA to keep the pool of KaiA that can interact with S small enough for S to function. This mechanism is less consistent with the results in [89], which shows that the inhibitory effect of CII KaiA binding on KaiB–KaiC complex formation is indeed weaker for D, but not to the extent suggested by the sampling.

The alternative, more realistic mechanism is that KaiA has a smaller but still noticeable inhibitory effect on KaiB–KaiC complex formation by D than on KaiB–KaiC complex formation by S, and that the rates of KaiA sequestration, desequestration, CII KaiA binding, and D formation are all balanced such that D replaces S as the primary sink for KaiA when nearing KaiA concentra-

tions that cause S to fail. This reasoning inspired the double-site minimal model (3.9). Compared to the more commonly sampled mechanism, this mechanism requires a stricter hierarchy of scales. Additional data that enters through either the prior or the likelihood could increase the proportion of sampled parameter sets that correspond to this mechanism.

3.4 Discussion

In this study, we have constructed a mechanistically agnostic, monomeric model of the Kai oscillator that we optimized to reproduce a wide range of experimental observations. Central to the function of our model is the ability for KaiA to promote its own release from the CI domain of KaiC. This mechanism has direct experimental support from [89], which shows that increasing the total concentration of KaiA promotes disassembly of the KaiB-KaiC complexes, even when the phosphorylation state of KaiC is kept fixed by phosphomimetic mutants. To our knowledge, this mechanism has not received significant attention in previous computational studies. From the full model (Figure 3.1), we obtain a reduced single-site minimal model that only possesses the S431 phosphorylation site (Figure 3.9). The single-site minimal model admits an asymptotic analysis in which the limit cycle can be divided into two slow segments, corresponding to the dephosphorylation and phosphorylation phases of the clock.

3.4.1 Relation to previous models of the Kai oscillator

Models of the Kai oscillator fall broadly into two families, distinguished by whether oscillations rely on positive feedback [54]. In monomeric models descended from [117], KaiA suppresses the formation of the KaiC state that acts as its inhibitor, resulting in a positive feedback loop that is essential for oscillations in these models. In allosteric models of the KaiC hexamer descended from [130], including the thermodynamically consistent model by [108], a sufficiently long delay between the phosphorylation-driven conformational switch and subsequent KaiA sequestration is enough to sustain oscillations without any positive feedback term.

Our models contain a variant of the positive feedback mechanism, in which KaiA promotes

its own release through the disruption of KaiB–KaiC complexes. If detailed balance is assumed for all reactions other than ATP hydrolysis, an assumption also made by [108] to construct their model, then the disruption of KaiB–KaiC complexes and the inhibition of the formation of KaiB–KaiC complexes are two segments of the same thermodynamic cycle governed by a single free energy parameter. Both of these mechanisms are present in the model by [108]. Following [130], their model assumes that all monomers within a hexamer undergo a concerted conformation switch, in the spirit of the Monod–Wyman–Changeux model. This same switching behavior is proposed by [89] as the source of robustness to KaiA concentrations in their model. Separately, [108] also includes a term proportional to the number of KaiA dimers bound to CII in their hexamer conformation free energy, which provides the capacity for KaiA to promote its own release from CI in their model. However, the model by [108] does not exhibit positive feedback due to a $10.0k_BT$ free energy penalty for KaiA binding to the CII domain of inactive KaiC and free energy penalties of $2.0k_BT$ and $1.0k_BT$ for KaiA binding per S monomer and D monomer, respectively. These penalties collectively result in a penalty on the order of $20k_BT$ for inactive hexamers, which tend to be strongly phosphorylated on S4321. A penalty of this size suppresses the binding rate of KaiA to CII by factor of about 5×10^{-8}. This is difficult to reconcile with the fluorescence polarization measurements by [90], which shows that KaiB–KaiC complexes formed when incubating KaiB with the KaiC S431E;T432E phosphomimetic mutant disassemble upon addition of KaiA.

Our explanation for robustness does not rely on cooperative, ultrasensitive switching of the KaiC hexamer. We show in (3.7) and (3.4) that the duration of the dephosphorylation and phosphorylation segments scale as $\log(A)$ and $\log(A_+ - A)$, respectively. The greater sensitivity of the phosphorylation segment to the total KaiA concentration can be exploited to achieve period robustness in the single-site minimal model by tuning parameters such that the duration of the dephosphorylation segment dominates the total period. However, the hierarchy of scales required for robustness in the single-site model results in unrealistic solutions. Restoring the T432 phosphorylation site results in a double-site minimal model (Figure 3.13) that exhibits robust oscillations over a wide range of KaiA concentration without the extreme scale separations required by the single-site minimal model. Our mechanism need not be mutually exclusive with ultrasensitive switching of the

hexamer concentration, as our model only models KaiC at the monomeric level and cannot demonstrate ultrasensitive switching. We do not expect that our minimal models quantitatively reproduce the dynamics of experimental system, but we expect that the qualitative behavior of the minimal models is accurate. The two key assumptions made in the minimal model, that the unbinding rate of KaiA from CI is large when KaiA is also bound on CII, and that the sequestration rate of KaiA by fully formed KaiB–KaiC complexes is high, are well supported by our sampling results (Figure 3.5). The other reactions in the minimal models represent longer loops in the full model that have been contracted. The rate constants of the members of these loops tend to be distributed over several orders of magnitude, so the assumed scales of the slow reactions in the reduced models are at least compatible with our sampling results.

3.4.2 Mixed positive/negative feedback and evolvability

A general argument for combining positive and negative feedback in a biochemical oscillator was made by [126]: a pure negative feedback oscillator's frequency and amplitude are difficult to tune independently, whereas adding positive feedback allows a wide range of tunable frequency with comparatively constant amplitude. This property of mixed positive/negative feedback oscillators seems to confer no obvious advantage for circadian oscillators, whose frequencies are fixed at one cycle per day. Noting that the cyanobacterial clock nonetheless appears to rely on positive feedback, they propose and test an alternative advantage: that mixed positive/negative feedback permits oscillations over a wider range of concentrations and parameters, making the oscillator both easier to evolve and more tolerant of imperfect components. Their Monte Carlo sampling of the parameters of three minimal architectures found oscillatory parameter sets in 0.36% of random draws for a triple negative feedback loop alone, rising to 2.1% with weak added positive feedback and 5.1% with strong positive feedback, while adding a second negative loop instead lowered the yield to 0.19% and 0.086%. This frames the comparison in the same terms by which robustness and evolvability are usually discussed. In [4], a distinction is made between robustness that requires fine tuning of specific parameter values from robustness that follows from the qualitative structure of a network across a broad region of parameter space; the former is fragile, since most perturbations move the system

out of the narrow supporting regime, while the latter is relatively insensitive to mutations. The general relationship between this kind of tolerance to mutations and a system's capacity for gradual evolutionary change is developed in [32]. The Monte Carlo results by [126] are a quantification of this distinction applied to oscillator topologies. All of our models fall under the mixed positive/negative feedback class and are not fine tuned in the strictest sense since the oscillations only require their parameters to be ordered, but our models with the T432 phosphorylation site present allow for robust oscillations without pushing any parameters to unrealistic extremes.

3.4.3 Broader connections to models of chemical oscillators

The term describing KaiA release that appears in our minimal models is mathematically identical to autocatalytic reactions that appear in models of chemical oscillators, which places this model within a much older tradition of minimal models of chemical oscillators built around an autocatalytic step inhibited by a slower, delayed feedback loop. The clearest point of comparison is the Oregonator model [41] of the Belousov–Zhabotinsky reaction [8], whose defining step, BrO_3^- reacting with $HBrO_2$ to produce more $HBrO_2$, contributes a term to the rate equation for $HBrO_2$ that is linear in $HBrO_2$ itself: mathematically the same structure as a promoting its own release. In both systems, this autocatalytic step is inhibited by a second, more slowly varying species (Br^- in the Oregonator, s_0 in our single-site minimal model) that builds up gradually and eventually shuts off the fast process, resetting the cycle. Our model differs from most classical models of chemical oscillators in that most chemical oscillators are formulated as open systems with buffered reservoir species that are not conserved together with the oscillating intermediates, whereas in our models the total KaiC and total KaiA are conserved. The closer chemical analogy of KaiC and KaiA is not to the BrO_3^-/malonate substrate pool of the BZ reaction, which is consumed over the course of a batch reaction, but to the BZ reaction's catalyst (cerium or ferroin), which cycles between oxidation states while its total amount stays fixed. ATP, which is hydrolyzed to provide the thermodynamic driving force for the KaiC conformational cycle, plays the role that the substrate pool plays in BZ chemistry.

This shared structure carries over into shared mathematical tools. The two-variable reduction of the Oregonator [129] is a standard example in singular perturbation theory of how a limit cycle can

be decomposed into slow segments connected by fast transitions, the same decomposition strategy used in our asymptotic treatment of the dephosphorylation and phosphorylation segments. It is also a standard example of how the specific route to oscillator failure (ordinary Hopf, canard explosion, or homoclinic/SNIC type divergences) depends sensitively on where in parameter space the model is relative to the folds of the underlying slow manifold. The geometric (Fenichel theory) methods developed in this literature can be applied directly to our reduced equations to distinguish these cases.

3.5 Methods

In this section, we describe the parameter sampling procedure for the full model, which closely follows [28]. We impose priors to keep the parameters consistent with existing, often qualitative, understanding of the Kai system. Even then, only a small fraction of parameter sets result in oscillations. Because limit cycles are computationally costly to evaluate, we first scan the space to identify parameter sets likely to result in oscillations. We do so in three stages: we first sample fixed points, then narrow these to Hopf points, which lie at the boundaries of oscillatory regions, and finally require the fixed point to be unstable at each condition at which oscillations are observed, using the imaginary parts of the associated eigenvalues as surrogates for the period to enforce consistency across conditions. We initiate searches for limit cycles from the last set of fixed points.

During sampling, we nondimensionalize the model. All concentrations are taken relative to the total KaiC concentration, so the population of all KaiC states sums to the dimensionless constant 1. All first-order rate constants are taken relative to the rate constant for CII ATP hydrolysis by the U phosphoform. In doing so, we assume that the CII ATP hydrolysis reaction by the U phosphoform always occurs with nonzero rate. All second-order rate constants are multiplied by the total KaiC concentration and divided by the rate constant for CII ATP hydrolysis by the U phosphoform. We sample the natural logarithm of the nondimensionalized reaction constants.

Using these dimensionless coordinates, the two conservation laws of the system are

$$\begin{aligned} 1 &= \sum y[\,\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\,] \\ A &= a + \sum y[\,\cdot\ \cdot\ \cdot\ \cdot\ 1\ \cdot\,] + \frac{1}{6}\sum y[\,\cdot\ \cdot\ 1\ \cdot\ \cdot\ \cdot\,], \end{aligned} \tag{3.11}$$

where A is the ratio of KaiA dimer to KaiC monomer, and the sum runs over all indices consistent with the wildcards $(\cdot)$. We enforce these constraints for a particular A by solving

$$\begin{aligned} y[0\ 0\ 0\ 0\ 0\ 0] &= 1 - \sum\{y[\,\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\,] \neq y[0\ 0\ 0\ 0\ 0\ 0]\} \\ a &= A - \sum y[\,\cdot\ \cdot\ \cdot\ \cdot\ 1\ \cdot\,] - \frac{1}{6}\sum y[\,\cdot\ \cdot\ 1\ \cdot\ \cdot\ \cdot\,]. \end{aligned} \tag{3.12}$$

3.5.1 Choice of priors

Here, we detail the log prior, which is added to the log likelihood to obtain the log posterior. To this end, we define the one-sided quadratic penalty

$$h(\mathbf{k}_1, \mathbf{k}_2) \equiv \|\mathbb{1}\{\mathbf{k}_1 > \mathbf{k}_2\} \odot (\mathbf{k}_1 - \mathbf{k}_2)\|^2, \tag{3.13}$$

where $\mathbb{1}\{\mathbf{k}_1 > \mathbf{k}_2\}$ compares $\mathbf{k}_1$ and $\mathbf{k}_2$ componentwise and returns a vector with components equal to 1 wherever $\mathbf{k}_1 > \mathbf{k}_2$ and 0 otherwise, and $\odot$ denotes componentwise multiplication between vectors.

To bound the parameter space, we subtract

$$50h(\mathbf{k}, \mathbf{k}_{\max}) + 50h(\mathbf{k}_{\min}, \mathbf{k}), \tag{3.14}$$

where $\mathbf{k}_{\max}$ and $\mathbf{k}_{\min}$ are vectors of upper and lower bounds for $\mathbf{k}$, a vector of the natural logarithm of the nondimensionalized rate constants. We set $\mathbf{k}_{\max}$ and $\mathbf{k}_{\min}$ to constant vectors with components

all equal to 7 and -7, respectively. To account for the results of [103], we subtract

$$\begin{aligned}
&50\sum h\{k[(\,0\;0\;\cdot\;0\;0\;0\,)\to(\,1\;0\;\cdot\;1\;0\;0\,)],\ k[(\,0\;0\;\cdot\;0\;0\;0\,)\to(\,0\;1\;\cdot\;1\;0\;0\,)]\}+\\
&50\sum h\{k[(\,1\;0\;\cdot\;1\;0\;0\,)\to(\,0\;0\;\cdot\;0\;0\;0\,)],\ k[(\,0\;1\;\cdot\;1\;0\;0\,)\to(\,0\;0\;\cdot\;0\;0\;0\,)]\}+\\
&50\sum h\{k[(\,1\;1\;\cdot\;1\;0\;0\,)\to(\,0\;1\;\cdot\;0\;0\;0\,)],\ k[(\,1\;1\;\cdot\;1\;0\;0\,)\to(\,1\;0\;\cdot\;0\;0\;0\,)]\}+\\
&50\sum h\{k[(\,1\;0\;\cdot\;0\;0\;0\,)\to(\,1\;1\;\cdot\;1\;0\;0\,)],\ k[(\,0\;1\;\cdot\;0\;0\;0\,)\to(\,1\;1\;\cdot\;1\;0\;0\,)]\},
\end{aligned}\tag{3.15}$$

where $k[a\to b]$ is the natural logarithm of a rate constant associated with a reaction in Table 3.1. The value of the wildcards in each term must match. These terms reflect the observation that phosphorylation on T432 is faster than phosphorylation on S431 when both sites are unphosphorylated, but phosphorylation on S431 inhibits phosphorylation on T432. Similarly, since it is observed in [89] that binding of KaiA to the CII domain of KaiC allosterically inhibits binding of KaiB to CI, we subtract

$$\begin{aligned}
&50\sum h\{k[(\,\cdot\;\cdot\;1\;\cdot\;0\;1\,)\to(\,\cdot\;\cdot\;1\;\cdot\;1\;1\,)],\ k[(\,\cdot\;\cdot\;0\;\cdot\;0\;1\,)\to(\,\cdot\;\cdot\;0\;\cdot\;1\;1\,)]\}+\\
&50\sum h\{k[(\,\cdot\;\cdot\;0\;\cdot\;1\;1\,)\to(\,\cdot\;\cdot\;0\;\cdot\;0\;1\,)],\ k[(\,\cdot\;\cdot\;1\;\cdot\;1\;1\,)\to(\,\cdot\;\cdot\;1\;\cdot\;0\;1\,)]\}.
\end{aligned}\tag{3.16}$$

These terms favor a lower rate of KaiA binding and higher rate of KaiA unbinding on the CI domain when KaiA is bound to the CII domain. Since the nucleotide exchange rate on the CII has not been experimentally observed to depend on the nucleotide and KaiA binding state of CI, we subtract the term

$$\begin{aligned}
&1000\,\mathrm{Var}\{k_a[(\,1\;0\;1\;1\;\cdot\;\cdot\,)\to(\,1\;0\;1\;0\;\cdot\;\cdot\,)],\ k_a[(\,1\;0\;1\;0\;\cdot\;\cdot\,)\to(\,1\;0\;1\;1\;\cdot\;\cdot\,)]\}+\\
&1000\,\mathrm{Var}\{k_a[(\,1\;1\;1\;1\;\cdot\;\cdot\,)\to(\,1\;1\;1\;0\;\cdot\;\cdot\,)],\ k_a[(\,1\;1\;1\;0\;\cdot\;\cdot\,)\to(\,1\;1\;1\;1\;\cdot\;\cdot\,)]\}.
\end{aligned}\tag{3.17}$$

This encourages the rate of nucleotide exchange on CII between KaiC states with the same phosphorylation state to be similar regardless of the state of CI in the absence of additional evidence. Likewise, since the ATP hydrolysis rate by the CII domain of KaiC has not been experimentally

observed to depend on KaiA binding, we subtract

$$500 \sum \{k[(\cdot\ \cdot\ 0\ 0\ \cdot\ \cdot) \to (\cdot\ \cdot\ 0\ 1\ \cdot\ \cdot)] - k[(\cdot\ \cdot\ 1\ 0\ \cdot\ \cdot) \to (\cdot\ \cdot\ 1\ 1\ \cdot\ \cdot)]\}^2. \tag{3.18}$$

This encourages the rate of ATP hydrolysis on the CII domain of different states of KaiC to have similar values in the absence of additional evidence. The log prior is the sum of the negative of 3.14–3.18, and we use it in the loss governing all sampling simulations. The regularization weights were chosen so that the restraints can be violated if doing so can significantly improve the agreement with experimental measurements. For example, a relatively low regularization weight of 50 is used in (3.15) since the rates involved can also be determined by the likelihood term (3.23), whereas the regularization weights in (3.17) and (3.18) are larger to reflect the lack of measurements that would provide direct evidence for or against significant differences between the rate constants involved.

3.5.2 Sampling fixed points

In the first stage, we sample fixed points of the model when the KaiC to KaiA ratio is 6/35 and at 100% ATP. That is, using the binary string notation in Table 3.1, we sample $(\mathbf{y}, \mathbf{k})$ with the constraints that

$$\frac{d\mathbf{y}}{dt} = \mathbf{f}(\mathbf{y}, \mathbf{k}) = \mathbf{0}. \tag{3.19}$$

Additionally, we require that the location of the fixed point $\mathbf{y}$ is consistent with the average fractions of pT, pS, and pD phosphoforms reported in [111]. These measurements are incorporated into the log likelihood by subtracting

$$\begin{aligned} &1000\Big(\bar{x}_{\mathrm{pT}} - \sum y[0\ 1\ \cdot\ \cdot\ 0\ 0]\Big)^2 + \\ &1000\Big(\bar{x}_{\mathrm{pS}} - \sum y[1\ 0\ \cdot\ \cdot\ \cdot\ \cdot]\Big)^2 + \\ &1000\Big(\bar{x}_{\mathrm{pD}} - \sum y[1\ 1\ \cdot\ \cdot\ \cdot\ \cdot]\Big)^2, \end{aligned} \tag{3.20}$$

where $\bar{x}$ is the average of the experimentally measured fraction of all KaiC monomers that a particular KaiC phosphoform occupies. Using a step size of 10^{-2}, friction coefficient of 10^{-1}, and

unit mass matrix, we sample 10 independent chains for 10^7 steps each for a total of 10^8 sampling steps.

3.5.3 Sampling Hopf points

After sampling fixed points, we sort the fixed points by

$$g_i = \min_j \left| \frac{\mathrm{Re}(\zeta_{ij})}{\mathrm{Im}(\zeta_{ij})} \right|, \tag{3.21}$$

where ζ_{ij} is the jth eigenvalue of the Jacobian of the ODE model evaluated at the ith fixed point sampled, Re denotes the real part of a complex number, and Im denotes the imaginary part of a complex number. We use g_i as a heuristic for finding fixed points that are close to a Hopf bifurcation; at a Hopf bifurcation, the Jacobian has a pair of purely imaginary eigenvalues, which gives $g_i = 0$. Although small g_i does not guarantee that the eigenvalue pair is near a crossing, we can verify by using these points as initial guesses to solve for the exact conditions for a Hopf bifurcation to occur: starting from $(\mathbf{y}_0, \mathbf{k}_0, \mathbf{v}_0, \omega)$, we apply Gauss–Newton iterations to the constraints

$$\begin{aligned} \mathbf{f}(\mathbf{y}, \mathbf{k}) &= \mathbf{0} \\ \mathbf{J}(\mathbf{y}, \mathbf{k})\,\mathrm{Re}(\mathbf{v}) + \omega\,\mathrm{Im}(\mathbf{v}) &= \mathbf{0} \\ \mathbf{J}(\mathbf{y}, \mathbf{k})\,\mathrm{Im}(\mathbf{v}) - \omega\,\mathrm{Re}(\mathbf{v}) &= \mathbf{0} \\ \mathrm{Re}(\mathbf{v})^{\mathsf{T}}\,\mathrm{Re}(\mathbf{v}) + \mathrm{Im}(\mathbf{v})^{\mathsf{T}}\,\mathrm{Im}(\mathbf{v}) &= 1, \end{aligned} \tag{3.22}$$

where $\mathbf{J}(\mathbf{y}, \mathbf{k})$ is the matrix of derivatives of $\mathbf{f}$ with respect to the components of $\mathbf{y}$, $\omega \equiv \mathrm{Im}(\zeta^*)$, and ζ^* is an eigenvalue with eigenvector $\mathbf{v}_0$ corresponding to parameters that produce an appropriately small g_i. We note that (3.22) usually contains a constraint to fix the complex phase of $\mathbf{v}$ and ω. We omit this constraint since our choice of Gauss–Newton solver resolves this phase degeneracy by returning a minimum norm solution. Solving (3.22) gives $(\mathbf{y}, \mathbf{k}, \mathbf{v}, \omega)$, where $\mathbf{y}$ is now the location of a Hopf point (a fixed point with a pair of purely imaginary eigenvalues), $\mathbf{k}$ are the corresponding rate constants for which the Hopf point exists, $\mathbf{v}$ is the eigenvector associated with imaginary eigenvalue $i\omega$. We sample $\mathbf{k}$ that satisfy the constraints (3.22) as a surrogate for parameters at which

the model exhibits oscillatory dynamics, since supercritical Hopf bifurcations lead to the onset of stable oscillations and solving the constraint equations (3.22) is much cheaper computationally than integrating the system of ODEs forward in time.

When sampling Hopf points, we require that the dephosphorylation dynamics of the model agree with the measurements reported in [111]. Experimentally, when KaiC is allowed to dephosphorylate in the absence of KaiA and KaiB, the time-dependent fraction of each KaiC phosphoform is measured over the course of one period of the full oscillator. We model this by considering the dynamics of the subsystem consisting of only the reactions between the subset of KaiC species that match the indices ($\cdot$ $\cdot$ 0 $\cdot$ 0 0). We incorporate the dephosphorylation data into the likelihood by subtracting from the log likelihood

$$\begin{aligned} &1500\left(\int_0^\tau (x_{\text{dephos,pT}}(t) - \sum y_{\text{dephos}}[0\ 1\ 0\ \cdot\ 0\ 0](t)^2 dt\right) + \\ &1500\left(\int_0^\tau (x_{\text{dephos,pS}}(t) - \sum y_{\text{dephos}}[1\ 0\ 0\ \cdot\ 0\ 0](t)^2 dt\right) + \\ &1500\left(\int_0^\tau (x_{\text{dephos,pD}}(t) - \sum y_{\text{dephos}}[1\ 1\ 0\ \cdot\ 0\ 0](t)^2 dt\right), \end{aligned} \tag{3.23}$$

where x_{dephos} is the experimentally measured time dependent KaiC phosphoform fraction and y_{dephos} is the population of a particular KaiC monomer in the model of the dephosphorylation subsystem. Since the dephosphorylation subsystem is linear, we evaluate $\mathbf{y}_{\text{dephos}}$ using the matrix exponential of the Jacobian of the subsystem. We assume that that the CII nucleotide binding site is fully occupied by ADP at time 0, so that

$$\begin{aligned} y_{\text{dephos}}[\ \cdot\ \cdot\ 0\ 1\ 0\ 0](0) &= x_{\text{dephos}}(0) \\ y_{\text{dephos}}[\ \cdot\ \cdot\ 0\ 0\ 0\ 0](0) &= \mathbf{0}. \end{aligned} \tag{3.24}$$

We approximate the period of the full oscillatory system as $\tau = 2\pi/\omega$ and rescale time such that τ corresponds to 24 hours in the experimental data. We then evaluate the integral in (3.23) numerically by the trapezoidal rule, taking the rescaled times of the data points as the quadrature nodes. We apply (3.23) in addition to the prior and likelihood terms detailed in Section 3.5.2. Using a step size of 10^{-2}, friction coefficient of 10^{-1}, and unit mass matrix, we sample 10 independent chains for

10^7 steps each for a total of 10^8 sampling steps.

3.5.4 Surrogate sampling for oscillations that are robust to variations in KaiA and ATP concentrations

After identifying parameters for which a Hopf point exists, we sample parameters subject to the constraints

$$\begin{aligned} \mathbf{f}(\mathbf{y}_c, \mathbf{k}) &= \mathbf{0} \\ \mathbf{J}(\mathbf{y}_c, \mathbf{k})\,\mathrm{Re}(\mathbf{v}_c) - \mathrm{Re}(\zeta_c)\,\mathrm{Re}(\mathbf{v}_c) + \mathrm{Im}(\zeta_c)\,\mathrm{Im}(\mathbf{v}_c) &= \mathbf{0} \\ \mathbf{J}(\mathbf{y}_c, \mathbf{k})\,\mathrm{Im}(\mathbf{v}_c) - \mathrm{Re}(\zeta_c)\,\mathrm{Im}(\mathbf{v}_c) - \mathrm{Im}(\zeta_c)\,\mathrm{Re}(\mathbf{v}_c) &= \mathbf{0} \\ \mathrm{Re}(\mathbf{v}_c)^{\mathsf{T}}\,\mathrm{Re}(\mathbf{v}_c) + \mathrm{Im}(\mathbf{v}_c)^{\mathsf{T}}\,\mathrm{Im}(\mathbf{v}_c) &= 1 \\ \mathrm{Re}(\zeta_c) &> 0. \end{aligned} \tag{3.25}$$

The index c represents a particular combination of total KaiA and %ATP for an oscillatory condition in Table 3.3. $\mathbf{y}_c$ is the location of a fixed point for a particular c, ζ_c is a Jacobian eigenvalue of $\mathbf{f}$ evaluated at $\mathbf{y}_c$, and $\mathbf{v}_c$ is its corresponding eigenvector. We assume that we always sample the eigenvalue with positive imaginary part in the pair of complex eigenvalues and use the natural logarithm of $\mathrm{Im}(\zeta_c)$. Since (3.22) enforces $\mathrm{Re}(\zeta_c) = 0$ the results obtained from sampling with the constraint (3.22) are not directly compatible with (3.25). We perform an initialization run with 10 independent chains for 10^6 steps each using with the soft restraint

$$1000h(5 \times 10^{-2}, \mathrm{Re}(\zeta_c)). \tag{3.26}$$

Once we find parameters for which $\mathrm{Re}(\zeta_c) > 0$ holds for all c, we enforce the inequality constraint on ζ_c by sampling $\log(\mathrm{Re}(\zeta_c))$. When sampling fixed points at each c, we include the log likelihood term in (3.20) for c at which there are available data [111]: $c = (0.6, 100\%)$ and $c = (0.6, 25\%)$. To reflect the observation that the period of the oscillator changes by at most 10% across the oscillatory conditions in Table 3.3, we penalize variance in $\mathrm{Im}(\zeta_c)$, which we use as a surrogate for the period,

by subtracting from the log likelihood

$$1000\,\mathrm{Var}(\mathrm{Im}(\zeta_c)). \tag{3.27}$$

The log likelihood terms detailed in Sections 3.5.2 and 3.5.3 are also present. To control the magnitude of ζ_c, we subtract from the log prior

$$\sum_c |\zeta_c|. \tag{3.28}$$

Using a step size of 5×10^{-3}, friction coefficient of 10^{-1}, and unit mass matrix, we sample 40 independent chains for 10^7 steps each for a total of 4×10^7 sampling steps.

3.5.5 Sampling limit cycles

The points sampled in Section 3.5.4 provide starting points from which we initiate sampling that accounts for the detailed time-dependent dynamics of the oscillator. We sample parameters subject to the constraint that there exists a periodic solution to the ODE

$$\begin{aligned} \frac{d\mathbf{y}}{dt} &= \mathbf{f}(\mathbf{y}, \mathbf{k}) \\ \mathbf{y}(0) &= \mathbf{y}(\tau). \end{aligned} \tag{3.29}$$

This boundary value problem imposes a continuous constraint. For numerical evaluation, we discretize (3.29) with fourth-order Gauss–Legendre collocation a mesh of size $N = 61$ and approximate the solution to (3.29) on the space of piecewise polynomials

$$\mathcal{P}_{m,\Delta t} = \left\{ \hat{\mathbf{y}} \in C[0,\tau] \;\middle|\; \hat{\mathbf{y}}_{[t_{i-1},t_i]} \in \mathcal{P}_m \right\}. \tag{3.30}$$

That is, $\mathcal{P}_{m,\Delta t}$ consists of continuous functions $\hat{\mathbf{y}}$ on the interval $[0,\tau]$ such that on each mesh subinterval $[t_{i-1}, t_i]$, $\hat{\mathbf{y}}$ is a polynomial of up to order m. We demand that on each subinterval, $\hat{\mathbf{y}}$ satisfies

$$\hat{\mathbf{y}}'(z_{i,j}) = \mathbf{f}(\hat{\mathbf{y}}(z_{i,j}), \mathbf{k}), \quad j \in \{0, \ldots, m-1\}, \quad i \in \{0, \ldots, N-1\}, \tag{3.31}$$

where $\{z_{i,j}\}$ are the collocation points and $\hat{\mathbf{y}}'(t)$ denotes the derivative of $\hat{\mathbf{y}}$ with respect to t. We choose the collocation points to be the scaled and shifted roots of the m-th Legendre polynomial. Our implementation constructs $\hat{\mathbf{y}}$ with Newton polynomials on each mesh subinterval. As in Section 3.5.4, we impose these constraints for solutions at each oscillatory condition listed in Table 3.3.

To avoid identifying constant solutions, which are trivially periodic, we subtract a term from the log prior that penalizes orbits with small arclength:

$$\frac{1}{4}\left(\frac{L_0^2}{L_c^2}-1\right)^2 \mathbb{1}\{L < L_0\}, \tag{3.32}$$

where the threshold $L_0 = 0.3$ and the arclength L_c is computed as

$$L_c = \int_0^\tau \|\hat{\mathbf{y}}_c'(t)\| dt. \tag{3.33}$$

We control the period of the solutions by subtracting from the log likelihood

$$1000\frac{\mathrm{Var}(\tau_c)}{\mu(\tau_c)^2}, \tag{3.34}$$

where $\mu(\cdot)$ denotes the arithmetic mean. Data for total KaiA and ATP fraction (0.6,1.0) and (0.6,0.25) are processed as described in the caption to Figure 3.2 and are included in the log likelihood term

$$\begin{aligned} &1500\left(\int_0^\tau (x_{\mathrm{pT}}(t) - \sum \hat{y}_c[0\ 1\ \cdot\ \cdot\ 0\ 0](t)^2 dt\right)+ \\ &1500\left(\int_0^\tau (x_{\mathrm{pS}}(t) - \sum \hat{y}_c[1\ 0\ \cdot\ \cdot\ \cdot\ \cdot](t)^2 dt\right)+ \\ &1500\left(\int_0^\tau (x_{\mathrm{pD}}(t) - \sum \hat{y}_c[1\ 1\ \cdot\ \cdot\ \cdot\ \cdot](t)^2 dt\right). \end{aligned} \tag{3.35}$$

Although there are no data for the KaiA and ATP conditions $c = (1.2, 1.0)$ and $c = (1.8, 1.0)$ reported in [111], experiments at these conditions were performed in [89], and these experiments show that when the KaiA dimer concentration is increased from 0.6 μM to 1.8 μM, the fraction of D increases linearly up to an additional 0.1 at 1.8 μM KaiA dimer while the fraction of the other KaiC phosphoforms does not show any clear trend. The data from [89] are not regular enough to be

processed with the segmentation procedure that we apply to the data from [111], so we substitute synthetic data generated by shifting the fraction of D in the data from [111] for the condition $(0.6, 1.0)$ linearly with the KaiA dimer concentration. These synthetic data are included in a log likelihood term with the same form as (3.35), but with the weight reduced from 1500 to 150 to reflect their reduced reliability.

When the reaction mixture contains only KaiA and KaiC but not KaiB, pS431 and pT432 increase to a steady state on the time scale of one period of the full system [111]. This KaiA-KaiC subsystem contains the subset of reactions between the subset of KaiC species that match the indices $(\,\cdot\,\cdot\,\cdot\,\cdot\,0\;0\,)$. We solve the ODEs corresponding to the KaiA-KaiC subsystem by forward integration using Kvaerno's 4/3 explicit singly diagonally implicit Runge Kutta method and incorporate the results by subtracting from the log likelihood

$$\begin{aligned} &1500\left(\int_0^\tau (x_{\mathrm{AC,pT}}(t) - \sum y_{\mathrm{AC}}[0\;1\;\cdot\;\cdot\;0\;0](t)^2 dt\right)+ \\ &1500\left(\int_0^\tau (x_{\mathrm{AC,pS}}(t) - \sum y_{\mathrm{AC}}[1\;0\;\cdot\;\cdot\;0\;0](t)^2 dt\right)+ \\ &1500\left(\int_0^\tau (x_{\mathrm{AC,pD}}(t) - \sum y_{\mathrm{AC}}[1\;1\;\cdot\;\cdot\;0\;0](t)^2 dt\right) \end{aligned} \tag{3.36}$$

where x_{AC} are experimental measurements for the phosphorylation fraction of KaiC in the KaiA-KaiC subsystem and y_{AC} is the population of a particular KaiC monomer in the model of the KaiA-KaiC subsystem. We assume that KaiC is fully dephosphorylated and the CII nucleotide binding site is occupied by ADP at the start of the experiment, i.e., we take the initial condition for the KaiA-KaiC subsystem to be

$$y[0\;0\;0\;1\;0\;0] = 1. \tag{3.37}$$

As is the case when computing (3.23), we rescale time in (3.35) and (3.36) such that τ corresponds to 24 hours in the experimental data. The integrals are computed numerically by the trapezoidal rule, with the rescaled data collection times as nodes, and the values of y are evaluated by polynomial interpolation of the ODE solution. Using a step size of 10^{-2}, friction coefficient of 10^{-1}, and unit

mass matrix, we sample a total of 1.6×10^6 steps over 68 independent chains. The steps are not uniformly distributed over the chains since in certain parameter regimes the constraint solver fails and causes the sampler to stall.

3.5.6 Experimental methods

Protein expression and purification

KaiA, KaiB, KaiC and KaiC T432V mutant were expressed in *Escherichia coli* and purified as previously described [90].

Fluorescent labeling

KaiB was labeled with a FAM fluorophore using sortase-mediated N-terminal ligation as previously described [26].

Kai oscillator reaction set-up and fluorescence polarization monitoring

A 30 μL reaction mixture consisting of different concentrations of KaiA (0.6 to 3.0 μM), 3.5 μM KaiB, 3.5 μM KaiC-T432V, and FAM-KaiB was loaded into a black non-binding 384-well plate (Corning cat. 3575) and covered with an optical film (Thermo Scientific cat. 235307). The reaction mixture was incubated at 30°C while fluorescence polarization signals being read by a microplate reader (Tecan Infinite F500) every 15 minutes for a total of 250 hours. Filter pairs were used to obtain an excitation wavelength of 485 nm and an emission wavelength of 535 nm.

KaiA injection phase response curve

We first created a set of oscillator reactions with staggered phase. This is achieved by first making a master mix consisting of 1.5 μM KaiA, 3.5 μM KaiB, 3.5 μM KaiC, and FAM-labelled KaiB and incubating it at 4°C for 24 hours, allowing it to reach a steady state at a low temperature which does not allow oscillations. Subsequently, an aliquot of the master mix was transferred to 30°C

(maintained using a thermocycler) every 2 hours during a 12-hour window, initiating oscillations. This procedure yielded a set of seven oscillator reactions with staggered initiation times, whose phases together spanned a 12-hour period. After a 24-hour relaxation period, each reaction in the set was then divided into three aliquots, two of which were perturbed 12 hours apart while the other served as an unperturbed control. Perturbation was applied by injecting purified KaiA from a high-concentration stock (27 μM), equivalent to 25% of the existing amount in the reaction mixture, into each staggered reaction. Fluorescence polarization signals were read from all reactions, including the perturbed ones and the unperturbed controls, for 72 hours under 30°C temperature control. Reaction phases were inferred by fitting a sinusoidal wave given by where zero point of the phase is defined at the trough, to the fluorescence polarization traces. The phase shift of a perturbed reaction is then estimated by subtracting its phase by that of the corresponding control reaction.

3.6 Appendix

3.6.1 Asymptotic expansions of the single-site model

After reducing the S only model by contracting the loops of reactions in Figure 3.6, we obtain the single-site minimal model (3.1). Since we assume that $\mathcal{O}(1) = k_{us} \approx k_{su} \approx k_{20} \ll k_{12} \ll k_{01}$, a natural approach for obtaining an approximate solution to (3.1) is to divide (3.1) through by k_{01}, define $\epsilon \equiv k_{01}^{-1}$ and take $\epsilon \to 0$. Doing so results in the system of equations

$$\frac{ds_2}{dt} = k_{us}ua + k_{12}s_1a - k_{20}s_2 \tag{3.38a}$$

$$0 = s_0a \tag{3.38b}$$

Equation (3.38b) gives either $s_0 = 0$ or $a = 0$. Assuming the case of $a = 0$, s_2 would have the asymptotic solution

$$s_2 \propto \exp(-k_{20}t) \tag{3.39}$$

However, the numerical solution of (3.1) is oscillatory. For any $\epsilon > 0$, (3.39) and the solution to (3.1) will have regions that differ by much more than ϵ; solutions to (3.1) do not converge

uniformly to (3.39) as $\epsilon \to 0$. The asymptotic behavior of (3.1) is singularly perturbed. We use the method of matched asymptotic expansions [79] to obtain approximate solutions to (3.1). The method involves identifying different regions in the limit cycle depending on which terms in (3.1) are dominant, performing a sequence of locally valid expansions in terms of an expansion parameter that may depend on the region being considered, and then enforcing that the expansions of neighboring regions agree within an overlap domain. A specific application of the method of matched asymptotic expansions arises in the WKB approximation for solving the time independent Schrödinger equation. In the following sections we identify different segments of the solution to (3.1) and provide closed form local solutions when analytically tractable. A comparison of the asymptotic and numerical phase portraits for the specific parameter choices $k_{01} = 10^6, k_{12} = 300, k_{su} = k_{us} = k_{20} = 1$ is shown in Figure 3.15

Main dephosphorylation segment

First, we consider the behavior of the numerical solution from $t \approx 0$ to $t \approx 10$ in Figure 3.10. Based on the behavior of the numerical solution in Figure 3.10, we use the variable scalings

$$\begin{aligned} s_0 &= s_{0,1}^{(0)} + \mathcal{O}(\epsilon) \\ s_2 &= s_{2,1}^{(0)} + \mathcal{O}(\epsilon) \\ s_1 &= s_{1,1}^{(0)} + \mathcal{O}(\epsilon) \\ a &= \epsilon a_1^{(0)} + \mathcal{O}(\epsilon^2) \\ u_1^{(0)} &= 1 - s_{0,1}^{(0)} - s_{2,1}^{(0)} - s_{1,1}^{(0)} \end{aligned} \tag{3.40}$$

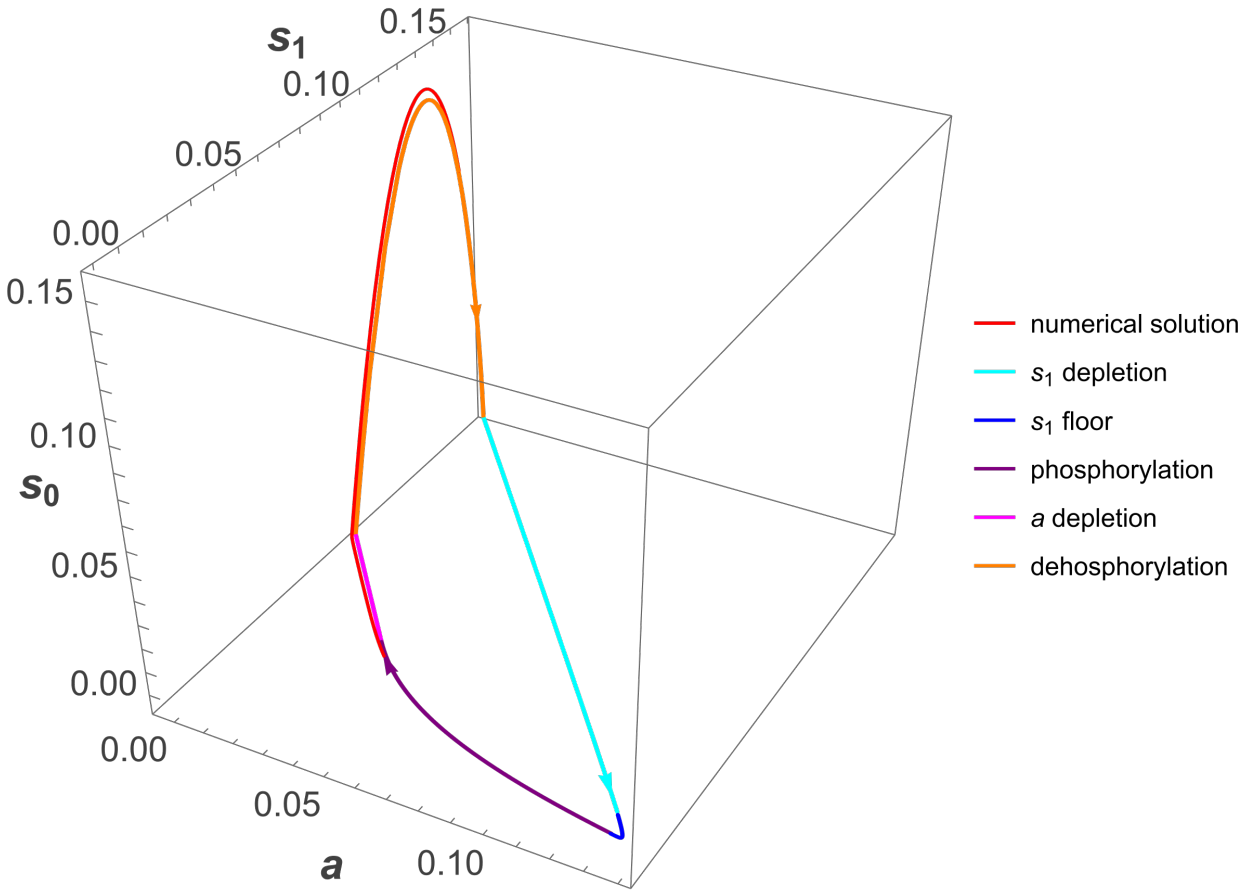


Figure 3.15: The phase portrait of the single-site minimal model (3.1) for the specific parameter choices $k_{01} = 10^6, k_{12} = 300, k_{su} = k_{us} = k_{20} = 1, A = 6/35$ overlaid with asymptotic solutions of the different segments of the limit cycle. The error from the approximation (3.121) causes the asymptotic solution of the dephosphorylation segment to deviate from the numerical solution even though the asymptotic expansion of the dephosphorylation segment is accurate to $\mathcal{O}(\epsilon)$. The segments describing the transition between the dephosphorylation segment and the s_1 depletion segment all occur within a $\mathcal{O}(\delta)$ neighborhood of the point $(a = 0, s_1 = A, s_0 = 0)$ and are not visible on this scale.

where the expansion coefficients $s_{0,1}^{(0)}, s_{2,1}^{(0)}, s_{1,1}^{(0)}, a_1^{(0)}$ are $\mathcal{O}(1)$ and $\epsilon \equiv k_{01}^{-1}$. Substituting (3.40) into (3.1) and retaining terms of $\mathcal{O}(1)$ gives

$$\frac{ds_{0,1}^{(0)}}{dt} = -k_{su}s_{0,1}^{(0)} + k_{20}s_{2,1}^{(0)} - s_{0,1}^{(0)}a_1^{(0)} \tag{3.41a}$$

$$\frac{ds_{2,1}^{(0)}}{dt} = -k_{20}s_{2,1}^{(0)} \tag{3.41b}$$

$$\frac{ds_{1,1}^{(0)}}{dt} = s_{0,1}^{(0)}a_1^{(0)}. \tag{3.41c}$$

From the conservation law (3.2b),

$$\begin{aligned} A &= \epsilon a_1^{(0)} + s_{1,1}^{(0)} + \frac{1}{6}s_{2,1}^{(0)} \\ s_{1,1}^{(0)} &= A - \frac{1}{6}s_{2,1}^{(0)} + \mathcal{O}(\epsilon). \end{aligned} \tag{3.42}$$

Differentiating (3.42) gives

$$\begin{aligned} \frac{d_{1,1}^{(0)}}{dt} &= -\frac{1}{6}\frac{ds_{2,1}^{(0)}}{dt} + \mathcal{O}(\epsilon) \\ s_{0,1}^{(0)}a_1^{(0)} &= \frac{1}{6}k_{20}s_{2,1}^{(0)} + \mathcal{O}(\epsilon). \end{aligned} \tag{3.43}$$

This leads to the quasi steady state relation

$$a_1^{(0)} = \frac{k_{20}s_{2,1}^{(0)}}{6s_{0,1}^{(0)}}. \tag{3.44}$$

Substituting (3.44) into (3.41) then gives the planar system

$$\begin{aligned} \frac{ds_{0,1}^{(0)}}{dt} &= -k_{su}s_{0,1}^{(0)} + \frac{5}{6}k_{20}s_{2,1}^{(0)} \\ \frac{ds_{2,1}^{(0)}}{dt} &= -k_{20}s_{2,1}^{(0)}. \end{aligned} \tag{3.45}$$

Equation (3.45) and the conservation law (3.42) has solutions

$$s_{0,1}^{(0)} = C_1 e^{-k_{su}t} + \frac{5C_0 k_{20}\left(e^{-k_{20}t} - e^{-k_{su}t}\right)}{6(k_{su} - k_{20})} \tag{3.46a}$$

$$s_{2,1}^{(0)} = C_0 e^{-k_{20}t} \tag{3.46b}$$

$$s_{1,1}^{(0)} = A - \frac{s_{2,1}^{(0)}}{6}. \tag{3.46c}$$

In the dephosphorylation segment, $a = \mathcal{O}(\epsilon)$ means that the phosphorylation reaction $k_{us}ua$ does not occur at leading order. Thus s_2 decays exponentially from its initial concentration upon entering the dephopshorylation segment. The conservation law (3.2b) directly relates the concentrations of s_1 and s_2 at leading order and leads to the quasi steady state (3.44). The quasi steady state then reduces the equation for s_0 to a linear first order ODE. Although the $k_{20}s_2$ reaction produces s_0, the large magnitude of k_{01} the 1/6 of an a released by the $k_{20}s_2$ reaction to instantly be sequestered through the $k_{01}s_0a$ reaction which consumes 1/6 of a s_0, so that the $k_{20}s_2$ reaction results in a net gain of 5/6 of a s_0.

Terminal dephosphorylation segment

The quasi steady state (3.44) was derived assuming that $s_0 a \gg \delta s_1 a$, where

$$\delta \equiv \frac{k_{12}}{k_{10}} = \epsilon k_{12}. \tag{3.47}$$

If s_0 falls to $\mathcal{O}(\delta)$, the rates $k_{01}s_0a$ and $k_{12}s_1a$ become comparable in magnitude, and (3.44) loses validity. This at $t \approx 10$ in Figure 3.10, where a remains $\mathcal{O}(\epsilon)$ but s_0 falls to $\mathcal{O}(\delta)$. This prompts the variable scalings

$$\begin{aligned} s_0 &= \delta s_{0,2}^{(0)} \\ s_2 &= \delta s_{2,2}^{(0)} \\ s_1 &= A + \delta s_{1,2}^{(0)} \\ a &= \epsilon a_2^{(0)}. \end{aligned} \tag{3.48}$$

Substituting (3.48) into (3.1) and retaining terms up to $\mathcal{O}(\delta)$ leads to the equations

$$\frac{ds_{0,2}^{(0)}}{dt} = -k_{su}s_{0,2}^{(0)} + k_{20}s_{2,2}^{(0)} - s_{0,2}^{(0)}a_2^{(0)} \tag{3.49a}$$

$$\frac{ds_{2,2}^{(0)}}{dt} = -k_{20}s_{2,2}^{(0)} + Aa_2^{(0)} \tag{3.49b}$$

$$\frac{ds_{1,2}^{(0)}}{dt} = s_{0,2}^{(0)}a_2^{(0)} - Aa_2^{(0)} \tag{3.49c}$$

Substituting (3.48) into the conservation law (3.2b) gives

$$\begin{aligned} A &= \frac{\delta s_{2,2}^{(0)}}{6} + \delta s_{1,2}^{(0)} + \epsilon a_2^{(0)} \\ s_{1,2}^{(0)} &= -\frac{s_{2,2}^{(0)}}{6} + \mathcal{O}(\xi), \end{aligned} \tag{3.50}$$

where

$$\xi \equiv \frac{1}{k_{12}} = \frac{\delta}{\epsilon}. \tag{3.51}$$

Differentiating (3.50) gives

$$\begin{aligned} \frac{ds_{1,2}^{(0)}}{dt} &= -\frac{1}{6}\frac{ds_{2,2}^{(0)}}{dt} + \mathcal{O}(\xi) \\ s_{0,2}^{(0)}a_2^{(0)} - Aa_2^{(0)} &= \frac{1}{6}\left(k_{20}s_{2,2}^{(0)} - Aa_2^{(0)}\right) + \mathcal{O}(\xi) \\ a_2^{(0)} &= \frac{k_{20}s_{2,2}^{(0)}}{6s_{0,2}^{(0)} - 5A} + \mathcal{O}(\xi) \end{aligned} \tag{3.52}$$

Since $a \sim \mathcal{O}(\epsilon) \ll \mathcal{O}(\delta)$, a is still at quasi steady state in the terminal dephosphorylation segment, but (3.52) acquires a factor of $-5A$ in the denominator compared to (3.44) to account for the $k_{12}s_1a$ reaction, which now has a rate comparable to $k_{01}s_0a$ and $k_{20}s_2$.

Equation (3.49) is not analytically tractable, but can still be shown to match with the main dephosphorylation segment. In as $t \to -\infty$ back into the main dephosphorylation segment,

$$\lim_{s_{0,2^{(0)}} \to \infty} a_2^{(0)} = \frac{k_{20}s_{2,2}^{(0)}}{6s_{0,2}^{(0)}} \tag{3.53}$$

so that the large negative time behavior of (3.49) is precisely (3.45).

Dephosphorylation QSS departure

Equation (3.52) defines the critical manifold for the fast $a_2^{(0)}$ subsystem. At $s_{0,2}^{(0)} = 5A/6$, the Jacobian of the critical manifold is

$$\xi \frac{\partial}{\partial a_2^{(0)}} \frac{da_2^{(0)}}{dt} = \frac{5A}{6} - s_{0,2}^{(0)}. \tag{3.54}$$

When $s_{0,2}^{(0)} < 5A/6$, the critical manifold is unstable and the quasi steady state given by (3.52) is repelling. The behavior of the system near this point can be mapped onto a dynamic fold bifurcation, which near the bifurcation has the canonical scalings

$$\begin{aligned}
t &= \xi^{2/3} t_3 \\
s_0 &= \frac{5\delta A}{6} + \delta \xi^{1/3} s_{0,3}^{(0)} + \mathcal{O}\left(\delta \xi^{2/3}\right) \\
s_2 &= \delta s_{2,3}^{(0)} + \delta \xi^{1/3} s_{2,3}^{(1)} + \mathcal{O}\left(\delta \xi^{2/3}\right) \\
s_1 &= A + \delta s_{1,3}^{(0)} + \delta \xi^{1/3} s_{1,3}^{(1)} + \mathcal{O}\left(\delta \xi^{2/3}\right) \\
a &= \delta \xi^{2/3} a_3^{(0)} + \mathcal{O}(\epsilon).
\end{aligned} \tag{3.55}$$

These variable scalings result in the equations

$$\frac{ds_{0,3}^{(0)}}{dt_3} = -\frac{5A a_3^{(0)}}{6} \tag{3.56a}$$

$$\frac{ds_{2,3}^{(0)}}{dt_3} = 0 \tag{3.56b}$$

$$\frac{ds_{2,3}^{(1)}}{dt_3} = A a_3^{(0)} \tag{3.56c}$$

$$\frac{ds_{1,3}^{(0)}}{dt_3} = 0 \tag{3.56d}$$

$$\frac{ds_{1,3}^{(1)}}{dt_3} = -\frac{A a_3^{(0)}}{6} \tag{3.56e}$$

$$\frac{da_3^{(0)}}{dt_3} = \frac{s_{2,3}^{(0)}}{6} - s_{0,3}^{(0)} a_3^{(0)} \tag{3.56f}$$

Solving (3.56a) for $a_3^{(0)}$ and substituting into (3.56f) gives

$$\frac{d^2 s_{0,3}^{(0)}}{dt_3^2} + s_{0,3}^{(0)} \frac{d s_{0,3}^{(0)}}{dt_3} + \frac{5 A s_{2,3}^{(0)}}{36} = 0. \tag{3.57}$$

Integrating (3.57) once with respect to t_3 gives a Riccati equation

$$\frac{d s_{0,2}^{(0)}}{dt_3} + \frac{1}{2}\left(s_{0,3}^{(0)}\right)^2 + \frac{5 A s_{2,3}^{(0)} t_3}{36} + C = 0, \tag{3.58}$$

where C is constant. The substitution $s_{0,3}^{(0)} = 2v'/v$ transforms the Riccati equation into an Airy equation

$$\frac{d^2 v}{dt_3^2} - \left(-\frac{5 A s_{2,3}^{(0)} t_3}{72} + \frac{C}{2}\right) v = 0. \tag{3.59}$$

Equation (3.59) has solutions in terms of the Airy functions Ai and Bi. Equations of this form also appear in the WKB approximation: (3.59) can be interpreted as a Schrödinger equation where v is the wavefunction and the time dependent coefficient to v is a potential linearized around the boundary between classically allowed and forbidden regions. Solving (3.59) and inverting the variable substitution gives

$$s_{0,3}^{(0)} = -\frac{\left(5 A k_{20} s_{2,3}^{(0)}\right)^{1/3} \left(\mathrm{Bi}'(z) + C_3 \,\mathrm{Ai}'(z)\right)}{3^{2/3} \left(\mathrm{Bi}(z) + C_3 \,\mathrm{Ai}(z)\right)} \tag{3.60a}$$

$$s_{2,3}^{(1)} = C_4 - \frac{6}{5} s_{0,3}^{(0)} \tag{3.60b}$$

$$s_{1,3}^{(1)} = C_5 + \frac{1}{5} s_{0,3}^{(0)} \tag{3.60c}$$

$$a_3^{(0)} = C_2 + \frac{k_{20} s_{2,3}^{(0)} t_3}{6} + \frac{\left(A k_{20} s_{2,3}^{(0)}\right)^{2/3} \left(\mathrm{Bi}'(z) + C_3 \,\mathrm{Ai}'(z)\right)^2}{15^{1/3} \left(\mathrm{Bi}(z) + C_3 \,\mathrm{Ai}(z)\right)^2} \tag{3.60d}$$

$$z \equiv -\frac{5^{1/3} A \left(6 C_2 + k_{20} s_{2,3}^{(0)} t_3\right)}{2 \left(3 A k_{20} s_{2,3}^{(0)}\right)^{2/3}}, \tag{3.60e}$$

where Ai' and Bi' are the derivatives of the Airy functions Ai and Bi. As $t_3 \to -\infty$ back into the terminal dephosphorylation region, $z \to \infty$. For large z, Bi grows exponentially whereas Ai decays exponentially to 0. For (3.60) to match with the terminal dephosphorylation region, we must have

$C_3 \to \infty$ so that

$$\begin{aligned} s_{0,3}^{(0)} &= -\frac{\left(5Ak_{20}s_{2,3}^{(0)}\right)^{1/3}\mathrm{Ai}'(z)}{3^{2/3}\,\mathrm{Ai}(z)} \\ a_3^{(0)} &= C_2 + \frac{k_{20}s_{2,3}^{(0)}t_3}{6} + \frac{\left(Ak_{20}s_{2,3}^{(0)}\right)^{2/3}\mathrm{Ai}'(z)^2}{15^{1/3}\,\mathrm{Ai}(z)^2} \end{aligned} \tag{3.61}$$

The constants C_4 and C_5 can be determined by expanding the phase plane parameterization of (3.49) as $s_{0,2}^{(0)} \to 5A/6 + \xi^{1/3}s_{0,3}^{(0)}$:

$$\begin{aligned} \frac{ds_{2,2}^{(0)}}{ds_{0,2}^{(0)}} &= \frac{1}{\xi^{1/3}}\frac{ds_{2,2}^{(0)}}{ds_{0,3}^{(0)}} \\ \frac{ds_{2,2}^{(0)}}{ds_{0,3}^{(0)}} &= -\frac{6}{5}\xi^{1/3} + \mathcal{O}\left(\xi^{2/3}\right) \\ \frac{ds_{1,2}^{(0)}}{ds_{0,2}^{(0)}} &= \frac{1}{\xi^{1/3}}\frac{ds_{1,2}^{(0)}}{ds_{0,3}^{(0)}} \\ \frac{ds_{1,2}^{(0)}}{ds_{0,3}^{(0)}} &= \frac{1}{5}\xi^{1/3} + \mathcal{O}\left(\xi^{2/3}\right). \end{aligned} \tag{3.62}$$

Integrating (3.62) gives

$$\begin{aligned} s_{2,2}^{(0)} &= s_{2,3}^{(0)} - \frac{6}{5}\xi^{1/3}s_{0,3}^{(0)} + \mathcal{O}\left(\xi^{2/3}\right) \\ s_{2,2}^{(0)} &= s_{1,3}^{(0)} + \frac{1}{5}\xi^{1/3}s_{0,3}^{(0)} + \mathcal{O}\left(\xi^{2/3}\right). \end{aligned} \tag{3.63}$$

Since no constant term appears at the $\mathcal{O}(\xi^{1/3})$ tier in (3.63), we must have

$$\begin{aligned} C_4 &= 0 \\ C_5 &= 0. \end{aligned} \tag{3.64}$$

The constants $s_{2,3}^{(0)}$ and $s_{1,3}^{(0)}$ must be determined by numerically integrating (3.49) starting from a point on the main dephosphorylation segment solution where $s_0^{(0)} \sim \delta$, $s_2^{(0)} \sim \delta$, $s_1^{(0)} \sim \delta$ until $s_{0,2}^{(0)} = 5A/6$.

s_0 depletion

At $t \approx 12$ in Figure 3.10, s_0 falls below $5\delta A/6$ and the quasi steady state (3.44) loses validity and a grows from $\mathcal{O}(\epsilon)$ to $\mathcal{O}(\delta)$. Define the transformed variable

$$s_A \equiv s_2 + s_1. \tag{3.65}$$

The updated variable scalings are

$$\begin{aligned} s_0 &= \delta s_{0,4}^{(0)} + \mathcal{O}(\epsilon) \\ s_A &= \delta s_{A,4}^{(0)} + \epsilon s_{A,4}^{(1)} + \mathcal{O}(\epsilon\xi) \\ s_1 &= A + \delta s_{1,4}^{(0)} + \mathcal{O}(\epsilon), \end{aligned} \tag{3.66}$$

With a at $\mathcal{O}(\delta)$, the rate of $k_{12}s_1a$ increases, starting a positive feedback loop that releases more a from s_1. At $t \approx 12$ in Figure 3.10, the remaining s_0 rapidly falls from $\mathcal{O}(\delta)$ to a floor of $\mathcal{O}(\epsilon)$ as it is consumed by the increasing pool of a. We parameterize the orbit spatially using s_0 as the independent variable

$$\begin{aligned} \frac{ds_1}{ds_0} &= \frac{k_{01}s_0a - k_{12}s_1a}{-k_{su}s_0 + k_{20}s_2 - k_{01}s_2a} \\ \frac{ds_A}{ds_0} &= \frac{-k_{20}s_2 + k_{01}s_0a + k_{us}ua}{-k_{su}s_0 + k_{20}s_2 - k_{01}s_2a}. \end{aligned} \tag{3.67}$$

If this segment of the orbit is parameterized in time, t should be rescaled as $t = \delta t_4$, but the time variable drops out of the spatial parameterization (3.67). Substituting (3.66) into (3.68) gives

$$\frac{ds_{1,4}^{(0)}}{ds_{0,4}^{(0)}} = -1 + \frac{A}{s_{0,4}^{(0)}} \tag{3.68a}$$

$$s_{A,4}^{(0)} = C_6 - s_{0,4}^{(0)} \tag{3.68b}$$

$$\frac{ds_{A,4}^{(1)}}{ds_{0,4}^{(0)}} = -\frac{k_{us}(1-A)}{s_{0,4}^{(0)}} - \frac{6k_{su}}{s_{A,4}^{(0)} + 5s_{1,4}^{(0)}}. \tag{3.68c}$$

Equation (3.68) has solutions

$$s_{1,4}^{(0)} = C_7 - s_{0,4}^{(0)} + A\log\left(s_{0,4}^{(0)}\right) \tag{3.69a}$$

$$s_{A,4}^{(1)} = C_8 - k_{us}(1-A)\log\left(s_{0,4}^{(0)}\right) - \int_0^{s_{0,4}^{(0)}} \frac{6k_{su}}{C_6 + 5C_7 - 6y + 5A\log(y)}. \tag{3.69b}$$

To determine C_7, we expand

$$s_{0,4}^{(0)} = \frac{5A}{6} + \xi^{1/3}s_{0,3}^{(0)} \tag{3.70a}$$

$$\begin{aligned}\frac{s_1 - A}{\delta} &= C_7 - \frac{5A}{6} + \xi^{1/3}s_{0,3}^{(0)} + A\log\left(\frac{5A}{6} + \xi^{1/3}s_{0,3}^{(0)}\right) \\ &\approx C_7 - \frac{5A}{6} - \xi^{1/3}s_{0,3}^{(0)} + A\left(\log\left[\frac{5A}{6}\right] + \frac{6\xi^{1/3}s_{0,3}^{(0)}}{5A}\right) \\ &= C_7 - \frac{5A}{6} + \frac{\xi^{1/3}s_{0,3}^{(0)}}{5} + A\log\left(\frac{5A}{6}\right).\end{aligned} \tag{3.70b}$$

The matching target for the spatial parameterization of the dephosphorylation QSS departure segment can be obtained directly from (3.60):

$$\frac{s_1 - A}{\delta} = s_{1,3}^{(0)} + \frac{\xi^{1/3}s_{0,3}^{(0)}}{5}. \tag{3.71}$$

To eliminate the constant terms in (3.70b), we choose

$$C_7 = s_{1,3}^{(0)} + A\left(\frac{5}{6} - \log\left[\frac{5A}{6}\right]\right). \tag{3.72}$$

The dephosphorylation QSS departure segment matching target for s_A is

$$\frac{s_A - A}{\delta} = s_{2,3}^{(0)} + s_{1,3}^{(0)} - \frac{5A}{6} - \xi^{1/3}s_{0,3}^{(0)}. \tag{3.73}$$

Substituting (3.70a) into (3.68b) gives

$$s_{A,4}^{(0)} = C_6 - \frac{5A}{6} - \xi^{1/3}s_{0,3}^{(0)}. \tag{3.74}$$

To balance the constant term in (3.73), we choose

$$C_6 = s_{2,3}^{(0)} + s_{1,3}^{(0)} \equiv s_{A,3}^{(0)} \tag{3.75}$$

To determine the remaining constant C_8, the expansion in the dephosphorylation QSS departure segment would need to be extended to $\mathcal{O}(\delta\xi^{3/3})$. For simplicity, we leave it unbalanced and set $C_8 = 0$.

s_0 floor

Towards the end of the s_0 depletion segment, s_0 approaches $\mathcal{O}(\epsilon)$, but s_2 and s_1 remain at $\mathcal{O}(\delta)$ since the conversion of s_0 to s_2 only changes the total concentration of s_2 by $\mathcal{O}(\delta)$. For small s_0, (3.69a) and (3.68b) show that s_0 decreases exponentially with increasing s_2. Past the s_0 depletion region, s_0 is at the quasi steady state (3.84) and depends linearly on s_2. The differing slopes of the s_0 depletion segment and the subsequent s_1 depletion segment must be bridged by an ultra fast region where s_0 adjusts down to its $\mathcal{O}(\epsilon)$ minimum. In Figure 3.10, this appears as a corner in solution for s_0 at $t \approx 12.02$. This segment has the variable scalings

$$\begin{aligned} s_0 &= \epsilon s_{0,5}^{(0)} \\ s_1 &= A + \delta \log(\xi) A + \delta s_{1,5}^{(0)} \\ s_A &\equiv s_2 + s_1 \\ &= A + \epsilon \log(\xi) s_{A,5}^{(0)} + \epsilon s_{A,5}^{(1)} \\ t &= \frac{\xi t_5}{\log(\xi)}. \end{aligned} \tag{3.76}$$

Note that since $\xi < 1$, $\log(\xi) < 0$ and the sign of t_5 is reversed relative to t. The variable scalings (3.76) result in the equations

$$\begin{aligned}
\frac{ds_{0,5}^{(0)}}{dt_5} &= A\left(\frac{5s_{0,5}^{(0)}}{6} - k_{20}\right) \\
\frac{ds_{1,5}^{(0)}}{dt_5} &= \frac{5A^2}{6} \\
\frac{ds_{A,5}^{(0)}}{dt_5} &= 0 \\
\frac{ds_{A,5}^{(1)}}{dt_5} &= k_{20}A - \frac{5k_{us}A(1-A)}{6} - \frac{5As_{0,5}^{(0)}}{6}.
\end{aligned} \tag{3.77}$$

Equation (3.77) has the solutions

$$s_{0,5}^{(0)} = \frac{6k_{20}}{5} + C_9 \exp\left(\frac{5At_5}{6}\right) \tag{3.78a}$$

$$s_{1,5}^{(0)} = C_{10} + \frac{5A^2t_5}{6} \tag{3.78b}$$

$$s_{A,5}^{(1)} = C_{11} - \frac{6k_{20}}{5} - \frac{5k_{us}A(1-A)t_5}{6} + C_9\left(1 - \exp\left[\frac{5At_5}{6}\right]\right). \tag{3.78c}$$

To match with the s_0 depletion region, we invert (3.78a) to obtain

$$\begin{aligned}
t_5 &= \frac{6}{5A}\log\left(\frac{s_{0,5}^{(0)}}{C_9} - \frac{6k_{20}}{5C_9}\right) \\
s_1 &= A + \delta\left(A\log(\xi) + C_{10} + A\log\left[\frac{s_{0,5}^{(0)}}{C_9} - \frac{6k_{20}}{5C_9}\right]\right) \\
&\approx A + \delta\left(A\log(\xi) + C_{10} + A\log\left[\frac{s_{0,5}^{(0)}}{C_9}\right]\right) \\
s_A &= A + \epsilon\left(\log(\xi)s_{A,5}^{(0)} + C_9 + C_{11} - s_{0,5}^{(0)} - k_{us}(1-A)\log\left[\frac{s_{0,5}^{(0)}}{C_9} - \frac{6k_{20}}{5C_9}\right]\right) \\
&\approx A + \epsilon\left(\log(\xi)s_{A,5}^{(0)} + C_9 + C_{11} - s_{0,5}^{(0)} - k_{us}(1-A)\log\left[\frac{s_{0,5}^{(0)}}{C_9}\right]\right).
\end{aligned} \tag{3.79}$$

To obtain the s_0 depletion segment matching target, we rewrite (3.69) in terms of $s_{0,5}^{(0)}$ as $s_{0,5}^{(0)} \to 0$

$$\begin{aligned} s_1 &= A + \delta s_{1,3}^{(0)} + \delta A \left(\frac{5}{6} - \log\left[\frac{5A}{6}\right] + \log\left[\xi s_{0,5}^{(0)}\right] \right) \\ s_A &= A + \delta s_{A,3}^{(0)} - \epsilon \left(s_{0,5}^{(0)} - k_{us}(1-A) \log\left[\xi s_{0,5}^{(0)}\right] \right). \end{aligned} \tag{3.80}$$

To complete the match, we choose

$$\begin{aligned} s_{A,3}^{(0)} &= 0 \\ C_9 &= 1 \\ C_{10} &= s_{1,3}^{(0)} + A \left(\frac{5}{6} - \log\left[\frac{5A}{6}\right] \right) \\ s_{A,5}^{(0)} &= -k_{us}(1-A) \\ C_{11} &= -1. \end{aligned} \tag{3.81}$$

s_1 depletion

From time $t \approx 12.1$ to $t \approx 20$ in Figure 3.10, $s_0 = \mathcal{O}(\epsilon)$ and $a = \mathcal{O}(1)$. This suggests the variable scalings

$$\begin{aligned} s_0 &= \epsilon s_0^{(0)} + \mathcal{O}(\epsilon\xi) \\ s_2 &= s_2^{(0)} + \mathcal{O}(\epsilon) \\ s_1 &= s_1^{(0)} + \mathcal{O}(\epsilon) \\ a &= a^{(0)} + \mathcal{O}(\epsilon) \\ u^{(0)} &= 1 - s_0^{(0)} - s_2^{(0)} + \mathcal{O}(\epsilon). \end{aligned} \tag{3.82}$$

Substituting (3.82) into (3.1) gives the leading order equations

$$0 = k_{20} s_2 - s_0^{(0)} a^{(0)} \tag{3.83a}$$

$$\frac{ds_2^{(0)}}{dt} = -k_{20} s_2^{(0)} + k_{us} u^{(0)} a^{(0)} + k_{12} s_1^{(0)} a^{(0)} \tag{3.83b}$$

$$\frac{ds_1^{(0)}}{dt} = s_0^{(0)} a^{(0)} - k_{12} s_1^{(0)} a^{(0)}. \tag{3.83c}$$

Equation (3.83a) imposes the quasi steady state

$$s_0^{(0)} = \frac{k_{20} s_2^{(0)}}{a^{(0)}}. \tag{3.84}$$

Substituting (3.84) into (3.83) gives a planar system:

$$\begin{aligned} \frac{ds_2^{(0)}}{dt} &= k_{us} u^{(0)} a^{(0)} + k_{12} s_1^{(0)} a^{(0)} - k_{20} s_2^{(0)} \\ \frac{ds_1^{(0)}}{dt} &= k_{20} s_2^{(0)} - k_{12} s_1^{(0)} a^{(0)}. \end{aligned} \tag{3.85}$$

The solution to equation (3.85) is uniformly valid for the next four segments but is not analytically tractable. We further expand (3.85) and the rescaled variables (3.82) in terms of ξ for analytically tractable solutions. Sequentially expanding (3.83) in terms of ξ after expanding (3.1) in terms of ϵ simplifies the algebra compared to the alternate approach of constructing a bivariate expansion of (3.1) in terms of ξ and ϵ, but we note that the sequential approach is only valid when the highest power of ξ that appears in the expansion is smaller than ϵ.

When the system enters the s_1 depletion segment,

$$\begin{aligned} s_1 &= \mathcal{O}(1) \\ s_2 &= \mathcal{O}(\xi). \end{aligned} \tag{3.86}$$

From these initial conditions, s_1 rapidly falls to $\mathcal{O}(\xi)$ and is converted to s_2, which rises to $\mathcal{O}(1)$. This transition occur over a $\mathcal{O}(\xi)$ timescale, during which $s_1 + s_2$ is approximately conserved. This leads to the variable scalings

$$\begin{aligned} s_1^{(0)} &= s_{1,6}^{(0)} + \mathcal{O}(\xi) \\ s_A^{(0)} &= A + \xi s_{A,6}^{(0)} + \mathcal{O}(\xi^2). \end{aligned} \tag{3.87}$$

In terms of $s_A^{(0)}, s_1^{(0)}$, (3.85) is

$$\begin{aligned} \frac{ds_A^{(0)}}{dt} &= k_{us} u^{(0)} a^{(0)} \\ \frac{ds_1^{(0)}}{dt} &= k_{20} s_2^{(0)} - k_{12} s_1^{(0)} a^{(0)}. \end{aligned} \tag{3.88}$$

The spatial parameterization of (3.88) is

$$\frac{ds_A^{(0)}}{ds_1^{(0)}} = \frac{k_{us}u^{(0)}a^{(0)}}{k_{20}s_2^{(0)} - k_{12}s_1^{(0)}a^{(0)}}. \tag{3.89}$$

Substituting (3.87) into (3.89) and expanding up to $\mathcal{O}(\xi)$ gives

$$\frac{ds_{A,6}^{(0)}}{ds_{1,6}^{(0)}} = -\frac{k_{us}(1-A)}{s_1^{(0)}}. \tag{3.90}$$

Equation (3.90) has solutions

$$s_{A,6}^{(0)} = C_{12} - k_{us}\,(1-A)\log\left(s_{1,6}^{(0)}\right). \tag{3.91}$$

To match with the s_0 floor, we consider the limit

$$\begin{aligned} s_1^{(0)} &= A + \delta s_{1,5} \\ s_{A,6}^{(0)} &= C_{12} - k_{us}(1-A)\log\left(A + \delta s_{1,5}\right) \\ &\approx C_{12} - k_{us}(1-A)\left(\log(A) + \frac{\delta s_{1,5}}{A}\right). \end{aligned} \tag{3.92}$$

To obtain the s_0 floor matching target for s_A, we solve

$$\begin{aligned} s_{1,5} &= \log(\xi)A + C_{10} + \frac{5A^2 t_5}{6} \\ t_5 &= \frac{6}{5}\left(\frac{s_{1,5} - C_{10}}{A^2} + \frac{\log(\xi)}{A}\right) \\ \frac{s_A - A}{\xi} &= \delta\left(-\frac{6k_{20}}{5} + 1 - \frac{k_{us}(1-A)\left(s_{1,5} - C_{10}\right)}{A}\right) - \epsilon\exp\left(\frac{s_{1,5} - C_{10}}{A} + \log(\xi)\right). \end{aligned} \tag{3.93}$$

As $t_5 \to -\infty$, $s_{1,5} \to -\infty$ such that

$$\frac{s_A - A}{\xi} = \delta\left(-\frac{6k_{20}}{5} + 1 - \frac{k_{us}(1-A)\left(s_{1,5} - C_{10}\right)}{A}\right). \tag{3.94}$$

The linear terms in (3.92) and (3.94) balance, the constant term in (3.92) can be eliminated by

choosing

$$C_{12} = k_{us}(1 - A)\log(A). \tag{3.95}$$

In the limit (3.92), the quasi steady state (3.84) becomes

$$\begin{aligned} \frac{k_{20}s_2^{(0)}}{a^{(0)}} &= \frac{k_{20}\left(s_A^{(0)} - s_1^{(0)}\right)}{A - s_A^{(0)}/6 - 5s_1^{(0)}/6} \\ &= \frac{-\delta k_{20}s_{1,5} + \mathcal{O}(\epsilon)}{-5\delta s_{1,5}/6 + \mathcal{O}(\epsilon)} \\ &= \frac{6k_{20}}{5}. \end{aligned} \tag{3.96}$$

Taking the limit as $t_5 \to -\infty$ in (3.78a) results in the same $6k_{20}/5$ minimum for $s_{0,5}^{(0)}$. To balance the remaining constant terms in (3.94), the expansion of s_A must be extended to $\mathcal{O}(\epsilon)$. For simplicity, we limit the expansion of (3.88) to $\mathcal{O}(\xi)$ and leave the $O(\delta)$ residual in (3.94) unbalanced.

s_1 floor

Although the s_1 depletion segment successfully resolves the approach of the orbit to the $s_1^{(0)}$ nullcline, (3.91) has a logarithmic singularity at $s_{1,6}^{(0)} = 0$ and cannot describe the slow accumulation of $s_1^{(0)}$ after reaching its minimum. Since $s_1^{(0)}$ decays to a constant at $\mathcal{O}(\xi)$, its slow accumulation should be a $\mathcal{O}(\xi^2)$ effect. We introduce a region parameterized by time, with variable scalings

$$\begin{aligned} t &= \xi t_7 \\ s_A^{(0)} &= A + \xi\log(\xi)s_{A,7}^{(0)} + \xi s_{A,7}^{(1)} \\ s_1^{(0)} &= \xi s_{1,7}^{(0)} + \xi^2\log(\xi)s_{1,7}^{(1)} + \xi^2 s_{1,7}^{(2)}. \end{aligned} \tag{3.97}$$

The variable scalings (3.97) result in the equations

$$\begin{aligned}
\frac{ds_{A,7}^{(0)}}{dt_7} &= 0 \\
\frac{ds_{A,7}^{(1)}}{dt_7} &= \frac{5k_{us}A\left(1-A\right)}{6} \\
\frac{ds_{1,7}^{(0)}}{dt_7} &= A\left(k_{20} - \frac{5}{6}s_{1,7}^{(0)}\right) \\
\frac{ds_{1,7}^{(1)}}{dt_7} &= k_{20}s_{A,7}^{(0)} + \frac{1}{6}\left(s_{A,7}^{(0)}s_{1,7}^{(0)} - 5As_{1,7}^{(1)}\right) \\
\frac{ds_{1,7}^{(2)}}{dt_7} &= k_{20}\left(s_{A,7}^{(1)} - s_{1,7}^{(0)}\right) + \frac{s_{1,7}^{(0)}}{6}\left(s_{A,7}^{(1)} + 5s_{1,7}^{(0)}\right) - \frac{5As_{1,7}^{(2)}}{6}.
\end{aligned} \tag{3.98}$$

Equation (3.98) has solutions

$$s_{A,7}^{(1)} = C_{13} + \frac{5k_{us}A\left(1-A\right)t_7}{6} \tag{3.99a}$$

$$s_{1,7}^{(0)} = \frac{6k_{20}}{5} + C_{14}\exp\left(-\frac{5At_7}{6}\right) \tag{3.99b}$$

$$s_{1,7}^{(1)} = \frac{36k_{20}s_{A,7}^{(0)}}{25A} + \exp\left(-\frac{5At_7}{6}\right)\left(\frac{s_{A,7}^{(0)}C_{14}t_7}{6} - C_{15}\right) \tag{3.99c}$$

$$\begin{aligned}
s_{1,7}^{(2)} = &-\frac{36k_{20}k_{us}\left(1-A\right)}{25A} + \frac{6k_{20}k_{us}(1-A)t_7}{5} + \frac{36k_{20}C_{13}}{25A} \\
&+ \frac{5k_{us}C_{14}A(1-A)t_7^2}{72}\exp\left(-\frac{5At_7}{6}\right) + C_{14}\left(k_{20} + \frac{C_{13}}{6}\right)t\exp\left(-\frac{5At_7}{6}\right) \\
&- \frac{C_{14}^2}{A}\exp\left(-\frac{5At}{3}\right) + C_{16}\exp\left(-\frac{5At}{6}\right).
\end{aligned} \tag{3.99d}$$

To determine $s_{A,7}^{(0)}$, C_{13}, and C_{14}, we express t_7 in terms of $s_{1,7}^{(0)}$:

$$\begin{aligned}
s_{1,7}^{(0)} &= \frac{6k_{20}}{5} + C_{14}\exp\left(-\frac{5At_7}{6}\right) \\
t_7 &= \frac{6}{5A}\log\left(\frac{5C_{14}}{5s_{1,7}^{(0)} - 6k_{20}}\right).
\end{aligned} \tag{3.100}$$

substituting (3.100) into (3.99a) gives

$$\begin{aligned}\frac{s_A^{(0)}-A}{\xi} &= \log(\xi)s_{A,7}^{(0)} + s_{A,7}^{(1)} \\ &= \log(\xi)s_{A,7}^{(0)} + C_{13} - k_{us}\,(1-A)\log\left(\frac{s_{1,7}^{(0)}}{C_{14}} - \frac{6k_{20}}{5C_{14}}\right) \\ &\approx \log(\xi)s_{A,7}^{(0)} + C_{13} - k_{us}\,(1-A)\log\left(\frac{s_{1,7}^{(0)}}{C_{14}}\right).\end{aligned} \tag{3.101}$$

From the s_1 depletion segment,

$$\begin{aligned}s_{A,6}^{(1)} &= -k_{us}\,(1-A)\log\left(\frac{s_1^{(0)}}{A}\right) \\ &= -k_{us}(1-A)\left(\log(\xi) + \log\left[\frac{s_{1,7}^{(0)}}{A}\right]\right).\end{aligned} \tag{3.102}$$

To complete the match, we choose

$$\begin{aligned}s_{A,7}^{(0)} &= -k_{us}(1-A) \\ C_{13} &= 0 \\ C_{14} &= A.\end{aligned} \tag{3.103}$$

The solutions to $s_{1,7}^{(1)}$ and $s_{1,7}^{(2)}$ contain terms multiplied by C_{15} and C_{16} that are not balanced by anything in (3.102), which is only resolved up to $\mathcal{O}(\xi)$, so we set

$$\begin{aligned}C_{15} &= 0 \\ C_{16} &= 0.\end{aligned} \tag{3.104}$$

Main phosphorylation segment

At the end of the s_1 depletion segment, s_1 reaches a minimum of $6\xi k_{20}/5$ and the system enters the main phosphorylation segment, where u is slowly converted to s_2. We use the spatial parameteri-

zation of the orbit, with

$$\begin{aligned} s_1^{(0)} &= \xi s_{1,8}^{(0)} + \xi^2 s_{1,8}^{(1)} + \mathcal{O}(\epsilon) \\ s_A^{(0)} &= s_{A,8}^{(0)} + \mathcal{O}(\xi). \end{aligned} \tag{3.105}$$

From (3.88)

$$\begin{aligned} \xi^{-1} s_1^{(0)} &= \frac{k_{20} s_2^{(0)}}{a^{(0)}} - \frac{1}{a^{(0)}} \frac{ds_1^{(0)}}{dt} \\ s_{1,8}^{(0)} + \xi s_{1,8}^{(1)} &= \frac{k_{20} s_2^{(0)}}{a^{(0)}} - \frac{\xi}{a^{(0)}} \frac{ds_{1,8}^{(0)}}{dt} \\ &= \frac{k_{20} s_2^{(0)}}{a^{(0)}} - \frac{\xi}{a^{(0)}} \frac{ds_{1,8}^{(0)}}{ds_{A,8}^{(0)}} \frac{ds_{A,8}^{(0)}}{dt} \\ &= \frac{k_{20} s_2^{(0)}}{a^{(0)}} - \xi u^{(0)} \frac{ds_{1,8}^{(0)}}{ds_{A,8}^{(0)}}. \end{aligned} \tag{3.106}$$

Equating terms of size $\mathcal{O}(1)$ in (3.106) gives

$$\begin{aligned} s_{1,8}^{(0)} &= \frac{k_{20} s_2^{(0)}}{a^{(0)}} \\ &= \frac{k_{20} \left(s_{A,8}^{(0)} - \xi s_{1,8}^{(0)} \right)}{A - s_{A,8}^{(0)}/6 - 5\xi s_{1,8}^{(0)}/6}. \end{aligned} \tag{3.107}$$

This is the nullcline for $s_1^{(0)}$. Equating terms of size $\mathcal{O}(\xi)$ in (3.106) gives

$$\begin{aligned} s_{1,8}^{(1)} &= -\frac{1}{a^{(0)}} \frac{ds_{1,8}^{(0)}}{ds_{A,8}^{(0)}} \frac{ds_{A,8}^{(0)}}{dt} \\ s_{1,8}^{(1)} &= -k_{us} u^{(0)} \frac{ds_{1,8}^{(0)}}{ds_{A,8}^{(0)}}. \end{aligned} \tag{3.108}$$

To determine $ds_{1,8}^{(0)}/ds_{A,8}^{(0)}$, we differentiate (3.107) with respect to $s_{A,8}^{(0)}$

$$\frac{ds_{1,8}^{(0)}}{ds_{A,8}^{(0)}} = \frac{k_{20} \left(A - \xi s_{1,8}^{(0)} - \xi (1 - A) \frac{ds_{1,8}^{(0)}}{ds_{A,8}^{(0)}} \right)}{\left(A - s_{A,8}^{(0)}/6 - 5\xi s_{1,8} \right)^2}. \tag{3.109}$$

Solving (3.107) for $s_{A,8}^{(0)}$ gives

$$s_{A,8}^{(0)} = \frac{s_{1,8}^{(0)}\left(A + \xi k_{20} - 5\xi s_{1,8}^{(0)}/6\right)}{k_{20} + s_{1,8}^{(0)}/6}. \tag{3.110}$$

Substituting (3.110) into (3.109) gives

$$\frac{ds_{1,8}^{(1)}}{ds_{A,8}^{(0)}} = \frac{\left(k_{20} + s_{1,8}^{(0)}/6\right)^2}{k_{20}A + \xi\left(k_{20}^2 - 5k_{20}s_{1,8}^{(0)}/3 - 5\left(s_{1,8}^{(0)}\right)^2\right)} \tag{3.111a}$$

$$s_{1,8}^{(1)} = \frac{\left(k_{20} + s_{1,8}^{(0)} + s_{1,8}^{(0)}\left(A + \xi k_{20} - 5\xi s_{1,8}^{(0)}/6\right)\right)\left(k_{20} + s_{1,8}^{(0)}\right)}{k_{20}A + \xi\left(k_{20}^2 - 5k_{20}s_{1,8}^{(0)}/3 - 5\left(s_{1,8}^{(0)}\right)^2\right)}. \tag{3.111b}$$

Equation (3.107) and (3.111b) are resummed solutions: the $\xi s_{1,8}^{(0)}$ term has been promoted to $\mathcal{O}(1)$. The relatively large value of ξ results in poor agreement with the numerical solution if the $s_{1,8}^{(0)}$ is properly treated as a $\mathcal{O}(\xi)$ term. To match the main dephosphorylation segment with the end of the s_1 depletion segment, we consider the large t_7 behavior of (3.99). From the s_1 floor,

$$s_{A,8}^{(0)} = A + \xi k_{us}(1 - A)\left(-\log(\xi) + \frac{5At_7}{6}\right). \tag{3.112}$$

Equation (3.112) can be inverted to express t_7 in terms of $s_{A,8}^{(0)}$

$$t_7 = \frac{6\left(s_{A,8}^{(0)} - A\right)}{5\xi k_{us}A(1 - A)} + \frac{6\log(\xi)}{5A}. \tag{3.113}$$

For large t_7,

$$\begin{aligned}\xi s_{1,7}^{(0)} + \xi^2\log(\xi)s_{1,7}^{(1)} + \xi^2 s_{1,7}^{(2)} &\approx \frac{6\xi k_{20}}{5} + \xi^2\left(-\frac{36(1 + \log(\xi))k_{20}k_{us}(1 - A)}{25A} + \frac{6k_{20}k_{us}(1 - A)t}{5}\right) \\ &= \frac{6\xi k_{20}}{5} + \frac{36\xi k_{20}\left(s_{A,8}^{(0)} - A\right)}{25A} - \frac{36\xi^2 k_{20}\,(1 - A)}{25A}.\end{aligned} \tag{3.114}$$

At the start of the main phosphorylation segment,

$$\begin{aligned}
s_{A,8}^{(0)} &\approx A + \xi s_{A,7}^{(0)} \\
\xi s_{1,8}^{(0)} &= \frac{\xi k_{20}\left(A + \xi s_{A,7}^{(0)} - \xi s_{1,8}^{(0)}\right)}{A - \left(A + \xi s_{A,7}^{(0)}\right)/6 - 5\xi s_{1,8}^{(0)}/6} \\
&\approx \left(\frac{6\xi k_{20}}{5} + \frac{6\xi^2 k_{20} s_{A,7}^{(0)}}{5A} - \frac{6\xi^2 k_{20} s_{1,8}^{(0)}}{5A}\right)\left(1 + \frac{\xi s_{A,7}^{(0)}}{5A} + \frac{\xi s_{1,8}^{(0)}}{A}\right) \\
&\approx \left(\frac{6\xi k_{20}}{5} + \frac{6\xi k_{20}\left(s_{A,8}^{(0)} - A\right)}{5A} - \frac{6\xi^2 k_{20} s_{1,8}^{(0)}}{5A}\right)\left(1 + \frac{s_{A,8}^{(0)} - A}{5A} + \frac{\xi s_{1,8}^{(0)}}{A}\right) \\
&\approx \frac{6\xi k_{20}}{5} + \frac{36\xi k_{20}\left(s_{A,8}^{(0)} - A\right)}{25A}.
\end{aligned} \tag{3.115}$$

The terms in (3.115) and (3.114) match. The remaining term is found by taking the limit of (3.111b) as $s_{1,7}^{(0)} \to 6k_{20}/5$

$$\begin{aligned}
\lim_{s_{1,8}^{(0)} \to 6k_{20}/5} \xi^2 s_{1,8}^{(1)} &= -\frac{36\xi^2 k_{20}(1-A)}{5(5A - 6\xi k_{20})} \\
&= -\frac{36\xi^2 k_{20}(1-A)}{25A} + \mathcal{O}\left(\xi^3\right).
\end{aligned} \tag{3.116}$$

s_1 fold

The $s_1^{(0)}$ nullcline reaches a fold at

$$\begin{aligned}
s_{A,\text{fold}} &= 6A + 66\xi k_{20} - 2\sqrt{5k_{20}(6\xi^2 k_{20} + \xi A)} \\
s_{1,\text{fold}} &= -6\xi k_{20} + \sqrt{\frac{k_{20}(6\xi^2 k_{20} + \xi A)}{5}} \\
a_{\text{fold}} &= A - \frac{1}{6} s_{A,\text{fold}} - \frac{5}{6} s_{1,\text{fold}}.
\end{aligned} \tag{3.117}$$

The form of (3.117) and the behavior of the numerical solution at $t \approx 20$ suggests a segment describing the depletion of a with the variable scalings

$$\begin{aligned} t &= \xi^{1/2} t_{10} \\ s_1^{(0)} &= s_{1,\text{fold}} + \xi^{1/2} s_{1,10}^{(0)} + \mathcal{O}(\xi) \\ s_A &= s_{A,\text{fold}} + \xi s_{A,10}^{(0)} + \mathcal{O}\left(\xi^{3/2}\right) \\ a^{(0)} &= a_{\text{fold}} + \xi^{1/2} s_{1,10}^{(0)} + \mathcal{O}(\xi). \end{aligned} \tag{3.118}$$

The approach to the fold and departure from the phosphorylation quasi steady state would be described by another segment with the turning point scalings

$$\begin{aligned} s_1^{(0)} &= s_{1,\text{fold}} + \xi^{2/3} s_{1,9}^{(0)} + \mathcal{O}\left(\xi^{1/2}\right) \\ s_A^{(0)} &= s_{A,\text{fold}} + \xi^{5/6} s_{A,9}^{(0)} + \mathcal{O}(\xi). \end{aligned} \tag{3.119}$$

The variable scalings (3.119) are analogous to (3.55), but the $\mathcal{O}(\xi^{1/2})$ location of the fold imposes a local expansion gauge of $\xi^{1/6}$. For the numerical values of the parameters that we consider, the $\xi^{1/6}$ expansion gauge and the size of the coefficients in (3.117) are problematic. If $\xi \sim 10^{-3}$ to 10^{-4}, $\xi^{1/6} \sim 0.316$ to 0.216. Formally $\xi^{1/6}$ is a small parameter, but numerically it is close to $\mathcal{O}(1)$. Likewise, the expansion of $s_{A,\text{fold}}$ to $\mathcal{O}(\xi^{1/2})$ is

$$6A + 12\sqrt{5\xi k_{20} A} + \mathcal{O}(\xi). \tag{3.120}$$

If $\xi \sim 10^{-3}$ to 10^{-4}, the $\mathcal{O}(\xi^{1/2})$ term is 0.351 to 0.111, which again, despite being formally treated as a small correction, is in reality $\mathcal{O}(1)$. Instead of performing an asymptotic expansion of these two segments, which would be inaccurate and converge slowly, we make the crude approximation that the main dephosphorylation segment tracks the s_1 nullcline exactly until the fold, at which the

system jumps to

$$\begin{aligned} s_{0,1}^{(0)}(0) &= 0 \\ s_{A,1}^{(0)}(0) &= s_{A,\text{fold}} \\ s_{1,1}^{(0)}(0) &= \frac{6A}{5} - \frac{1}{5} s_{A,\text{fold}} \\ &= -66\xi k_{20} + 2\sqrt{\frac{k_{20}(6\xi^2 k_{20} + \xi A)}{5}}. \end{aligned} \tag{3.121}$$

We take (3.121) as the entry conditions to the main dephosphorylation segment, which fixes

$$\begin{aligned} C_0 &= 6A + \frac{392}{5}\xi k_{20} - 12\sqrt{5k_{20}(6\xi^2 k_{20} + \xi A)} \equiv s_2(0) \\ C_1 &= 0. \end{aligned} \tag{3.122}$$

This approximation also skips an ultra fast segment analogous to the s_0 floor, visible as a corner at $t \approx 20.9$ in Figure 3.10, which would have the variable scalings

$$\begin{aligned} t &= \epsilon t_{11} \\ s_0 &= \epsilon s_{0,11}^{(0)} + \mathcal{O}(\epsilon^2) \\ s_2 &= s_{2,11}^{(0)} + \mathcal{O}(\epsilon) \\ s_1 &= s_{1,11}^{(0)} + \mathcal{O}(\epsilon) \\ a &= \epsilon a_{11}^{(0)} + \mathcal{O}(\epsilon^2). \end{aligned} \tag{3.123}$$

3.6.2 Asymptotic duration of the phosphorylation segment

When numerically integrating the single-site model (3.1) with $k_{01} = 10^6, k_{12} = 300, k_{su} = k_{us} = k_{20} = 1$, the duration of the phosphorylation segment is highly dependent on the total KaiA concentration; increasing A from 6/35 to 9/35 triples the duration of the phosphorylation segment. However, the duration of the dephosphorylation segment depends only slightly on the total KaiA concentration. An asymptotic expression for the duration of the phosphorylation segment can be

obtained by computing

$$\begin{aligned}\tau_{\text{phos}}(s_A) &= \int \left(\frac{ds_A}{dt}\right)^{-1} ds_A \\ &= \int \frac{1}{(1-s_A)(a - s_A/6 - 5s_1/6)} ds_A.\end{aligned} \tag{3.124}$$

From (3.107), the value of s_1 in the dephosphorylation segment is

$$\begin{aligned}s_1 &\approx \xi s_{1,8}^{(0)} \\ &= \frac{k_{20}(s_A - s_1)}{k_{12}\left(A - s_A/6 - 5s_1/6\right)} \\ &= \frac{3\left(k_{20} + k_{12}(A - s_A/6)\right) \pm \sqrt{-10k_{20}k_{12}s_A/3 + \left(k_{20} + k_{12}(A - s_A/6)\right)^2}}{5k_{12}}.\end{aligned} \tag{3.125}$$

Substituting (3.125) and evaluating (3.124) gives

$$\begin{aligned}t_{\text{phos}} = -\frac{1}{12k_{us}k_{20}(1-A)}\Bigg[&2G\arctan\left(\frac{k_{12}(1-s_A)+F}{G}\right) \\ &+ 2(6k_{20} - 5k_{12}A)\log\left(-6k_{20} + k_{12}(6A - s_A) + F\right) \\ &- \left(6k_{20} - k_{12}(1+4A)\right)\log\left(66k_{20} + k_{12}(6A - s_A) + F\right) \\ &- \left(6k_{20} + k_{12}(1-6A)\right)\log\left(k_{12}(1-s_A)\right)\Bigg] \\ F \equiv \sqrt{36(k_{20} + k_{12}A)^2 - 12k_{12}(11k_{20} + k_{12}A)s_A + \left(k_{12}s_A\right)^2} \\ G \equiv \sqrt{-36k_{20}^2 + 12k_{20}k_{12}(11 - 6A) - k_{12}^2(6A-1)^2}.\end{aligned} \tag{3.126}$$

The phosphorylation segment begins when $s_1 \sim 6k_{20}/(5k_{12})$, i.e. at

$$s_{A,\text{start}} = A + \frac{6k_{20}}{5k_{12}}. \tag{3.127}$$

and ends when reaching the fold in the $s_1^{(0)}$ nullcline, which occurs at

$$s_{A,\text{fold}} = 6A + \frac{6(11k_{20} - 2\sqrt{5k_{20}(6k_{20} + k_{12}A)})}{k_{12}}. \tag{3.128}$$

The duration of the phosphorylation segment is then

$$\tau_{\text{phos}} = t_{\text{phos}}(s_{A,\text{fold}}) - t_{\text{phos}}(s_{A,\text{start}}). \tag{3.129}$$

The asymptotic period diverges at a value of A that causes

$$1 - s_{A,\text{fold}} = 0. \tag{3.130}$$

This occurs at

$$A_{\pm} = \frac{1}{6} - \frac{k_{20}}{k_{12}} \pm \frac{1}{3}\sqrt{\frac{30k_{20}}{k_{12}}}. \tag{3.131}$$

Thus, A_+ gives an estimate for the value of A at which oscillations fail. In the asymptotic limit, oscillations fail at A_+ where the dephosphorylation segment intersects a saddle node. Solving for the roots of (3.85) gives a trivial root at $s_1 = A, s_2 = 0$, and two conjugate roots at

$$\begin{aligned} s_1 &= \frac{6k_{20} - k_{12}(1-6A) \pm \sqrt{-120k_{20}k_{12} + (6k_{20} - k_{12}(1-A))^2}}{10k_{12}} \\ s_2 &= -\frac{6k_{20} - k_{12}(11-6A) \pm \sqrt{-120k_{20}k_{12} + (6k_{20} - k_{12}(1-A))^2}}{10k_{12}}. \end{aligned} \tag{3.132}$$

which are complex for $A^- < A < A_+$. At $A = A_+$, a saddle node appears on the s_1 nullcline, and the asymptotic period diverges to ∞. For $0 < A < A^-$, the conjugate roots are also real, but the location of the root is at negative values of s_1, making this case physically irrelevant.

3.6.3 Asymptotic duration of the dephosphorylation segment

Using the approximation (3.121), the solution for s_0 in the main dephosphorylation segment is

$$\begin{aligned} s_0(t) &= K\exp\left(-k_{su}t\right) - K\exp\left(-k_{20}t\right) \\ K &\equiv \frac{5k_{20}s_2(0)}{6(k_{20} - k_{su})}. \end{aligned} \tag{3.133}$$

To find the duration of the dephosphorylation segment, we solve

$$s_0(t) = \delta \equiv \frac{k_{12}}{k_{01}}. \tag{3.134}$$

For general k_{20} and k_{su}, $s_0(t)$ does not have an inverse in terms of known special functions. However, since $\delta \ll 1$, s_0 should reach $\mathcal{O}(\delta)$ at large t, so that if $k_{us} < k_{20}$, then $K \exp(-k_{su})$ dominates, leading to the approximation

$$\begin{aligned} s_0(t) &\approx K \exp\left(-k_{us}\tau^{(0)}_{\text{dephos}}\right) \\ &\approx \delta \\ \tau^{(0)}_{\text{dephos}} &= -\frac{1}{k_{su}} \log\left(\frac{\delta}{K}\right). \end{aligned} \tag{3.135}$$

This approximation can be refined by computing

$$\begin{aligned} K \exp\left(-k_{su}\tau^{(1)}_{\text{dephos}}\right) &= \delta + K \exp\left(-k_{20}\tau^{(0)}_{\text{dephos}}\right) \\ \tau^{(1)}_{\text{dephos}} &= -\frac{1}{k_{su}} \log\left(\frac{\delta}{K} + \exp\left(-k_{20}\tau^{(0)}_{\text{dephos}}\right)\right). \end{aligned} \tag{3.136}$$

Substituting (3.135) gives the 1 step iterative correction

$$\tau^{(1)}_{\text{dephos}} = -\frac{1}{k_{su}} \left(\log\left[\frac{\delta}{K}\right] + \log\left[1 + \left(\frac{\delta}{K}\right)^{\frac{k_{20}}{k_{su}}-1}\right]\right). \tag{3.137}$$

Similarly, in the case of of $k_{su} > k_{20}$,

$$\tau^{(1)}_{\text{dephos}} = -\frac{1}{k_{20}} \left(\log\left[-\frac{\delta}{K}\right] - \log\left[1 - \left(-\frac{\delta}{K}\right)^{\frac{k_{su}}{k_{20}}-1}\right]\right). \tag{3.138}$$

In the case of $k_{su} = k_{20} = \kappa$, τ_{dephos} has an exact solution in terms of the secondary branch of the Lambert W function

$$\tau_{\text{dephos}} = -\frac{1}{\kappa} W_{-1}\left(-\frac{6\delta}{5s_2(0)}\right). \tag{3.139}$$

Since $\delta \ll 1$,

$$\frac{1}{\kappa}W_{-1}\left(-\frac{6\delta}{5s_2(0)}\right) \approx -\frac{1}{\kappa}\left[\log\left(\frac{6\delta}{5s_2(0)}\right) - \log\left(-\log\left(\frac{6\delta}{5s_2(0)}\right)\right)\right]. \tag{3.140}$$

In (3.137), $K > 0$, and in (3.138), $K < 0$ but appears with a multiplicative factor of -1, so the quantity that enters in both cases is $|K|$. As $k_{20} \to k_{su}$, (3.137) and (3.138) diverge since the denominator of $k_{20} - k_{su} \to 0$ inside the logarithm. We can construct a composite solution by replacing $|k_{20} - k_{su}|$ in the denominator of $|K|$ with the regularized expression

$$|k_{20} - k_{su}| \approx \sqrt{(k_{20} - k_{su})^2 + \eta^2}. \tag{3.141}$$

Substituting (3.141) gives

$$\tau^{(0)}_{\text{dephos}} = -\frac{1}{\min(k_{su}, k_{20})}\log\left(\frac{6\delta\sqrt{(k_{20} - k_{su})^2 + \eta^2}}{5k_{20}s_2(0)}\right). \tag{3.142}$$

We then require that at $k_{su} = k_{20} = \kappa$,

$$\begin{aligned}
\tau^{(0)}_{\text{dephos}} &= -\frac{1}{\kappa}\log\left(\frac{6\delta\eta}{5\kappa s_2(0)}\right) = -\frac{1}{\kappa}W_{-1}\left(-\frac{6\delta}{5s_2(0)}\right) \\
\eta &= \frac{5\kappa s_2(0)}{6\delta}\exp\left(W_{-1}\left[-\frac{6\delta}{5s_2(0)}\right]\right) \\
&= \kappa W_{-1}\left(-\frac{6\delta}{5s_2(0)}\right)^{-1},
\end{aligned} \tag{3.143}$$

where we have applied the identity

$$\frac{1}{z}\exp\left(W(z)\right) = \frac{1}{W(z)}. \tag{3.144}$$

Equation (3.142) can then be substituted into (3.136) to obtain an iterative correction that is exact for $k_{su} = k_{20}$. We plot the composite $\tau^{(1)}_{\text{phos}}$ in Figure 3.12.

CHAPTER 4

TRANSITION PATH THEORY FOR EXCITABLE DYNAMICS

4.1 Abstract

Transition path theory (TPT) provides a means to compute path-dependent quantities of a Markov process using only local quantities. For systems exhibiting excitable dynamics, quantities of interest can depend on the history of the process. As TPT relies on the Markov property, such history-dependent quantities cannot be computed using TPT without modification. By appropriately augmenting the state space, TPT can be used as an alternative to other methods such as large deviation theory to compute statistics of excitable systems when the noise is large, and TPT can also provide mechanistic information about the dominant contributions to quantities of interest. We demonstrate the use of transition path theory to compute path-dependent quantities of a model of noise-induced bursting in neurons.

4.2 Introduction

Suppose that a system with intrinsically noisy dynamics is modeled as an Itô diffusion

$$d\mathbf{x} = \mathbf{f}(\mathbf{x})dt + \sigma d\mathbf{W}. \tag{4.1}$$

The infinitesimal generator of (4.1) is

$$\mathcal{L} \equiv \mathbf{f} \cdot \nabla + \frac{\sigma^2}{2}\nabla^2. \tag{4.2}$$

Transition path theory (TPT) concerns the computation of expectations over realizations of (4.1) using the Kolmogorov Backward Equation (KBE):

$$\frac{du}{dt} = -\mathcal{L}u(\mathbf{x}, t). \tag{4.3}$$

Conventionally, TPT has been used to compute statistics concerning transitions between two well-separated subsets A and B of the state space S. One such statistic is the mean first passage time $T(\mathbf{x})$, i.e., the expected time it takes for the system to hit set B given that the system has initial state $\mathbf{x}$. The mean first passage time satisfies

$$T(\mathbf{x}_0) = \mathbb{E}[T(\mathbf{x}_t)|\mathbf{x}_0] + t. \tag{4.4}$$

That is, if the system was initially at state $\mathbf{x}_0$ and has not hit B by time t, the expected time required to hit B is t, the amount of time the system has already spent outside of B, in addition to the expected additional time that the system must wait given that it is now in state $\mathbf{x}_t$. In terms of the infinitesimal generator, $T(\mathbf{x})$ satisfies

$$\begin{aligned} \mathbb{E}[T(\mathbf{x}_t)|\mathbf{x}_0] - T(\mathbf{x}_0) &= -t \\ \lim_{t\downarrow 0} \frac{\mathbb{E}[T(\mathbf{x}_t)|\mathbf{x}_0] - T(\mathbf{x}_0)}{t} &= -1 \\ \implies \mathcal{L}T(\mathbf{x}) &= -1. \end{aligned} \tag{4.5}$$

At the set boundary ∂B, the system is already in B, leading to the boundary value problem

$$\begin{aligned} \mathcal{L}T(\mathbf{x}) &= -1 \\ T(\partial B) &= 0. \end{aligned} \tag{4.6}$$

Another quantity central to TPT is the committor $q(\mathbf{x})$, defined as the probability that the system hits B before hitting A given that the system has initial state $\mathbf{x}$. The committor satisfies

$$\begin{aligned} \mathbb{E}[q(\mathbf{x}_t)|\mathbf{x}_0] &= q(\mathbf{x}_0) \\ \lim_{t\downarrow 0} \frac{\mathbb{E}[q(\mathbf{x}_t)|\mathbf{x}_0] - q(\mathbf{x}_0)}{t} &= 0 \\ \implies \mathcal{L}q(\mathbf{x}) &= 0. \end{aligned} \tag{4.7}$$

Equation (4.7) is a statement of conservation of probability. At the set boundaries,

$$\begin{aligned} q(\partial A) &= 0 \\ q(\partial B) &= 1. \end{aligned} \tag{4.8}$$

The boundary conditions (4.8) simply state that the system will always hit B before A if it is already in B, and has no chance of hitting B before A if it is already in A.

The mean first passage time and committor are both expectations of functionals of trajectories of (4.1) that satisfy

$$f = g(t) + f \circ \theta_t, \tag{4.9}$$

where the functional f can be decomposed into the sum of $g(t)$, a contribution from the trajectory at all times up to t, and $f \circ \theta_t$ is the same functional evaluated after shifting the starting time of the trajectory to t, i.e., the same functional evaluated on a fresh copy of the trajectory restarted at $\mathbf{x}_t$. For the first passage time, $g(t) = t$ is the time the trajectory has spent outside of B up until now. In the case of the committor f is 1 if the trajectory hits B before A and 0 otherwise, and $g = 0$; the value of the functional is determined entirely by the terminal state of the trajectory. Taking the expectation of (4.9) gives

$$\begin{aligned} u(\mathbf{x}_0) &= \mathbb{E}[g(t)] + \mathbb{E}[u(\mathbf{x}_t)|\mathbf{x}_0] \\ \mathbb{E}[u(\mathbf{x}_t)|\mathbf{x}_0] - u(\mathbf{x}_0) &= -\mathbb{E}[g(t)] \\ \lim_{t \downarrow 0} \frac{\mathbb{E}[u(\mathbf{x}_t)|\mathbf{x}_0] - u(\mathbf{x}_0)}{t} &= \lim_{t \downarrow 0} -\frac{\mathbb{E}[g(t)]}{t} \\ \mathcal{L}u(\mathbf{x}) &= -g'(0). \end{aligned} \tag{4.10}$$

Functionals that cannot be additively decomposed as in (4.9) cannot be computed as a boundary value problem (BVP) involving the generator of the process but can still be of physical interest. In such cases, augmenting the original state space with additional history information can restore additivity and allow expectations of the functional to be computed as a BVP involving the generator of the augmented process. We provide two examples of such quantities that may be of interest when

studying systems that exhibit excitable dynamics, such as those that appear in neuroscience and immunology.

4.2.1 Expected maximum deviation

Let A be a neighborhood of a reference state $\mathbf{x}^*$. In the deterministic limit, $\mathbf{x}^*$ might be the fixed point of an ODE. Let T_A be the first time that the process hits A, and assume that $T_A < \infty$ with probability 1. Let $d_A(\mathbf{x})$ be the distance of $\mathbf{x}$ to A under some metric. We seek to compute

$$m_A(\mathbf{x}) = \mathbb{E}\left[\sup_{0\le t\le T_A} d_A(\mathbf{x}_t)\middle|\mathbf{x}_0 = \mathbf{x}\right]. \tag{4.11}$$

Equation (4.11) can be decomposed into contributions from all past times and a contribution from a fresh trajectory initiated from $\mathbf{x}_t$, but the decomposition is not additive:

$$m_A(\mathbf{x}) = \mathbb{E}\left[\max\left(\sup_{0\le s\le t} d_A(\mathbf{x}_s), m_A(\mathbf{x}_t)\right)\middle|\mathbf{x}_0 = \mathbf{x}\right]. \tag{4.12}$$

The term $\sup_{0\le s\le t} d_A(\mathbf{x}_s)$ is the maximum excursion size up until t, but it is combined with the contribution from the restarted trajectory through $\max(\cdot,\cdot)$ and not through addition. Additivity can be achieved if we define a process on the augmented state space

$$\Omega = \{\mathbf{x}, r : d_A(\mathbf{x}) \le r\}, \tag{4.13}$$

where r tracks the running supremum of $d_A(\mathbf{x})$ in the original process. The augmented state space Ω has the structure of a cone (Figure 4.1). A slice through the cone for fixed $r = \rho$ contains all the points $\mathbf{x}$ in the original state space at most ρ away from A:

$$B_\rho = \{\mathbf{x} : d_A(\mathbf{x}) \le \rho\}. \tag{4.14}$$

Within each B_ρ, $\mathbf{x}_t$ has the same dynamics as in the original system. If a transition $\mathbf{x}_t \to \mathbf{x}_{t+dt}$ would result in $d_A(\mathbf{x}_{t+dt}) > r_t$, then r_{t+dt} is updated to $d_A(\mathbf{x}_{t+dt})$. Otherwise, $r_{t+dt} = r_t$. By

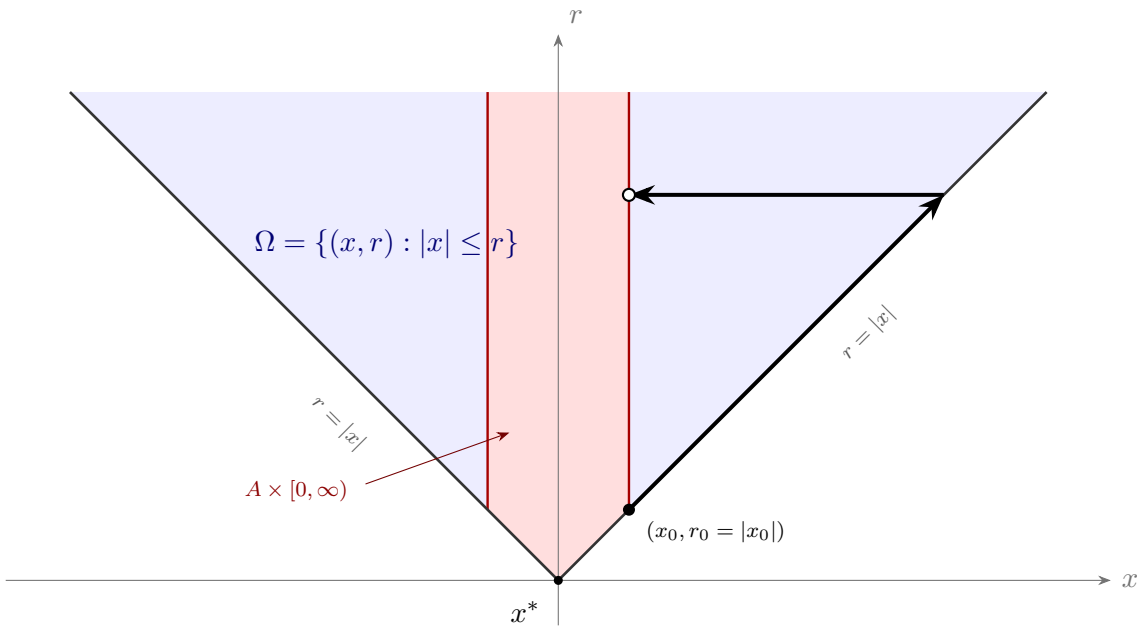


Figure 4.1: The state space of the augmented process when computing the expected maximum deviation for a one-dimensional x. For a fixed $r = \rho$, the augmented process freely moves between all x with $|\mathbf{x}| < \rho$. The history variable r increments whenever a transition is made from x that would cause $|x| > r$. The absorbing set A becomes $A \times [0, \infty)$ in the augmented state space.

construction, r is nondecreasing. The functional f on trajectories of the augmented process is simply the value of r_{T_A}, and $g = 0$ as in the case for the committor. This leads to the boundary value problem over the augmented state space

$$\begin{aligned} \mathcal{L}_{\text{aug}} m_A(\mathbf{x}, r) &= 0 \\ m_A(\partial A, r) &= r, \end{aligned} \tag{4.15}$$

where $\mathcal{L}_{\text{aug}}$ is the generator of the augmented process.

In practice, there is a more robust way to compute m_A than solving (4.15). Since $r_{T_A} \geq r_t$ for every realization of the augmented process,

$$\begin{aligned} r_{T_A} - r_0 &= \int_{r_0}^{\infty} \mathbb{1}(r_{T_A} > \rho) d\rho \\ m_A(\mathbf{x}_0, r_0) &= r_0 + \mathbb{E}[r_{T_A} - r_0 | \mathbf{x}_0, r_0] \\ &= r_0 + \int_{r_0}^{\infty} \mathbb{E}[\mathbb{1}(r_{T_A} > \rho) | \mathbf{x}_0, r_0] d\rho \\ &= r_0 + \int_{r_0}^{\infty} \mathbb{P}(r_{T_A} > \rho | \mathbf{x}_0, r_0) d\rho \end{aligned} \tag{4.16}$$

The integrand $\mathbb{P}(r_{T_A} > \rho | \mathbf{x}_0, r_0)$ is the probability that, given a trajectory is at $\mathbf{x}_0$ and has previously hit B_{r_0}, the trajectory hits B_ρ before being absorbed in A. If $r \geq \rho$, this probability must be 1. However, the integration limits of (4.16) require that $d_A(\mathbf{x}) \leq r_0 \leq \rho$. In this case, $\mathbb{P}(r_{T_A} > \rho | \mathbf{x}_0, r_0)$ as a hitting probability is independent of $\mathbf{x}_t$ for $t < 0$ and thus of r_0, and $\mathbb{P}(r_{T_A} > \rho | \mathbf{x}_0) = q_\rho(\mathbf{x}_0)$ is the committor for B_ρ so that

$$\begin{aligned} m_A(\mathbf{x}_0, r_0) &= r_0 + \int_{r_0}^{\infty} q_\rho(\mathbf{x}_0) d\rho \\ m_A(\mathbf{x}_0) &\equiv m_A(\mathbf{x}_0, d_A(\mathbf{x}_0)). \end{aligned} \tag{4.17}$$

Using (4.17), $m_A(\mathbf{x})$ can be computed from a sequence of regular committors using the generator of the original system; no explicit augmentation with r is required. The committor $q_\rho(\mathbf{x})$ has the interpretation of the complementary cumulative distribution function that a trajectory starting from $\mathbf{x}$ makes an excursion of at least size ρ. Numerically, it is easier to solve $q_\rho(\mathbf{x})$ for a nested sequence

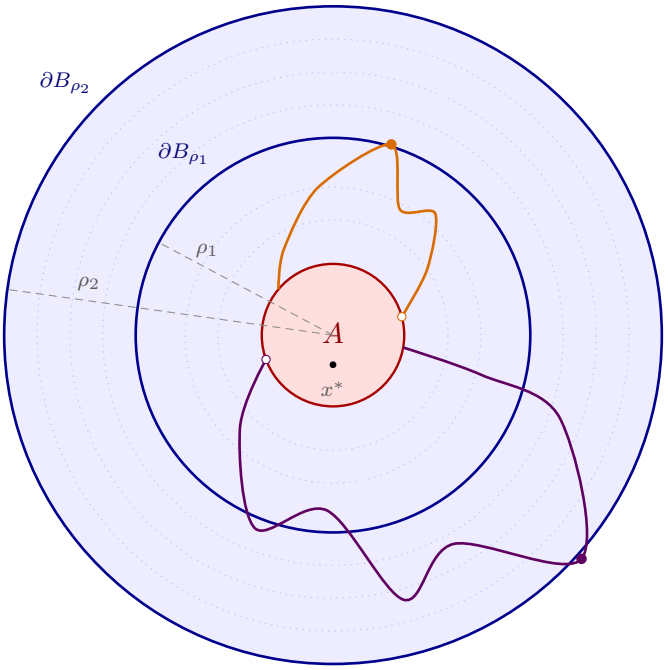


Figure 4.2: A diagram of the nested sets involved in the committor sequence when solving for the expected maximum deviation using (4.17). The probability that a trajectory makes an excursion of at least ρ before returning to A is given by $q_\rho(\mathbf{x})$, the committor to the set B_ρ, which can be interpreted as the complementary cumulative distribution function that a trajectory makes an excursion of at least size ρ. The expected maximum deviation is computed by integrating $q_\rho(\mathbf{x})$ over ρ.

of B_ρ (Figure 4.2) and then integrate over ρ with quadrature than it is to discretize and solve (4.15) directly.

4.2.2 Expected excursion count

For this quantity the presentation is clearer when working with the discretized process. Let $\mathbf{L}$ denote the generator of a discrete Markov process, i.e., the transition rate matrix. Some neurons exhibit bursting dynamics, where the neuron fires several times in rapid succession before returning to a resting state. The transition between bursting and resting states can also be driven by noise. Abstractly, this can described with three subsets A, B, C of the state space S. We seek to compute the expected number of $A \to B$ transitions made by the system before being absorbed in C. Let N_t be the number of $A \to B$ transitions made by time t and T_C be the first time that the system

hits C. At first glance, the transition count can be decomposed as

$$N_{T_C} = N_t + (N_{T_C} - N_t). \tag{4.18}$$

Where $N_{T_C} - N_t$ would be the number of future transitions for a process currently at $\mathbf{x}_t$. This decomposition, however, is not actually of the form (4.9). The number of future transitions $N_{T_C} - N_t$, cannot be computed from a trajectory restarted from $\mathbf{x}_t$; N_{T_C} depends on which of A or B the process last visited. If the process last visited A, then N increments if the process next visits B, whereas if the process last visited B, then the process must first visit A before visiting B again for N to increment. This ambiguity can be resolved by defining the augmented state space

$$\Omega \equiv S \times \{0, 1\}. \tag{4.19}$$

The augmented state variables are then

$$\begin{aligned} &(\mathbf{x}, w) \in \Omega \\ &w = \begin{cases} 0 & \text{if system was last in } A \\ 1 & \text{if system was last in } B \end{cases}. \end{aligned} \tag{4.20}$$

On the augmented state space, (4.18) can be unambiguously written as a sum between past contributions and a contribution from a trajectory restarted at $(\mathbf{x}_t, w)$. The augmented state space has the structure shown in Figure 4.3. The augmented state space has a discrete axis corresponding to the value of the binary variable w. The sets S_1 and S_2 are copies of $S \backslash B$ and $S \backslash A$, respectively, that contain all of the internal transitions in the original S. A_1 and B_2 are copies of A and B stripped of all internal transitions. This reflects the change in y upon making a transition from a state A or B. The directional flow allows the expected number of $A \to B$ transitions to be computed as the expected number of times that a trajectory hits B_1 before being absorbed in C in the augmented

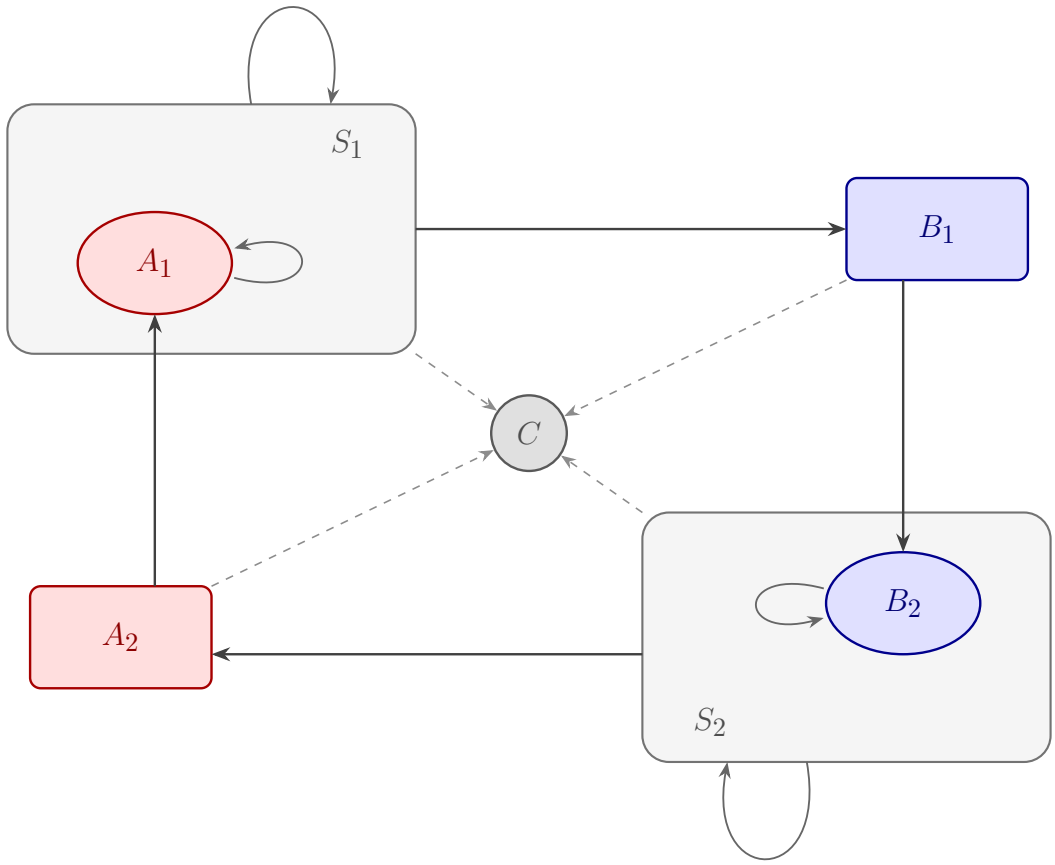


Figure 4.3: The state space of the augmented process when computing the expected excursion count. S_1 and S_2 are copies of $S \setminus B$ and $S \setminus A$, respectively. A_2 and B_1 are copies of A and B stripped of all internal transitions as to avoid overcounting; states in A_2 and B_1 can only transition to their neighbors in $A_1 \subset S_1$ and $B_2 \subset S_2$. All states in S_1 and S_2 (including those in A_1 and B_2) retain their internal transitions. Transitions to the absorbing state C occur at some state-dependent rate.

system. The generator of the augmented process is

$$\mathbf{L}_{\text{aug}} = \begin{pmatrix} \mathbf{L}_{11} & \mathbf{L}_{1B} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{L}_{BB} & \mathbf{L}_{B2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{L}_{22} & \mathbf{L}_{2A} \\ \mathbf{L}_{A1} & \mathbf{0} & \mathbf{0} & \mathbf{L}_{AA} \end{pmatrix}, \tag{4.21}$$

where $\mathbf{L}_{11}$ and $\mathbf{L}_{22}$ contain the rates of transitions internal to S_1 and S_2. The block $\mathbf{L}_{1B}$ contains all transitions from S_1 to B_1, and the blocks $\mathbf{L}_{B2}, \mathbf{L}_{2A}, \mathbf{L}_{A1}$ are defined analogously. The blocks $\mathbf{L}_{AA}$ and $\mathbf{L}_{BB}$ are diagonal such that

$$\begin{aligned} (\mathbf{L}_{BB})_{ii} &= -\sum_j (\mathbf{L}_{B2})_{ij} \\ (\mathbf{L}_{AA})_{ii} &= -\sum_j (\mathbf{L}_{A1})_{ij}. \end{aligned} \tag{4.22}$$

The expected number of visits to B_1 is then

$$\mathbb{E}[N|\mathbf{x}_0, w_0] = \text{diag}(\mathbf{L}_{\text{aug}})\mathbf{L}_{\text{aug}}^{-1}\mathbb{1}(B_1), \tag{4.23}$$

where $\text{diag}(\mathbf{L}_{\text{aug}})$ is a diagonal matrix sharing diagonal elements with $\mathbf{L}_{\text{aug}}$. The matrix $\text{diag}(\mathbf{L}_{\text{aug}})\mathbf{L}_{\text{aug}}^{-1}$ is the fundamental matrix of the augmented process, i.e., the inverse of the generator of embedded discrete time chain. Having computed (4.23), the second moment can be computed as

$$\mathbb{E}\left[N^2 \middle| \mathbf{x}_0, w_0\right] = \text{diag}(\mathbf{L}_{\text{aug}})\mathbf{L}_{\text{aug}}^{-1}\text{diag}(\mathbb{1}(B_1))(2\mathbb{E}[N|\mathbf{x}_0, y_0] - \mathbf{1}). \tag{4.24}$$

4.3 Numerical examples

We consider the hybrid system

$$dx = (y - \alpha x^3 + \beta x^2 - z + I)dt + \sigma dW_x \tag{4.25a}$$

$$dy = \left(\gamma - cx^2 - y\right) dt + \sigma dW_y \tag{4.25b}$$

$$z \in \{0, 3\}, \tag{4.25c}$$

Table 4.1: Parameters for (4.25) and (4.26).

Parameter	Numerical value
α	1
β	3
I	3
γ	1
c	5
κ_{13}	0.01
λ_{13}	2
κ_{31}	0.1
λ_{13}	1.5
σ	0.1

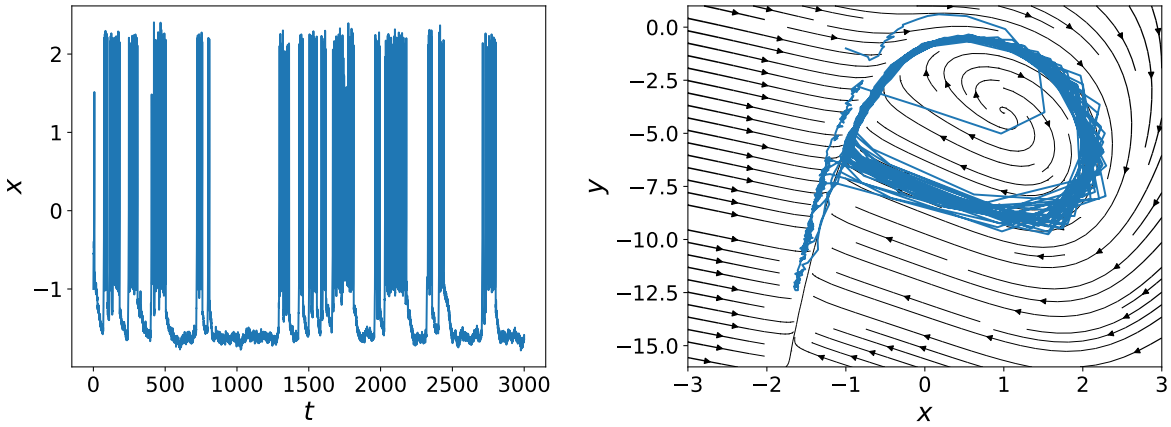


Figure 4.4: A realization of (4.25). The phase portrait is overlaid on the vector field in the case of $z = 0$.

where z transitions between 0 and 3 with rates

$$\begin{aligned} k_{13} &= \kappa_{13}(a_{13} + \exp(\lambda_{13}x)) \\ k_{31} &= \kappa_{31}(a_{31} + \exp(\lambda_{31}x)). \end{aligned} \tag{4.26}$$

Values for the parameters that appear in equations (4.25) and (4.26) are given in Table 4.1. Equation (4.25) originates from the Hindmarsh–Rose model of bursting neurons. The z variable can be interpreted as an extremely coarse discretization of the dynamic bifurcation parameter z that dictates whether the system is quiet or bursting in the original continuous dynamical system. We further discretize (4.25) using upwind finite volumes. A realization of (4.25) is plotted in Figure 4.4. We define the sets

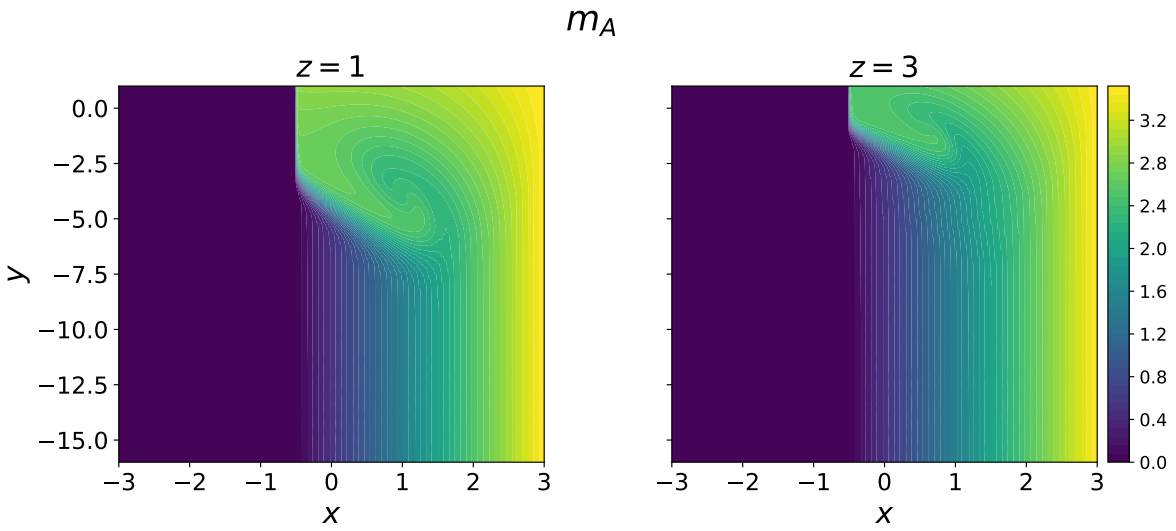


Figure 4.5: The expected maximum excursion size computed for (4.25) using the set definitions (4.27) and the metric (4.28).

$$\begin{aligned} A &= \{(x, y, z) : x \leq -1/2\} \\ B &= \{(x, y, z) : x \geq 2\} \\ C &= \{(x, y, 3)\} \end{aligned} \tag{4.27}$$

For the expected maximum excursion, we use the metric

$$d_A(x, y, z) = \inf_A \{|x_A - x| : (x_A, y_A, z_A) \in A\}. \tag{4.28}$$

The expected excursion size is plotted in Figure 4.5 and the expected excursion count is plotted in Figure 4.6. Both quantities have the most structure around $(1, -4)$, which is a fixed point of the vector field in the case of $z = 1$. The vector field shown in in Figure 4.4 spirals outward from $(1, -4)$. Where the vector field points in the $+x$ direction, trajectories are sent towards larger values of x and m_A is significantly larger than the baseline of d_A. Starting points near $(1, -4)$ with y slightly larger than -4 produce trajectories that turn around sooner, resulting in smaller values of m_A. A similar effect can be observed in Figure 4.6 for N_{T_C}. When $w = 0$, trajectories initiated near $(1, -4)$ that turn around too soon can miss B and get sent back towards A. Such trajectories must make another attempt at reaching B, during which they can also be absorbed

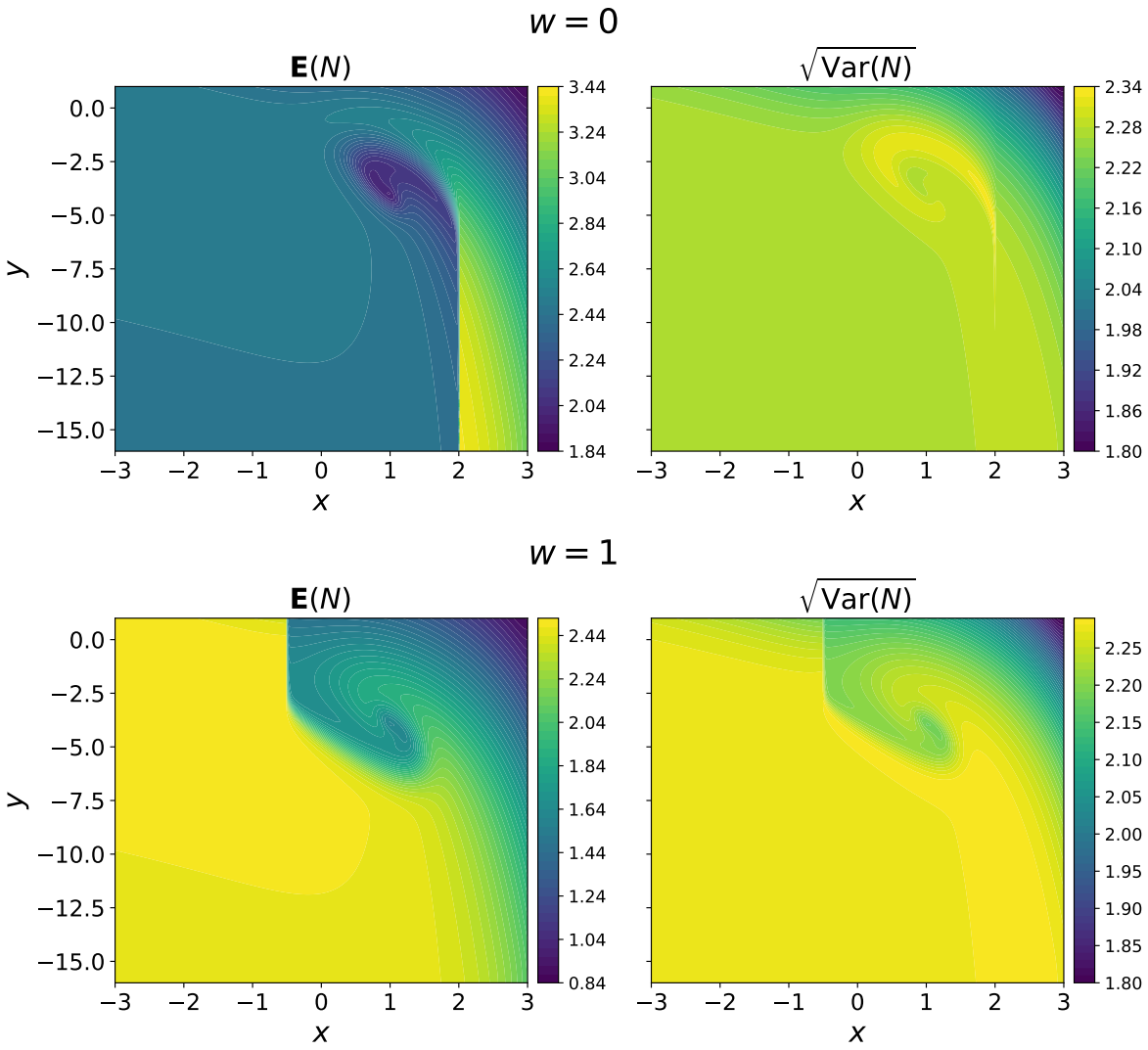


Figure 4.6: The expectation and standard deviation of the number of $A \to B$ excursions for the system (4.25) with the set definitions (4.27).

in C, resulting in a lower value of N_{T_C} relative to trajectories initiated in the rest of the domain. When $w = 1$, the opposite is true; trajectories must hit A and then B to register another excursion. Consequently, trajectories that turn around sooner can reach A and attempt a transition towards B sooner, whereas trajectories that make it deeper into B take longer to hit A and are more likely to be absorbed before registering another excursion. The requirement that the sets A and B must be visited sequentially also results in a difference in scale between the plots for $w = 0$ and $w = 1$: when $w = 0$, N_{T_C} is approximately 1 greater than when $w = 1$, since when $w = 1$, the system must return to A before registering the next excursion. The expected excursion count is slightly lower than the trajectory in Figure 4.4 suggests, since z occasionally flips rapidly from $3 \to 1 \to 3$, but x does not make an excursion, lowering the average.

4.4 Discussion

In this chapter we have introduced a framework for computing expectations of functionals of realizations of a stochastic dynamical system. By appropriately augmenting the state space to track the history of trajectories, a richer set of functionals can be computed than is possible with conventional TPT. For high dimensional problems, the uniform discretization used to compute the numerical examples becomes computationally costly. This can be addressed by computing Monte Carlo estimates of the integrals involved in discretizing the BVP as is done in the Dynamical Galerkin Approximation (DGA) [125] when applying TPT to molecular simulation. As an alternative approach, large deviations theory has been applied by [73] to approximate the firing rate of a noisy Morris-Lecar neuron as the asymptotic rate of escape from a quasipotential well around the rest state of the neuron. This approach scales well with the dimension and does not require solving a sampling problem as is required for DGA, but is only asymptotically accurate in the noise level. The two approaches can be used complementarily: large deviations theory can provide an asymptotic estimate of the stationary density of the process and most probable transition path. These estimates can then be used to initiate sampling for a DGA calculation.

CHAPTER 5

CONCLUSIONS AND OUTLOOK

In this dissertation, we have developed and applied methods for analyzing the dynamics of biological systems. Specifically, in Chapter 2, we developed a specialized MCMC method for inferring parameters of systems of ODEs that scales more favorably to higher dimensional problems than conventional methods, which we apply to a model of the Kai oscillator in Chapter 3. Our sampling results make an unexpected prediction regarding the mechanism by which period robustness arises in the Kai oscillator. We confirm the prediction experimentally and investigate the mechanism in more detail using asymptotic analysis. In Chapter 4 we consider the case of noisy dynamics, and show how history dependent statistics can be computed by augmenting the state space of the system.

The MCMC sampler developed in Chapter 2 borrows heavily from ideas used in numerical bifurcation theory and trajectory optimization. More ideas can be incorporated for further improvements to the efficiency of the sampler. Branches of solutions emerge from certain bifurcations of the ODE. One case directly relevant to our work is the emergence of a branch of limit cycles from Hopf bifurcations. Branch switching methods in numerical bifurcation analysis initiate continuation of a branch of limit cycles upon encountering a Hopf bifurcation; for our applications, this can be used to perform constrained sampling of fixed points until encountering a Hopf bifurcation, and then switch to sampling limit cycles to rapidly move through regions of low posterior probability without needing to perform multiple stages of sampling. The switching procedure would need special treatment to preserve the reversibility of the sampling process. When we solve the BVP constraint during the limit cycle sampling stage, we include the mesh points as state variables to be sampled and move them such that the mesh is the finest where a monitor function is high. This is to improve the convergence of the BVP solver in parameter regimes where the ODE solution exhibits sharp transitions. However, the parameters of the ODE can be tuned to make the problem arbitrarily stiff, and a mesh of fixed size will eventually encounter regions that cause the BVP solver to fail, stalling the entire sampler. The BVP solver can be made more robust by introducing an h adaptive mesh [1], where additional mesh intervals are introduced where the solution has the sharpest

features. Such a Monte Carlo move would alter the dimension of the space being sampled, and the sampler would need to incorporate a transdimensional move as in Reversible Jump MCMC or be implemented as a Multilevel Monte Carlo algorithm [57, 61]. A complementary approach is to switch to an asymptotic approximation of the ODE when such an approximation would be accurate. Stiff solutions tend to arise where the parameters of the ODE impose a large scale separation, which conveniently is where asymptotic approximations of the system are accurate. Even if an asymptotic approximation does not admit analytical solutions, the asymptotic ODEs are often scale free and do not exhibit the stiffness that plagues the exact system. In general, it is difficult to automatically identify the dominant balances for matched asymptotic expansions of a generic ODE, but for ODEs with polynomial vector fields of moderate size there has been some success in automatically identifying dominant balances using the Newton polytope of the polynomial. In the event that the sampler does stall, the sampler could switch to a sampling a distribution where the BVP constraints have been relaxed to soft restraints in hopes of returning to a region in parameter space where the constraints can again be imposed exactly, but as before this requires special treatment to preserve the reversibility of the sampler.

Our results regarding the Kai oscillator in Chapter 3 raise further questions about the effect of the T432 phosphorylation site on the dynamics of the clock. For our optimized parameters, oscillations tend to fail through a supercritical Hopf bifurcation at high KaiA concentrations, where the limit cycles shrinks to a point. This amplitude scaling is consistent with experimental measurements as the KaiA concentration is increased [26]. Deleting the T432 phosphorylation site changes the bifurcation structure of the model. Instead of failing through a supercritical Hopf bifurcation, the single site minimal model often fails through SNIC or homoclinic bifurcations that cause the period to diverge. The period of T432V KaiC does lengthen significantly as KaiA is increased while the amplitude does not appear to change by much, but our data for T432 KaiC was not collected on a fine enough grid of KaiA concentrations to make a definite statement about the scaling of the period and amplitude. Further measurements at a finer resolution of KaiA concentrations can enable a more precise comparison with the behavior of the model and provide evidence regarding whether the T432V mutation alters the bifurcation structure of the physical system.

The examples in Chapter 4 were computed by discretizing the Kolmogorov Backward Equation using the finite volume method and then solving the resulting linear systems. For higher dimensional models, such as models of neurons whose dynamics are driven by ion channels that undergo state transitions, direct discretization is numerically intractable. This problem has been encountered in molecular simulation, where TPT is used to analyze protein dynamics. The Dynamical Galerkin Approximation [125], in which the inner products that appear in the weak form of the KBE are computed using MCMC, has been successfully applied to systems with hundreds of dimensions [58]. The same approach can be adapted for the analysis of stochastic dynamical systems.

REFERENCES

[1] *Adaptive Mesh Refinement - Theory and Applications: Proceedings of the Chicago Workshop on Adaptive Mesh Refinement Methods, Sept. 3–5, 2003*. Springer Berlin Heidelberg, 2005.

[2] H. C. Andersen. Rattle: A "velocity" version of the shake algorithm for molecular dynamics calculations. *Journal of Computational Physics*, 52(1):24–34, oct 1983.

[3] C. Andrieu, A. Doucet, and R. Holenstein. Particle Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 72(3):269–342, jun 2010.

[4] N. Barkai and S. Leibler. Robustness in simple biochemical networks. *Nature*, 387(6636):913–917, 1997.

[5] G. Barles, C. Daher, and M. Romano. Optimal control on the l^∞ norm of a diffusion process. *SIAM Journal on Control and Optimization*, 32(3):612–634, May 1994.

[6] E. Barth, K. Kuczera, B. Leimkuhler, and R. D. Skeel. Algorithms for constrained molecular dynamics. *Journal of Computational Chemistry*, 16(10):1192–1209, oct 1995.

[7] D. Battogtokh, D. K. Asch, M. E. Case, J. Arnold, and H.-B. Schüttler. An ensemble method for identifying regulatory circuits with special reference to the qa gene cluster of Neurospora crassa. *Proceedings of the National Academy of Sciences*, 99(26):16904–16909, 2002.

[8] B. BELOUSOV. A periodic reaction and its mechanism. *Ref. Radiats. Med.*, 1958.

[9] A. Beskos, N. Pillai, G. Roberts, J.-M. Sanz-Serna, and A. Stuart. Optimal tuning of the hybrid Monte Carlo algorithm. *Bernoulli*, 19(5A):1501–1534, nov 2013.

[10] J. T. Betts. Survey of numerical methods for trajectory optimization. *Journal of Guidance, Control, and Dynamics*, 21(2):193–207, Mar. 1998.

[11] M. J. Blandamer and S. H. Morris. Investigation into the effect of temperature and added t-butyl alcohol on the dynamic properties of the belousov reaction. *Journal of the Chemical Society, Faraday Transactions 1: Physical Chemistry in Condensed Phases*, 71(0):2319, 1975.

[12] O. Bokanowski, A. Picarelli, and H. Zidani. Dynamic programming and error estimates for stochastic control problems with maximum cost. *Applied Mathematics amp; Optimization*, 71(1):125–163, May 2014.

[13] J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, G. Necula, A. Paszke, J. VanderPlas, S. Wanderman-Milne, and Q. Zhang. JAX: composable transformations of Python+NumPy programs, 2018.

[14] W. C. Bray and H. A. Liebhafsky. REACTIONS INVOLVING HYDROGEN PEROXIDE, IODINE AND IODATE ION. i. INTRODUCTION. *Journal of the American Chemical Society*, 53(1):38–44, jan 1931.

[15] C. Brettschneider, R. J. Rose, S. Hertel, I. M. Axmann, A. J. R. Heck, and M. Kollmann. A sequestration feedback determines dynamics and temperature entrainment of the kaiabc circadian clock. *Molecular Systems Biology*, 6(1), Jan. 2010.

[16] K. S. Brown and J. P. Sethna. Statistical mechanical approaches to models with many poorly known parameters. *Physical Review E*, 68(2):021904, 2003.

[17] M. Brubaker, M. Salzmann, and R. Urtasun. A family of MCMC methods on implicitly defined manifolds. *Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics, PMLR*, 22:161–172, 2012.

[18] T. Bui-Thanh and M. Girolami. Solving large-scale PDE-constrained Bayesian inverse problems with Riemann manifold Hamiltonian Monte Carlo. *Inverse Problems*, 30(11):114014, Oct. 2014.

[19] T. Bui-Thanh and Q. P. Nguyen. FEM-based discretization-invariant MCMC methods for PDE-constrained Bayesian inverse problems. *Inverse Problems and Imaging*, 10(4):943–975, Oct. 2016.

[20] A. Cabezas, A. Corenflos, J. Lao, R. Louf, A. Carnec, K. Chaudhari, R. Cohn-Gordon, J. Coullon, W. Deng, S. Duffield, G. Durán-Martín, M. Elantkowski, D. Foreman-Mackey, M. Gregori, C. Iguaran, R. Kumar, M. Lysy, K. Murphy, J. C. Orduz, K. Patel, X. Wang, and R. Zinkov. BlackJAX: Composable Bayesian inference in JAX. *arXiv preprint arXiv:2402.10797*, 2024.

[21] E. Campillo-Funollet, C. Venkataraman, and A. Madzvamuse. Bayesian parameter identification for Turing systems on stationary and evolving domains. *Bulletin of Mathematical Biology*, 81(1):81–104, Oct. 2018.

[22] G. F. Carey. *Computational Grids: Generation, Adaptation, and Solution Strategies*. CRC Press, Boca Raton, United States, 1997.

[23] T. F. C. Chan and H. B. Keller. Arc-length continuation and multigrid techniques for nonlinear elliptic eigenvalue problems. *SIAM Journal on Scientific and Statistical Computing*, 3(2):173–194, 1982.

[24] Y.-G. Chang, S. E. Cohen, C. Phong, W. K. Myers, Y.-I. Kim, R. Tseng, J. Lin, L. Zhang, J. S. Boyd, Y. Lee, S. Kang, D. Lee, S. Li, R. D. Britt, M. J. Rust, S. S. Golden, and A. LiWang. A protein fold switch joins the circadian oscillator to clock output in cyanobacteria. *Science*, 349(6245):324–328, jun 2015.

[25] G. Chauvet. A comparison of pivotal sampling and unequal probability sampling with replacement. *Statistics & Probability Letters*, 121:1–5, feb 2017.

[26] A. G. Chavan, J. A. Swan, J. Heisler, C. Sancar, D. C. Ernst, M. Fang, J. G. Palacios, R. K. Spangler, C. R. Bagshaw, S. Tripathi, P. Crosby, S. S. Golden, C. L. Partch, and A. LiWang. Reconstitution of an intact clock reveals mechanisms of circadian timekeeping. *Science*, 374(6564), Oct. 2021.

[27] J. Chew, E. Leypunskiy, J. Lin, A. Murugan, and M. J. Rust. High protein copy number is required to suppress stochasticity in the cyanobacterial circadian clock. *Nature Communications*, 9(1), aug 2018.

[28] C. Chi, J. Weare, and A. R. Dinner. Sampling parameters of ordinary differential equations with constrained langevin dynamics. *SIAM/ASA Journal on Uncertainty Quantification*, 13(3):1374–1405, Aug. 2025.

[29] S. E. Cohen and S. S. Golden. Circadian rhythms in cyanobacteria. *Microbiology and Molecular Biology Reviews*, 79(4):373–385, sep 2015.

[30] S. L. Cotter, G. O. Roberts, A. M. Stuart, and D. White. MCMC methods for functions: Modifying old algorithms to make them faster. *Statistical Science*, 28(3), Aug. 2013.

[31] K. Cranmer, J. Brehmer, and G. Louppe. The frontier of simulation-based inference. *Proceedings of the National Academy of Sciences*, 117(48):30055–30062, 2020.

[32] J. A. G. M. de Visser, J. Hermisson, G. P. Wagner, L. A. Meyers, H. Bagheri-Chaichian, J. L. Blanchard, L. Chao, J. M. Cheverud, S. F. Elena, W. Fontana, G. Gibson, T. F. Hansen, D. Krakauer, R. C. Lewontin, C. Ofria, S. H. Rice, G. v. Dassow, A. Wagner, and M. C. Whitlock. Perspective: Evolution and detection of genetic robustness. *Evolution*, 57(9):1959–1972, 2003.

[33] P. Deuflhard. *Newton Methods for Nonlinear Problems*. Springer Berlin Heidelberg, Berlin, Germany, 2011.

[34] A. R. Dinner, E. H. Thiede, B. V. Koten, and J. Weare. Stratification as a general variance reduction method for Markov chain Monte Carlo. *SIAM/ASA Journal on Uncertainty Quantification*, 8(3):1139–1188, 2020.

[35] E. DOEDEL, H. B. KELLER, and J. P. KERNEVEZ. NUMERICAL ANALYSIS AND CONTROL OF BIFURCATION PROBLEMS (II): BIFURCATION IN INFINITE DIMENSIONS. *International Journal of Bifurcation and Chaos*, 01(04):745–772, dec 1991.

[36] M. Domijan and M. Kirkilionis. Bistability and oscillations in chemical reaction networks. *Journal of Mathematical Biology*, 59(4):467–501, nov 2008.

[37] A. Doucet, N. Freitas, and N. Gordon, editors. *Sequential Monte Carlo Methods in Practice*. Springer New York, New York, United States, 2001.

[38] M. B. Elowitz and S. Leibler. A synthetic oscillatory network of transcriptional regulators. *Nature*, 403(6767):335–338, 2000.

[39] H. Errami, M. Eiswirth, D. Grigoriev, W. M. Seiler, T. Sturm, and A. Weber. Detection of hopf bifurcations in chemical reaction networks using convex coordinates. *Journal of Computational Physics*, 291:279–302, jun 2015.

[40] H. Eydgahi, W. W. Chen, J. L. Muhlich, D. Vitkup, J. N. Tsitsiklis, and P. K. Sorger. Properties of cell death models calibrated and compared using Bayesian approaches. *Molecular Systems Biology*, 9(1):644, 2013.

[41] R. J. Field and R. M. Noyes. Oscillations in chemical systems. iv. limit cycle behavior in a model of a real chemical reaction. *Journal of Chemical Physics*, 60(5):1877–1884, 1974.

[42] P. Flaherty, M. L. Radhakrishnan, T. Dinh, R. A. Rebres, T. I. Roach, M. I. Jordan, and A. P. Arkin. A dual receptor crosstalk model of G-protein-coupled signal transduction. *PLOS Computational Biology*, 4(9):e1000185, 2008.

[43] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman. emcee: The MCMC hammer. *Publications of the Astronomical Society of the Pacific*, 125(925):306–312, Mar 2013.

[44] Y. Furuike, A. Mukaiyama, D. Ouyang, K. Ito-Miwa, D. Simon, E. Yamashita, T. Kondo, and S. Akiyama. Elucidation of master allostery essential for circadian clock oscillation in cyanobacteria. *Science Advances*, 8(15), Apr. 2022.

[45] M. Gabrié, G. M. Rotskoff, and E. Vanden-Eijnden. Adaptive Monte Carlo augmented with normalizing flows. *Proceedings of the National Academy of Sciences*, 119(10):e2109420119, 2022.

[46] A. Garbuno-Inigo, N. Nüsken, and S. Reich. Affine invariant interacting Langevin dynamics for Bayesian inference. *SIAM Journal on Applied Dynamical Systems*, 19(3):1633–1658, Jan. 2020.

[47] K. Gatermann, M. Eiswirth, and A. Sensse. Toric ideals and graph theory to analyze hopf bifurcations in mass action systems. *Journal of Symbolic Computation*, 40(6):1361–1382, dec 2005.

[48] A. Gelman and D. B. Rubin. Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4), Nov. 1992.

[49] A. George, M. T. Heath, and E. Ng. Solution of sparse underdetermined systems of linear equations. *SIAM Journal on Scientific and Statistical Computing*, 5(4):988–997, dec 1984.

[50] C. J. Geyer. Practical Markov chain Monte Carlo. *Statistical Science*, 7(4), Nov. 1992.

[51] D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry*, 81(25):2340–2361, dec 1977.

[52] M. Girolami and B. Calderhead. Riemann manifold Langevin and Hamiltonian Monte Carlo methods. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 73(2):123–214, Mar. 2011.

[53] A. Goldbeter and R. Lefever. Dissipative structures for an allosteric model. *Biophysical Journal*, 12:1302–1315, 1972.

[54] A. Golden, J. Paijmans, and D. K. Lubensky. Distinguishing feedback mechanisms in clock models. *arXiv: Molecular Networks*, 2016.

[55] J. Goodman and J. Weare. Ensemble samplers with affine invariance. *Communications in Applied Mathematics and Computational Science*, 5(1):65–80, Jan 2010.

[56] W. J. F. Govaerts. *Numerical Methods for Bifurcations of Dynamical Equilibria*. Society for Industrial and Applied Mathematics, Philadelphia, United States, 2000.

[57] P. J. GREEN. Reversible jump markov chain monte carlo computation and bayesian model determination. *Biometrika*, 82(4):711–732, 1995.

[58] S. C. Guo, R. Shen, B. Roux, and A. R. Dinner. Dynamics of activation in the voltage-sensing domain of ciona intestinalis phosphatase ci-vsp. *Nature Communications*, 15(1), Feb. 2024.

[59] R. N. Gutenkunst, J. J. Waterfall, F. P. Casey, K. S. Brown, C. R. Myers, and J. P. Sethna. Universally sloppy parameter sensitivities in systems biology models. *PLOS Computational Biology*, 3(10):e189, Oct. 2007.

[60] T. S. Hatakeyama and K. Kaneko. Generic temperature compensation of biological clocks by autonomous regulation of catalyst concentration. *Proceedings of the National Academy of Sciences*, 109(21):8109–8114, may 2012.

[61] S. Heinrich. *Multilevel Monte Carlo Methods*, page 58–67. Springer Berlin Heidelberg, 2001.

[62] M. E. HENDERSON. MULTIPLE PARAMETER CONTINUATION: COMPUTING IMPLICITLY DEFINED k-MANIFOLDS. *International Journal of Bifurcation and Chaos*, 12(03):451–476, mar 2002.

[63] M. D. Hoffman, A. Radul, and P. Sountsov. An adaptive-MCMC scheme for setting trajectory lengths in Hamiltonian Monte Carlo. In *International Conference on Artificial Intelligence and Statistics*, 2021.

[64] M. D. Homan and A. Gelman. The No-U-Turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, 15(1):1593–1623, jan 2014.

[65] L. Hong, D. O. Lavrentovich, A. Chavan, E. Leypunskiy, E. Li, C. Matthews, A. LiWang, M. J. Rust, and A. R. Dinner. Bayesian modeling reveals metabolite-dependent ultrasensitivity in the cyanobacterial circadian clock. *Molecular Systems Biology*, 16(6), jun 2020.

[66] A. M. Horowitz. A generalized guided Monte Carlo algorithm. *Physics Letters B*, 268(2):247–252, Oct. 1991.

[67] W. Huang and R. D. Russell. *Adaptive Moving Mesh Methods*. Springer New York, 2011.

[68] S. Hug, A. Raue, J. Hasenauer, J. Bachmann, U. Klingmüller, J. Timmer, and F. Theis. High-dimensional Bayesian parameter estimation: Case study for a model of JAK2/STAT5 signaling. *Mathematical Biosciences*, 246(2):293–304, 2013.

[69] Y. Hwang, H. J. Kim, W. Chang, C. Hong, and S. N. MacEachern. Bayesian model calibration and sensitivity analysis for oscillating biological experiments. *Technometrics*, pages 1–11, 2025.

[70] M. Ishiura, S. Kutsuna, S. Aoki, H. Iwasaki, C. R. Andersson, A. Tanabe, S. S. Golden, C. H. Johnson, and T. Kondo. Expression of a gene cluster kaiabc as a circadian feedback process in cyanobacteria. *Science*, 281(5382):1519–1523, 1998.

[71] H. Ito, H. Kageyama, M. Mutsuda, M. Nakajima, T. Oyama, and T. Kondo. Autonomous synchronization of the circadian KaiC phosphorylation rhythm. *Nature Structural amp; Molecular Biology*, 14(11):1084–1088, Oct. 2007.

[72] A. Kazarnikov and H. Haario. Statistical approach for parameter identification by Turing patterns. *Journal of Theoretical Biology*, 501:110319, Sept. 2020.

[73] J. P. Keener and J. M. Newby. Perturbation analysis of spontaneous action potential initiation by stochastic ion channels. *Physical Review E*, 84(1), 2011.

[74] Y. B. Kiyohara, M. Katayama, and T. Kondo. A novel mutation in KaiC affects resetting of the cyanobacterial circadian clock. *Journal of Bacteriology*, 187(8):2559–2564, Apr. 2005.

[75] A. Kramer, B. Calderhead, and N. Radde. Hamiltonian Monte Carlo methods for efficient parameter estimation in steady state dynamical systems. *BMC Bioinformatics*, 15(1), jul 2014.

[76] B. Kramer and K. E. Willcox. Nonlinear model order reduction via lifting transformations and proper orthogonal decomposition. *AIAA Journal*, 57(6):2297–2307, jun 2019.

[77] B. Krauskopf, H. M. Osinga, and J. Galán-Vioque. *Numerical continuation methods for dynamical systems*. Springer Netherlands, Dordrecht, The Netherlands, 2007.

[78] A. Kröner, A. Picarelli, and H. Zidani. Infinite horizon stochastic optimal control problems with running maximum cost. *SIAM Journal on Control and Optimization*, 56(5):3296–3319, Jan. 2018.

[79] C. Kuehn. *Multiple Time Scale Dynamics*. Springer International Publishing, 2015.

[80] G. Kurosawa, K. Aihara, and Y. Iwasa. A model for the circadian rhythm of cyanobacteria that maintains oscillation without gene expression. *Biophysical Journal*, 91(6):2015–2023, Sept. 2006.

[81] A. Kværnø. Singly diagonally implicit Runge–Kutta methods with an explicit first stage. *BIT Numerical Mathematics*, 44(3):489–502, 2004.

[82] S.-H. Lee, K. Palmo, and S. Krimm. WIGGLE: A new constrained molecular dynamics algorithm in Cartesian coordinates. *Journal of Computational Physics*, 210(1):171–182, Nov. 2005.

[83] B. Leimkuhler and C. Matthews. *Molecular Dynamics: With Deterministic and Stochastic Numerical Methods*. Springer New York, New York, United States, 2015.

[84] B. Leimkuhler and C. Matthews. Efficient molecular dynamics using geodesic integration and solvent–solute splitting. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 472(2189):20160138, may 2016.

[85] B. Leimkuhler, C. Matthews, and J. Weare. Ensemble preconditioning for Markov chain Monte Carlo simulation. *Statistics and Computing*, 28(2):277–290, Feb. 2017.

[86] T. Lelièvre, M. Rousset, and G. Stoltz. Hybrid Monte Carlo methods for sampling probability measures on submanifolds. *Numerische Mathematik*, 143(2):379–421, jun 2019.

[87] T. Lelièvre, G. Stoltz, and W. Zhang. Multiple projection Markov chain Monte Carlo algorithms on submanifolds. *IMA Journal of Numerical Analysis*, 43(2):737–788, Mar. 2022.

[88] E. Leypunskiy, J. Lin, H. Yoo, U. Lee, A. R. Dinner, and M. J. Rust. The cyanobacterial circadian clock follows midday in vivo and in vitro. *eLife*, 6, July 2017.

[89] J. Lin, J. Chew, U. Chockanathan, and M. J. Rust. Mixtures of opposing phosphorylations within hexamers precisely time feedback in the cyanobacterial circadian clock. *Proceedings of the National Academy of Sciences*, 111(37):E3937–E3945, sep 2014.

[90] Y. Liu, G. K. Pattanayak, C. Chi, H. Yoo, A. R. Dinner, and M. J. Rust. An underlying bistability sets amplitude and explains temperature compensation in the cyanobacterial circadian clock. *Biorxiv: the Preprint Server for Biology*, 2026.

[91] D. J. MacKay. *Information Theory, Inference and Learning Algorithms*. Cambridge University Press, Cambridge, United Kingdom, 2003.

[92] O. Mangoubi and N. K. Vishnoi. Dimensionally tight bounds for second-order Hamiltonian Monte Carlo. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, NIPS'18, page 6030–6040, Red Hook, NY, USA, 2018. Curran Associates Inc.

[93] C. Matthews, J. Weare, A. Kravtsov, and E. Jennings. Umbrella sampling: A powerful method to sample tails of distributions. *Monthly Notices of the Royal Astronomical Society*, 480(3):4069–4079, 2018.

[94] B. A. Mello, W. Pan, G. L. Hazelbauer, and Y. Tu. A dual regulation mechanism of histidine kinase CheA identified by combining network-dynamics modeling and system-level input-output data. *PLOS Computational Biology*, 14(7):e1006305, 2018.

[95] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, 21(6):1087–1092, June 1953.

[96] F. Miyoshi, Y. Nakayama, K. Kaizu, H. Iwasaki, and M. Tomita. A mathematical model for the kai-protein-based chemical oscillator and clock gene expression rhythms in cyanobacteria. *Journal of Biological Rhythms*, 22:69–80, 2007.

[97] P. D. Moral. Measure-valued processes and interacting particle systems. application to nonlinear filtering problems. *The Annals of Applied Probability*, 8(2):438–495, may 1998.

[98] F. Morgan. *Geometric measure theory: a beginner's guide*. Academic press, London, United Kingdom, 2016.

[99] Y. Murakami. Bayesian parameter inference and model selection by population annealing in systems biology. *PLOS ONE*, 9(8):e104057, 2014.

[100] M. Nakajima. Reconstitution of circadian oscillation of cyanobacterial KaiC phosphorylation in vitro. *Science*, 308(5720):414–415, apr 2005.

[101] R. M. Neal. Annealed importance sampling. *Statistics and Computing*, 11:125–139, 2001.

[102] Y. Nesterov. Gradient methods for minimizing composite functions. *Mathematical Programming*, 140(1):125–161, Dec. 2012.

[103] T. Nishiwaki, Y. Satomi, Y. Kitayama, K. Terauchi, R. Kiyohara, T. Takao, and T. Kondo. A sequential program of dual phosphorylation of kaic as a basis for circadian rhythm in cyanobacteria. *The EMBO Journal*, 26(17):4029–4037, Aug. 2007.

[104] B. Novák and J. J. Tyson. Design principles of biochemical oscillators. *Nature Reviews Molecular Cell Biology*, 9(12):981–991, 2008.

[105] R. M. Noyes, L. V. Kalachev, and R. J. Field. Mathematical model of the bray-liebhafsky oscillations. *Journal of Physical Chemistry*, 99(11):3514–3520, 1995.

[106] M. Ottobre, N. S. Pillai, F. J. Pinski, and A. M. Stuart. A function space HMC algorithm with second order Langevin diffusion limit. *Bernoulli*, 22(1):60–106, 2016.

[107] J. Paijmans, D. K. Lubensky, and P. R. ten Wolde. Period robustness and entrainability of the kai system to changing nucleotide concentrations. *Biophysical Journal*, 113(1):157–173, Jul 2017.

[108] J. Paijmans, D. K. Lubensky, and P. R. ten Wolde. A thermodynamically consistent model of the post-translational kai circadian clock. *PLOS Computational Biology*, 13(3):e1005415, mar 2017.

[109] R. Pattanayek and M. Egli. Protein–protein interactions in the cyanobacterial circadian clock: Structure of kaia dimer in complex with c-terminal kaic peptides at 2.8 Å resolution. *Biochemistry*, 54:4575–4578, 2015.

[110] R. Pattanayek, J. Wang, T. Mori, Y. Xu, C. H. Johnson, and M. Egli. Visualizing a circadian clock protein: Crystal structure of kaic and functional insights. *Molecular Cell*, 15:0–388, 2004.

[111] C. Phong, J. S. Markson, C. M. Wilhoite, and M. J. Rust. Robust and tunable circadian rhythms from differentially sensitive catalytic domains. *Proceedings of the National Academy of Sciences*, 110(3):1124–1129, dec 2012.

[112] T. Plewa, T. Linde, and V. G. Weirs. *Adaptive Mesh Refinement - Theory and Applications: Proceedings of the Chicago Workshop on Adaptive Mesh Refinement Methods, Sept. 3–5, 2003.* Springer Berlin Heidelberg, Berlin, Germany, 2005.

[113] N. Pullen and R. J. Morris. Bayesian model comparison and parameter inference in systems biology using nested sampling. *PLOS ONE*, 9(2):e88419, 2014.

[114] L. Pyta and D. Abel. Nonlinear model reduction of the navier-stokes-equations. In *2016 American Control Conference (ACC)*. IEEE, jul 2016.

[115] L. Riou-Durand and J. Vogrinc. Metropolis adjusted Langevin trajectories: a robust alternative to Hamiltonian Monte Carlo. *arXiv preprint arXiv:2202.13230*, 2022.

[116] M. J. Rust, S. S. Golden, and E. K. O. Shea. Light-driven changes in energy metabolism directly entrain the cyanobacterial circadian oscillator. *Science*, 331(6014):220–223, jan 2011.

[117] M. J. Rust, J. S. Markson, W. S. Lane, D. S. Fisher, and E. K. O'Shea. Ordered phosphorylation governs oscillation of a three-protein circadian clock. *Science*, 318:809–812, 2007.

[118] J.-P. Ryckaert, G. Ciccotti, and H. J. Berendsen. Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics*, 23(3):327–341, Mar. 1977.

[119] L. Shepp and A. N. Shiryaev. The russian option: Reduced regret. *The Annals of Applied Probability*, 3(3), Aug. 1993.

[120] L. A. Shepp and A. N. Shiryaev. A new look at pricing of the "russian option". *Theory of Probability amp; Its Applications*, 39(1):103–119, Jan. 1995.

[121] H. Smith. Oscillations and multiple steady states in a cyclic gene model with repression. *Journal of Mathematical Biology*, 25(2):169–190, June 1987.

[122] A. Sokal. *Monte Carlo Methods in Statistical Mechanics: Foundations and New Algorithms*, page 131–192. Springer US, 1997.

[123] Stan Development Team. Stan modeling language users guide and reference manual, version 2.36.0, 2024.

[124] R. H. Swendsen and J.-S. Wang. Replica Monte Carlo simulation of spin-glasses. *Physical Review Letters*, 57(21):2607–2609, Nov. 1986.

[125] E. H. Thiede, D. Giannakis, A. R. Dinner, and J. Weare. Galerkin approximation of dynamical quantities using trajectory data. *The Journal of Chemical Physics*, 150(24), 2019.

[126] T. Y.-C. Tsai, Y. S. Choi, W. Ma, J. R. Pomerening, C. Tang, and J. E. Ferrell. Robust, tunable biological oscillations from interlinked positive and negative feedback loops. *Science*, 321(5885):126–129, 2008.

[127] R. Tseng, N. F. Goularte, A. Chavan, J. Luu, S. E. Cohen, Y.-G. Chang, J. Heisler, S. Li, A. K. Michael, S. Tripathi, S. S. Golden, A. LiWang, and C. L. Partch. Structural basis of the day-night transition in a bacterial circadian clock. *Science*, 355(6330):1174–1180, mar 2017.

[128] C. Tsitouras. Runge–Kutta pairs of order 5(4) satisfying only the first column simplifying assumption. *Computers & Mathematics with Applications*, 62(2):770–775, 2011.

[129] J. J. Tyson and P. C. Fife. Target patterns in a realistic model of the Belousov–Zhabotinskii reaction. *Journal of Chemical Physics*, 73:2224–2237, 1980. 10.1063/1.440418.

[130] J. S. van Zon, D. K. Lubensky, P. R. H. Altena, and P. R. ten Wolde. An allosteric model of circadian KaiC phosphorylation. *Proceedings of the National Academy of Sciences*, 104(18):7420–7425, May 2007.

[131] V. Vyshemirsky and M. A. Girolami. Bayesian ranking of biochemical system models. *Bioinformatics*, 24(6):833–839, 2008.

[132] S. Wang, K. Fan, N. Luo, Y. Cao, F. Wu, C. Zhang, K. A. Heller, and L. You. Massive computational acceleration by using neural networks to emulate mechanism-based biological models. *Nature Communications*, 10(1):4354, 2019.

[133] A. Weerasinghe. Controlling the running maximum of a diffusion process and an application to queueing systems. *SIAM Journal on Control and Optimization*, 56(2):1412–1440, Jan. 2018.

[134] K. Xu and M. Holmes-Cerfon. Monte Carlo on manifolds in high dimensions. *Journal of Computational Physics*, 506:112939, June 2024.

[135] T.-R. Xu, V. Vyshemirsky, A. Gormand, A. von Kriegsheim, M. Girolami, G. S. Baillie, D. Ketley, A. J. Dunlop, G. Milligan, M. D. Houslay, and W. Kolch. Inferring signaling pathway topologies from multiple perturbation measurements of specific biochemical species. *Science Signaling*, 3(113):ra20–ra20, 2010.

[136] T. Yoshida, Y. Murayama, H. Ito, H. Kageyama, and T. Kondo. Nonparametric entrainment of the in vitro circadian phosphorylation rhythm of cyanobacterial KaiC by temperature cycle. *Proceedings of the National Academy of Sciences*, 106(5):1648–1653, jan 2009.

[137] Y. Zhang and C. Sutton. Quasi-Newton methods for Markov chain Monte Carlo. In J. Shawe-Taylor, R. Zemel, P. Bartlett, F. Pereira, and K. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 24. Curran Associates, Inc., 2011.