Long-time analytic approximation of large stochastic oscillators: simulation, analysis and inference

Giorgos Minas[1] and David A Rand[1,2,*]

[1] Systems Biology Centre and [2] Mathematics Institute,

University of Warwick, Coventry CV4 7AL, UK*

Abstract

There is currently a great need for analytical tools and accurate approximation methods for large complex stochastic dynamical models such as those oscillators studied in systems biology. We present a new stochastic approximation of biological oscillators that allows such an approach. To do this we analyse the failure of the fast and analytically tractable Linear Noise Approximation (LNA) and use this understanding and dynamical systems perturbation theory to develop a modified LNA, called phase-corrected LNA (pcLNA) that overcomes the main limitations of the standard LNA providing approximations uniformly accurate for long times, which are still fast and analytically tractable. As part of this, we develop analytical expressions for key probability distributions and associated quantities, such as the Fisher Information Matrix and Kullback-Leibler divergence, which can be used to analyse the system's stochastic sensitivities and information geometry. We also present algorithms for statistical inference and for long-term simulation of oscillating systems. We use a model of the drosophila circadian clock for illustration and comparisons of pcLNA with exact simulations.

Keywords: stochastic models — oscillations — Linear Noise Approximation — biological systems — Fisher information — Kalman filter

Abbreviations: pcLNA, phase-corrected Linear Noise Approximation; KS, Kolmogorov-Smirnov; FIM, Fisher Information Matrix

1 Introduction

Dynamic cellular oscillating systems such as the cell cycle, circadian clock and other signaling and regulatory systems have complex structures, highly nonlinear dynamics and are subject to

*to whom correspondence may be addressed

both intrinsic and extrinsic stochasticity. Moreover, current models of these systems have high-dimensional phase spaces and many parameters. Modelling and analysing them is therefore a challenge, particularly if one wants to take account of stochasticity and develop a more analytical approach enabling the development of theory and quantification in a more controlled way than is possible by simulation alone. The stochastic kinetics that arise due to random births, deaths and interactions of individual species give rise to Markov jump processes that, in principle, can be analyzed by means of master equations. However, these are rarely tractable and although an exact numerical simulation algorithm is available [1], for the large systems we are interested in, this is very slow.

It is therefore important to develop accurate approximation methods for such systems that enable a more analytical approach as well as offering faster simulation and better algorithms for data fitting and parameter estimation. One obvious candidate for this is the Linear Noise Approximation (LNA). This is based on a systematic approximation of the master equation by means of van Kampen's Ω-expansion [4]. Its large volume ($\Omega \to \infty$) validity has been shown in [5], in the sense that the distribution of the Markov jump process at a fixed finite time converges, as $\Omega \to \infty$, to the LNA probability distribution. The latter distribution is analytically tractable allowing for fast estimation and simulation algorithms. However, the LNA has significant limitations, particularly, as we show below, in approximating long-term behaviour of oscillatory systems.

Our aim in this paper is to introduce a modified LNA, called the phase-corrected LNA, or pcLNA, for oscillatory systems that overcomes the most important shortcomings of the LNA. As part of this, (a) we present a simulation algorithm that is uniformly accurate for all times and very fast compared to exact simulation, (b) we show that we can effectively calculate and analyse the probability distributions of phase states (such as timing of maxima or minima) of the system and associated quantities, such as the Fisher Information Matrix and Kullback-Leibler divergence, thus allowing calculation of its stochastic sensitivities and analysis of its information geometry, and (c) we facilitate estimation of system parameters θ by showing that, given appropriate data Y, we can accurately approximate the likelihood function $L(\theta; Y)$ via a Kalman filter.

A key advantage of this approach is that one obtains approximations that are uniform in time. To understand how this is possible consider a simpler problem, a system that in the large volume, $\Omega \to \infty$, limit has a stable attracting equilibrium. Under very general conditions, for time $T > 0$ large and large Ω, the LNA will give a good approximation of the true distribution $P(Y(t)|t \gg 0)$. Thus, under these conditions, the distribution that provides a good approximation for a specific large T will be a similarly good approximation for all $T \gg 0$. Our approach exploits the fact that the distributions for a general class of systems with a stable attracting limit cycle in the $\Omega \to \infty$ limit have a similar property if one conditions this distribution on appropriate transversal sections to the limit cycle. For these systems and any transversal section $\mathcal{S}$ to the limit cycle our modified LNA (pcLNA) provides a distribution that can be analytically calculated and that is a good approximation to the true distribution $P(Y(t)|Y(t) \in \mathcal{S}, t \gg 0)$ uniformly in time.

To do this we build on previous work of Boland et al. [6]. They use the 2-dimensional Brusselator system as an exemplar to investigate the failure of the LNA in approximating long-term behavior of oscillatory systems and present a method for computing power spectra and comparing exact simulations with LNA predictions of the same phase rather than time. We extend these results in a number of ways including the following: (i) we develop a theory that treats the general case and provide analytical arguments which justify our approximations and enable computation of trajectory distributions, (ii) we show that the approach is practicable for large nonlinear systems and (iii) we provide practical algorithms to simulate systems, estimate parameters and analytically calculate key quantities including probability distributions and information metrics. The approach in [6] uses transversal sections which are normal to the limit cycle. We show that for most considerations one can use any transversal to the limit cycle.

To illustrate and validate our approach we apply it to a relatively large published stochastic model of the Drosophila circadian clock due to Gonze et al. [7] (see SI). This model involves 10 state variables and 30 reactions. The oscillations are driven by the negative feedback exerted on the *per* and *tim* genes by the complex formed from PER and TIM proteins following phosphorylation. *per* mRNA (M_P) and *tim* mRNA (M_T) is transported into the cytosol where it is degraded and translated into protein (P_0 and T_0). These proteins are multiply phosphorylated (PER: $P_0 \to P_1 \to P_2$; TIM: $T_0 \to T_1 \to T_2$) and these modifications can be reversed by a phosphatase. The fully phosphorylated form of the proteins is targeted for degradation and forms a complex, which is transported into the nucleus in a reversible manner where the nuclear form of the PER–TIM complex represses the transcription of *per* and *tim* genes. The large system limit is given by the differential equation system of 10 kinetic equations that are listed in the supplementary information (SI) along with the reaction scheme of the system.

The 2-dimensional Brusselator system is also used for illustrating our methods and the results can be found in the SI. The SI also contains technical derivations and further illustrative figures which we refer to in this paper.

1.1 Mathematical preliminaries

Stochastic models of cellular processes in signaling and regulatory systems are usually described in terms of reaction networks. A system of multiple different molecular subpopulations has state vector, $Y(t) = (Y_1(t), \ldots, Y_n(t))^T$ where $Y_i(t)$, $i = 1, \ldots, n$, denotes the number of molecules of each species. These molecules undergo a number of possible reactions (e.g. transcription, degradation of the mRNA, translation, degradation of the protein) where the reaction of index j changes $Y(t)$ to $Y(t) + \nu_j$ with the vectors $\nu_j \in \mathbb{R}^n$ called stoichiometric vectors. Each reaction occurs at a rate $w_j(Y(t))$ (often called the intensity of the reaction) which is a function of $Y(t)$. The stoichiometry matrix S is the matrix whose columns are the ν_j, $j = 1, 2, \ldots, m$.

We explicitly take into account the volume Ω of the reaction system by assuming that the rates $w_j(Y(t))$ depend upon the volume Ω as $w_j(Y) = u_j(Y/\Omega, \Omega)$ where the dependence of u_j

on Ω is of canonical form in the sense of [4], that is $u_j(x,\Omega) = f(\Omega)\sum_{m=0}^{\infty}\Omega^{-m}u_j^{(m)}(x)$. Under this assumption, the LNA as formulated by [5] is derived directly from the underlying Markov jump process and is valid for any time interval of finite fixed length. It is based on the Ansatz

$$X(t) = \frac{Y(t)}{\Omega} = x(t) + \frac{\xi(t)}{\sqrt{\Omega}} \tag{1}$$

where $x(t)$ is a solution of the differential equation that describes the limiting $\Omega \to \infty$ system i.e.

$$\dot{x} = F(x), \quad F(x) = \sum_j \nu_j u_j(x(t)). \tag{2}$$

We will be interested in the case where the solution $x(t)$ of interest is a stable limit cycle of minimal period $\tau > 0$ given by $x = g(t)$, $0 \le t \le \tau$.

To describe the statistics of ξ we need some mathematical preliminaries. Let $J(x)$ denote the $n \times n$ Jacobian matrix of F so that $(J(x))_{ij} = \partial F_i/\partial x_j$ where the partial derivative is evaluated at x. Let $C(s,t)$ be the family of $n \times n$ fundamental matrices which are the solutions of the differential equation

$$\frac{d}{dt}C(s,t) = J(g(t))C(s,t), \quad C(s,s) = I. \tag{3}$$

The fundamental or transition matrix has the properties $C(t_1,t)C(t_0,t_1) = C(t_0,t)$ for all $t_0 \le t_1 \le t$ and $C(s,t) = C(0,t)C(0,s)^{-1}$.

Since g is periodic, $J(g(t+\tau)) \equiv J(g(t))$ and therefore by Floquet theory [9] $C(s,t) = Z(t)e^{(t-s)R}$ where $Z(t+\tau) \equiv Z(t)$, R only depends upon s and where the eigenvalues λ_i of $C(s,s+\tau) = e^{\tau R}$ are independent of s. In the case of a non-degenerate attracting limit cycle of a free-running oscillator (i.e. when F does not depend directly upon t), the matrix $e^{\tau R}$ has $\lambda_1 = 1$ as a simple eigenvalue and the remaining eigenvalues satisfy $|\lambda_i| < 1$. The tangent vector at the initial value x_0 is an eigenvector associated with λ_1. In the case of periodically forced oscillators the limiting differential equation is $\dot{x} = F(t,x)$ and for an entrained non-degenerate attracting limit cycle solution all the eigenvalues λ satisfy $|\lambda| < 1$.

In the LNA, the stochastic variable ξ in (1) satisfies

$$\xi(t) = C(t_0,t)\xi(t_0) + \eta(t_0,t) \tag{4}$$

for all $t_0 < t$ where $\eta(t_0,t) \sim \mathrm{MVN}(0, V(t_0,t))$ is multivariate normal with mean 0 and covariance matrix

$$V(t_0,t) = \int_{t_0}^{T} C(s,t)E(s)E(s)^T C(s,t)^T ds \tag{5}$$

and $E(s)$ the square root of the matrix product of the stoichiometry matrix S introduced above and $W(x(s))$, a diagonal matrix with main diagonal the reaction rates $u_j(x(s))$. Henceforth, we write $\zeta \sim \mathrm{MVN}(m,S)$ to mean that a random variable ζ is multivariate normal with mean m and covariance matrix S.

The locus γ in $\mathbb{R}^n$ of the points $g(t)$ on the limit cycle is topologically a circle (possibly with intersection points in the case of entrained forced oscillator). We will be particularly interested in the probability distributions of the intersections of a stochastic trajectory with transversal sections to γ. By a transversal section through $x \in \gamma$ we mean a $(n-1)$-dimensional linear hyperplane $\mathcal{S}_x$ containing x and transversal to the tangent vector, $F(x)$, to γ at x. A particular example is the hyperplane normal to γ at x. A transversal system is a family $\mathcal{S}_{g(t)}$ of transversal sections which vary smoothly with t in the sense that the unit normal vector to $\mathcal{S}_{g(t)}$ varies smoothly with t. The *normal transversal system* is the one where $\mathcal{S}_{g(t)}$ is the normal to γ at $g(t)$. We use this for our computed examples below. However, we will also consider other interesting transversal systems in the SI. We will see below that, under general conditions and for free-running oscillators, knowledge of trajectory distributions in one transversal system allows accurate approximation of them in others. A transversal system defines a mapping G of a neighborhood of γ onto γ where if $X \in \mathcal{S}_x$ then $G(X) = x \in \gamma$. In cases where $X(s)$ lies in more than one transversal sections, $\mathcal{S}_{x(t')}$, $t' = t_1, t_2, \ldots$, then $G(X(s)) = x(t)$ with $t = \min_i |t_i - s|$ the closest time to s. We denote this mapping for the normal transversal system by G_N.

When we discuss MVN distributions on a transversal sections $\mathcal{S}_{g(t)}$ and give the covariance matrix we are always doing this (unless otherwise stated) in an adapted coordinate system. An adapted coordinate system at $g(t)$, $\mathcal{C}_{g(t)}$, is one determined by a set of orthonormal basis vectors $e_1(t), \ldots, e_n(t)$ with $e_1(t)$ the unit normal vector to $\mathcal{S}_{g(t)}$ and the vectors $e_2(t), \ldots, e_n(t)$ forming an orthonormal basis of $\mathcal{S}_{g(t)}$.

2 Motivation

2.1 Exact simulations versus the LNA

We next motivate our methodological results with some observations regarding the dynamics of oscillatory systems. We use the circadian clock system introduced above. Similar results for the Brusselator can be found in the SI. We use the Gillespie algorithm to exactly simulate the system and produce $R = 3000$ samples of stochastic trajectories (see figure 1) for a time-length of $8.5 \times \tau$, where $\tau \approx 26.98$h is the period of the limit cycle. The system volume used is $\Omega = 300$ imposing moderate to high levels of stochasticity. Results for system volumes $\Omega = 200$, 500 and 1000 are also reported in the SI.

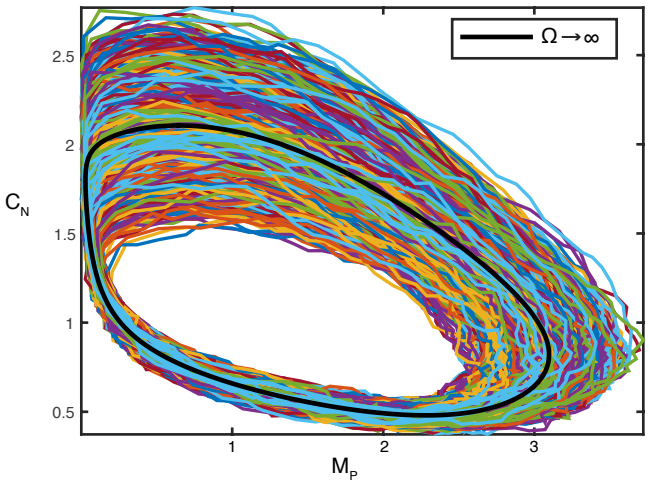


Figure 1: Stochastic trajectories obtained from the Gillespie simulation algorithm. Two (out of 10) of the species are displayed (*per* mRNA (x-axis) and nuclear PER-TIM complex (y-axis)). The volume size is $\Omega = 300$. A subsample of 300 trajectories is displayed ($R = 3000$). The black solid curve is the large volume, $\Omega \to \infty$, limit cycle solution.

We use the Kolmogorov-Smirnov (KS) statistic to quantify comparisons between different probability distributions. The KS statistic measures the distance between two scalar cumulative distribution functions (cdf's) and can be used to test the hypothesis that two distributions are equal. The latter hypothesis can be rejected on the basis of the KS test only if the value of the KS statistic is larger than a threshold level. In our examples the KS statistic is applied to each of the scalar distributions for the individual species populations as a measure of how close the multivariate distributions are.

We first compare the distribution $P(X(t)|X(0) = x_0)$ for the LNA and the Gillespie simulated samples at a sequence of times $t = \tau, 2\tau, \ldots, 8\tau$ and for an arbitrary (fixed) initial state $x_0 \in \gamma$. As we can see in figure 2, the LNA fits the Gillespie simulations relatively well in the short run ($t = \tau$), but as time progresses the KS distance increases substantially beyond the threshold level. The LNA predictions spread along the tangental direction and therefore fail to accurately reflect the Gillespie samples that have spread along the curved limit cycle.

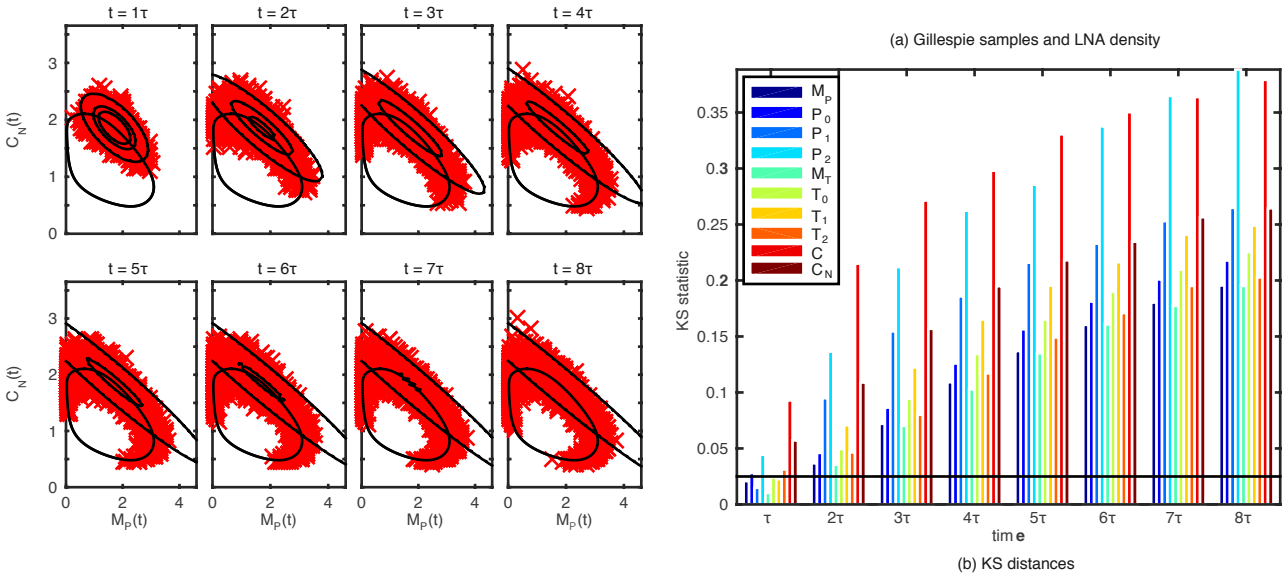


Figure 2: Comparison between LNA and exact simulations. (a) Samples obtained from the Gillespie simulation algorithm (red crosses) and $0.05, 0.5, 0.95$ contours of the LNA probability density (black ellipsoids) at fixed times, $t = \tau, 2\tau, \ldots, 8\tau$ (τ: minimal period), for the circadian clock system. The limit cycle ODE solution is also displayed (black solid line). (b) KS distance between the empirical distribution of Gillespie samples and LNA distribution of each species (different colors, see legend) at the fixed times. The threshold level is also displayed (black solid line).

2.2 Transversal distributions

Consider an exact simulation and a stochastic trajectory $X(t)$ of an oscillatory system starting from a point $X(0)$ at time t_0. If the system size is not too small, although there will be some reversals, $G(X(t))$ will move around γ in the direction of the deterministic flow given by the $\Omega \to \infty$ limit system. We are interested in the earliest intersection $Q_{x_1}^{(r)}$ of this trajectory with a given transversal section $\mathcal{S}_{x_1}$ to γ when $G(X(t))$ has done a prescribed number r of revolutions around γ. Practically, in exact simulations, the intersection points $Q_{x_1}^{(r)}$, are derived by interpolation between the last state before and the first state after the intersection (see SI section S1.1).

The transversal distribution $P(Q_{x_0}^{(r)}|X(0) = x_0)$, $r = 1, 2, \ldots$, describes the stochastic behaviour of the system in the $(n-1)$-dimensional space of system states with phase x_0, i.e. $G(X(t)) = x_0$. As we can see in figure 3, these transversal distributions are approximately normal and hardly distinguishable between different rounds which indicates a fast convergence to a fixed, approximately normal transversal distribution (see also KS distances in SI figures S18 and S19). The exact transversal distributions are also accurately approximated by the LNA

transversal distributions which are analytically derived in later sections (see (6)). Therefore, we see that, even though the full n-dimensional LNA distribution is not accurate, when we restrict it to transversals it is effectively indistinguishable from that of exact simulations. This restricted LNA distribution is multivariate normal.

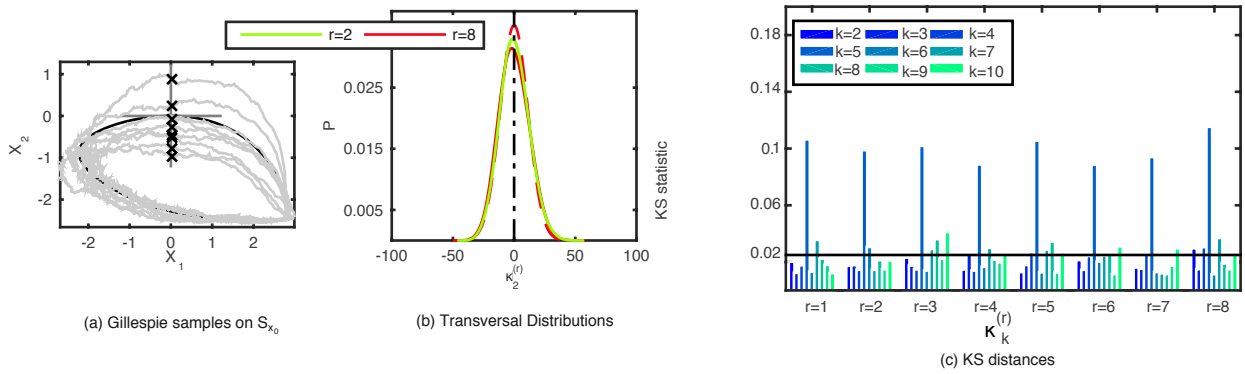


Figure 3: Distribution of oscillatory systems at transversal sections. (a) A single trajectory (grey line) of Gillespie simulation algorithm and its intersections (crosses) to a transversal section (grey cross) of the limit cycle (black line) projected on adapted coordinates e_1 and e_2. (b) The exact empirical distribution (solid lines) and LNA distribution (dashed line) of the (centered) intersections, $\kappa^{(r)} = Q_{x_0}^{(r)} - x_0$, projected on the first transversal direction e_2 (c) KS distances between exact empirical and LNA distribution of $\kappa_k^{(r)}$, $r = 1, 2, \ldots, 8$, in the k-th transversal coordinate, $k = 2, 3, \ldots, 10$with k set in order of decreasing scale.

We also study the distribution of the times of intersection, $T^{(r)} = T_{x_0}^{(r)}$, of the stochastic trajectories with the transversal section $\mathcal{S}_{x_0}$. As we can see in figure 5(a), they appear to be approximately normally distributed with mean approximately $r\tau$, but, crucially, as time progresses, their variance grows linearly. This is a characteristic of free-running oscillators. For entrained forced oscillators, the variance converges to a finite limit, but still may be sufficiently large to cause inaccuracies.

This variability accumulation explains the failure of LNA in the long run. To elaborate this further, let $\kappa(t) = \sqrt{\Omega}(X(t) - G_N(X(t)))$ so that $\kappa(t)$ is normal to γ at $G_N(X(t))$ (see figure 4). Consider the times $t = t_1, t_2, \ldots$ where $X(t)$ meets the section $\mathcal{S}_{x_0}$. Then the above observation suggests that the distribution of the $\kappa(t_i)$ is approximately normal and is independent of t providing t is not too small. Thus, while $(\kappa(t)|G_N(X(t)) = x_0)$ has a long-time stable distribution, the noise process used in the LNA, $\xi(t) = \sqrt{\Omega}(X(t) - g(t))$, does not have this property because

$$\xi(t) = \kappa(t) + \Omega^{1/2}(g(t) - G_N(X(t))),$$

and the variance of $g(t) - G_N(X(t))$ grows linearly with time. This implies the result in figure 5(c)). Therefore, long-term predictions of $\xi(t)$ have large variance due to the $\Omega^{1/2}(g(t) -$

$G_N(X(t)))$ term, but long-term prediction of $\kappa(t)$ conditional on the phase of $X(t)$ does not suffer from this. Moreover, since the LNA has a multivariate normal distribution for $X(t)$ and since the dominant variance σ_1^2 will have size which is $\mathrm{O}(t)$, when σ_1 is larger than the diameter of the limit cycle, it is clear that the LNA will be inaccurate for large times.

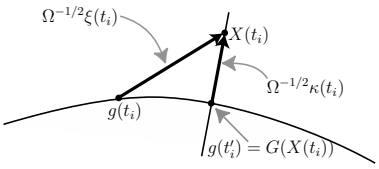
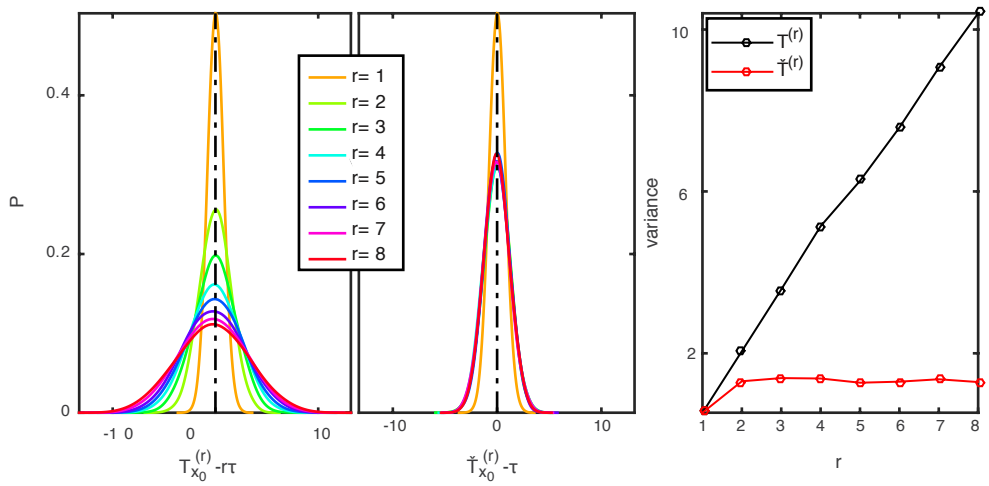


Figure 4: This illustrates the relation between $X(t_i) = Y(t_i)/\Omega$, $g(t_i')$, $\xi(t_i)$ and $\kappa(t_i)$. To each $X(t_i)$ we can associate the point on γ, $G(X(t_i))$ and the vector $\kappa(t_i)/\sqrt{\Omega}$ from this point to $X(t_i)$. However, $G(X(t_i)) = g(t_i')$ for some $0 \leq t_i' < \tau$ and thus to $X(t_i)$ we can also associate the pair $(g(t_i'), \kappa(t_i))$.

This understanding of the nature of the LNA's failure along with the observation that the corrected return times, $\check{T}_{x_0}^{(r)} = T_{x_0}^{(r)} - T_{x_0}^{(r-1)}$, have a stable normal distribution across different rounds (figure 5), motivates the following simulation algorithm to approximate exact solutions.



(a) Intersection Times

(b) Corrected times of intersection

(c) Variance of intersection times

Figure 5: Times of intersection to a transversal section. Empirical distribution of (a) the intersection times, $T_{x_0}^{(r)}$, and (b) the corrected intersection times, $\check{T}_{x_0}^{(r)}$ and their empirical variance (c) for samples derived by Gillespie simulation of the circadian clock system.

3 Results

3.1 pcLNA Simulation Algorithm

Our simulation algorithm is based on that developed in [6] for the Brusselator system. The approach is to use resetting of $g(t)$ to $G_N(X(t))$ to cope with the growth in the variance of $g(t) - G_N(X(t))$ and keep the LNA fluctuation in the transversal direction. The fact that $g(t) - G_N(X(t))$ is expected to be approximately normal enables us to do this in a well-defined way.

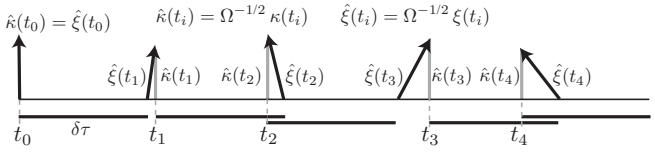


Figure 6: This illustrates the main step in the pcLNA algorithm. The solid horizontal bars below the horizontal axis are all of length $\delta\tau$. The black arrows show $\hat{\kappa}(t_i) = \Omega^{-1}\xi(t_i)$ and the grey arrows $\hat{\kappa}(t_i) = \Omega^{-1}\kappa(t_i)$.

The pcLNA simulation algorithm is described next. Its main step is illustrated in figure 6.

1. Choose a time-step size $\delta\tau > 0$.
2. Input **initial condition** $\kappa(t_0)$ and $X(t_0) = g(t_0) + \Omega^{-1/2}\kappa(t_0)$.
3. **For iteration** $i = 1, 2, \ldots$
   (a) sample $\xi(t_{i-1} + \delta\tau)$ from $\mathrm{MVN}(C_i\kappa(t_{i-1}), V_i)$;
   (b) compute $X_i = g(t_{i-1} + \delta\tau) + \Omega^{-1/2}\xi(t_{i-1} + \delta\tau)$;
   (c) set t_i to be such that $G_N(X_i) = g(t_i)$ and $\kappa(t_i) = \Omega^{1/2}(X_i - g(t_i))$.

In the for loop $C_i = C(t_{i-1}, t_{i-1} + \delta\tau)$ and $V_i = V(t_{i-1}, t_{i-1} + \delta\tau)$.

Stochastic trajectories of $N = 3000$ samples of the circadian clock system observed over a period of time-length 8.5τ are derived using the pcLNA simulation algorithm (see SI figure 2). The distribution $P(X(t)|X(0) = x_0)$ derived by the pcLNA simulation algorithm is compared with Gillespie algorithm exact simulations and the standard LNA in figure 7. As we can see, the pcLNA accurately approximates the Gillespie samples for all the considered time points, substantially improving the performance of standard LNA. Similar results are obtained for volume sizes $\Omega = 200$, 500 and 1000 (see SI figures 24, 25).

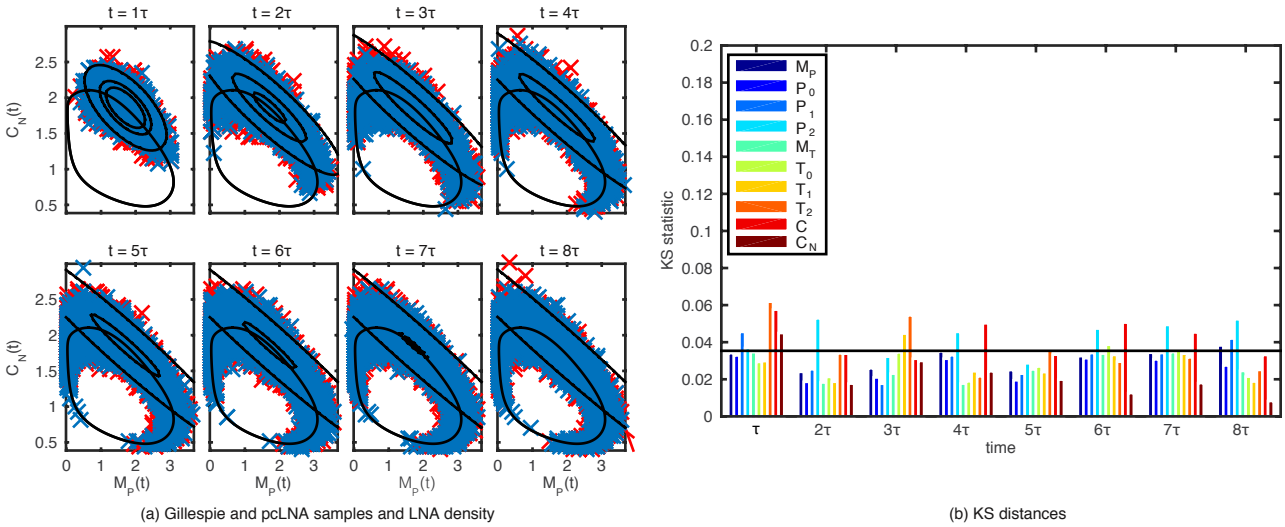


Figure 7: Comparison between Gillespie and pcLNA simulated distribution. (a) Samples obtained from the Gillespie simulation algorithm (red crosses) and the pcLNA simulation algorithm (blue crosses) and $0.05, 0.5, 0.95$ contours of the LNA distribution (black ellipsoids) at fixed times, $t = \tau, 2\tau, \ldots, 8\tau$ (τ: minimal period), for the circadian clock system. (b) KS statistics of the distance between the empirical distribution of Gillespie and pcLNA samples of each species (different colors, see legend) at the fixed times. The threshold level is also displayed (black solid line).

To understand why this algorithm is fast it is important to note that there is an efficient method to compute and store a subset of the matrices $C(s,t)$ for this system and the given parameters once and for all so that all such matrices can be computed extremely quickly from this subset. This subset is computed up-front. A Matlab routine to do this is contained with the package PeTSSy which is freely available.[1]. Thus, the only extra computational cost compared to the standard LNA is due to the derivation of $G(X)$ which can be practically performed through an optimization procedure (e.g. Newton-Raphson algorithm). In the above simulation, phase correction is performed every 6 hours ($\tau/6 \approx 4.5$ corrections in every round of the limit cycle). The effect of less frequent correction is studied in the SI section S2.5.1.

3.2 Calculation and convergence of transversal distributions

The stochastic behaviour of oscillatory systems at specific phase-states such as maxima or minima of pivotal species is of paramount importance in analysing the stochastic sensitivities of the system, its information geometry and various other aspects. The transversal distributions which provide the necessary information for such analysis are generally intractable. However,

[1] http://www2.warwick.ac.uk/fac/sci/systemsbiology/research/software/

as we showed in figure 3 (see also SI figure S4), transversal distributions can be accurately approximated by the LNA. In the following, we compute these LNA transversal distributions, which are multivariate normal and have parameters in a reasonably simple form, and show that they converge to a fixed point under fairly general conditions.

Consider q phase states of the limit cycle $x_i = g(t_i)$, $i = 1, \ldots, q$ on γ where $0 \leq t_1 < t_2 < \ldots < t_q < \tau$ and associated transversals $\mathcal{S}_{x_i}$ which are not necessarily normal and assume that these are part of a transversal system $\mathcal{S}_{g(t)}$ so that the mapping G is defined. If $X(t)$ is a stochastic trajectory, we consider how it meets the transversal sections at the x_i as t increases as described above. Suppose it first meets $\mathcal{S}_{x_i}$ in $Q^{(1)}_{x_i}$ for $i = 1, \ldots, q$. If $i < q$ then we let $Q^{(k)}_{x_{i+1}}$ denote the first point in $\mathcal{S}_{x_{i+1}}$ that X meets after it leaves $Q^{(k)}_{x_i}$. If $i = q$ then the next transversal it meets is $\mathcal{S}_{x_0}$ and the intersection point is $Q^{(k+1)}_{x_0}$. In this way we derive a sequence of points $\underline{Q} = Q^{(1)}_{x_0}, \ldots, Q^{(1)}_{x_q}, Q^{(2)}_{x_0}, \ldots, Q^{(2)}_{x_q}, \ldots, Q^{(m)}_{x_0}, \ldots, Q^{(m)}_{x_q}$.

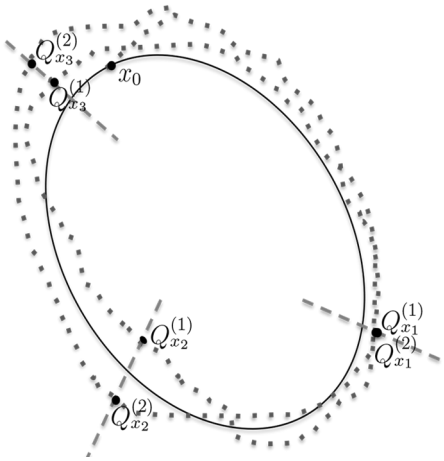


Figure 8: This illustrates the sequence Q in two-dimensions. The stochastic trajectory $X(t)$ (grey dotted line) initiated from x_0 intersects the transversal sections $\mathcal{S}_{x_1}, \mathcal{S}_{x_3}$ and $\mathcal{S}_{x_3}$ (grey dashed lines) of the limit cycle (black solid line) for the first time at $Q^{(1)}_{x_1}$, $Q^{(1)}_{x_2}$ and $Q^{(1)}_{x_3}$, respectively, and then $\mathcal{S}_{x_1}, \mathcal{S}_{x_3}$ and $\mathcal{S}_{x_3}$ at $Q^{(2)}_{x_1}$, $Q^{(2)}_{x_2}$ and $Q^{(2)}_{x_3}$, respectively, for the second time.

We shall be interested in the distribution

$$P(\underline{Q}|X(t_0)) = P(Q^{(1)}_{x_0}, \ldots, Q^{(m)}_{x_q}|X(t_0)) \tag{6}$$

and will derive an analytic expression for it.

For this calculation we need to introduce some new matrices. At each of the points x_k we fix a coordinate system $\mathcal{C}_{x_k}$ defined by a set of orthonormal vectors $e_1(x_k), \ldots, e_n(x_k)$ where $e_1(x_k)$ is normal to the hyperplane $\mathcal{S}_{x_k}$. We write coordinates in $\mathcal{C}_{x_k}$ in the form $(y_1, \boldsymbol{y}_2)$ where $y_1 \in \mathbb{R}$ and $\boldsymbol{y}_2 \in \mathbb{R}^{n-1}$ and $\boldsymbol{y}_2$ are coordinates on $\mathcal{S}_{x_0}$. The matrices $C(s,t)$ and $V(s,t)$ in this

coordinate systems are

$$C(s,t) = \begin{pmatrix} C_{11} & C_{12} \\ \\ C_{21} & C_{22} \end{pmatrix} \quad \text{and} \quad V(s,t) = \begin{pmatrix} V_{11} & V_{12} \\ \\ V_{21} & V_{22} \end{pmatrix},$$

where $V_{12} = V_{21}^T$. Let $\check{C}(s,t) = C_{22} - V_{21}V_{11}^{-1}C_{12}$ and $\check{V}(s,t) = V_{22} - V_{21}V_{11}^{-1}V_{12}$. Note that $\check{V}(s,t)$ is the Schur complement of V_{22} in $V(s,t)$ so if $(y_1, \boldsymbol{y}_2)$ is MVN with covariance $V(s,t)$ then the distribution of $\boldsymbol{y}_2$ conditional on y_1 having a specific value has covariance $\check{V}(s,t)$.

Firstly, we consider $P(Q_{x_j}^{(l)}|Q_{x_i}^{(k)})$ where either $k < l$ or $k = l$ and $i < j$. We show in the Section S1.3 of the SI that this is very well approximated by a MVN distribution with mean $\check{C}_j Q_{x_i}^{(k)}$ and covariance $\check{V}_j$ where

$$\check{C}_j = \check{C}(t_j, t_j - t_i + (l-m)\tau),\ \check{V}_j = \check{V}(t_j, t_j - t_i + (l-m)\tau).$$

We also prove that as $l - k \to \infty$, $P(Q_{x_j}^{(l)}|Q_{x_i}^{(k)})$ converges to a limit P_{x_j} which is independent of $Q_{x_i}^{(k)}$ and does this exponentially fast in the sense that the mean and covariance of $P(Q_{x_j}^{(l)}|Q_{x_i}^{(k)})$ converge exponentially fast to their limiting values.

Fixed point distribution. The special case where the transitions are returns to the same transversal sections ($i = j$) is especially interesting. Clearly, $P(Q_{x_i}^{(k)}|Q_{x_i}^{(1)})$ converges to the same limit P_{x_i} as above and this distribution is a fixed point in the sense that the distribution $P(Q_{x_i}^{(k+1)}|Q_{x_i}^{(k)} \sim P_{x_i}) = P_{x_i}$. We therefore denote it $P_{x_i}^{(\mathrm{fp})}$. Because of the aforementioned convergence the transversal distribution has the following important property

$$\lim_{k\to\infty} P(Q_{x_i}^{(k)}|X(t_0)) = P_{x_i}^{(\mathrm{fp})}$$

for any initial condition $X(t_0)$. That is, the distribution of the system at any transversal section converges to a limiting multivariate normal distribution. For the circadian clock the convergence is fast: the L^2-norm of the difference between the covariance matrices of the limiting distribution, $P_{x_0}^{(\mathrm{fp})}$ and $P(Q_{x_0}^{(r)}|X(t_0))$ for $r = 1, 2, \ldots, 5$ is respectively $(1250, 70, 3.9, 0.2) \cdot 10^{-3}$.

This distribution can be easily calculated numerically. To derive its mean and covariance matrix one solves the equations $m = \check{C}_i m$ and $S = \check{C}_i S \check{C}_i^{\,T} + \check{V}_i$, respectively. The latter can be solved by using the fact that $\mathrm{vec}(\check{C}_i S \check{C}_i^T) = (\check{C}_i \otimes \check{C}_i)\mathrm{vec}(S)$ which implies that $\mathrm{vec}(S) = (I - \check{C}_i \otimes \check{C}_i)^{-1}\mathrm{vec}(\check{V}_i)$. Here $\mathrm{vec}(S)$ is the vector obtained by stacking all the colums of S on top of each other and $\otimes$ is the Kronecker or tensor product.

3.3 The distribution $P(\underline{Q}|X(t_0))$

We now consider the distribution of the sequence of transitions between transversal section of different phases $P(\underline{Q}|X(t_0))$. We relabel $Q_{x_0}^{(1)}, \ldots, Q_{x_q}^{(1)}, \ldots, Q_{x_0}^{(m)}, \ldots, Q_{x_q}^{(m)}$ as $Q_1, \ldots, Q_N$

where $N = m(q+1)$. To each $Q_{x_i}^{(k)}$ there is a corresponding time $t_i + (k-1)\tau$. We label these times in increasing order as $T_1, \ldots T_N$ so that T_n corresponds to Q_n. Let $T_0 = t_0$. With this notation if follows from the above that $P(Q_{i+1}|Q_i)$ is approximately MVN with mean $\check{C}_i Q_i$ and covariance $\check{V}_i$ where $\check{C}_i = \check{C}(T_i, T_{i+1})$ and $\check{V}_i = \check{V}(T_i, T_{i+1})$. Consequently, if Q_i has covariance matrix $\check{S}_i$ then Q_{i+1} has covariance matrix given by $S_{i+1} = \check{V}_i + \check{C}_i S_i \check{C}_i^T$. As explained in Section S4 of the SI, it follows that $P(\underline{Q}|Q_0)$ is MVN with mean

$$\mu = (\check{C}_0 Q_0, \ldots, \check{C}_{N-1} \cdots \check{C}_0 Q_0)$$

and covariance Σ where the precision matrix Σ^{-1} is a tridiagonal matrix whose only non-zero entries are the main diagonal with matrix entries $\Sigma_{i,i} = \check{C}_i^T \check{V}_i^{-1} \check{C}_i + \check{V}_{i-1}^{-1}$ if $1 \leq i < N$ and $\Sigma_{N,N} = \check{V}_{N-1}^{-1}$, the upper diagonal with entries $-\check{C}_i^T \check{V}_i^{-1}$ and the lower diagonal with entries $-\check{V}_i^{-1} \check{C}_i$, $i = 1, \ldots, N-1$.

3.4 Fisher Information

Fisher Information quantifies the information that an observable random variable carries about an unknown parameter θ. If $P(X, \theta)$ is a probability distribution depending on parameters θ the Fisher Information Matrix (FIM) $I = I_P$ has entries

$$I_{ij} = E\left[\frac{\partial \ell}{\partial \theta_i}\frac{\partial \ell}{\partial \theta_i}\right] = -E\left[\frac{\partial^2 \ell(\theta; X)}{\partial \theta_i \theta_j}\right] \tag{7}$$

where $\ell = \log P$, and θ_i and θ_j are the i-th and j-th components of the parameter θ. If P is MVN with mean and covariance $\mu = \mu(\theta)$ and $\Sigma = \Sigma(\theta)$ then

$$I_{ij} = \frac{\partial \mu}{\partial \theta_i}^T \Sigma^{-1} \frac{\partial \mu}{\partial \theta_j} + \frac{1}{2}\mathrm{tr}\left(\Sigma^{-1}\frac{\partial \Sigma}{\partial \theta_i}\Sigma^{-1}\frac{\partial \Sigma}{\partial \theta_j}\right).$$

The FIM measures the sensitivity of P to a change in parameters in the sense that

$$D_{KL}(P(\cdot, \theta + \delta\theta), P(\cdot, \theta)) = \frac{1}{2}\delta\theta^T I \delta\theta + \mathrm{O}(||\delta\theta||^3)$$

where D_{KL} is the Kullback-Leibler divergence. The significance of the FIM for sensitivity and experimental design follows from its role in (7) as an approximation to the Hessian of the log-likelihood function at a maximum. Assuming non-degeneracy, if θ^* is a parameter value of maximum likelihood there is a $s \times s$ orthogonal matrix U such that, in the new parameters $\theta' = U \cdot (\theta - \theta^*)$,

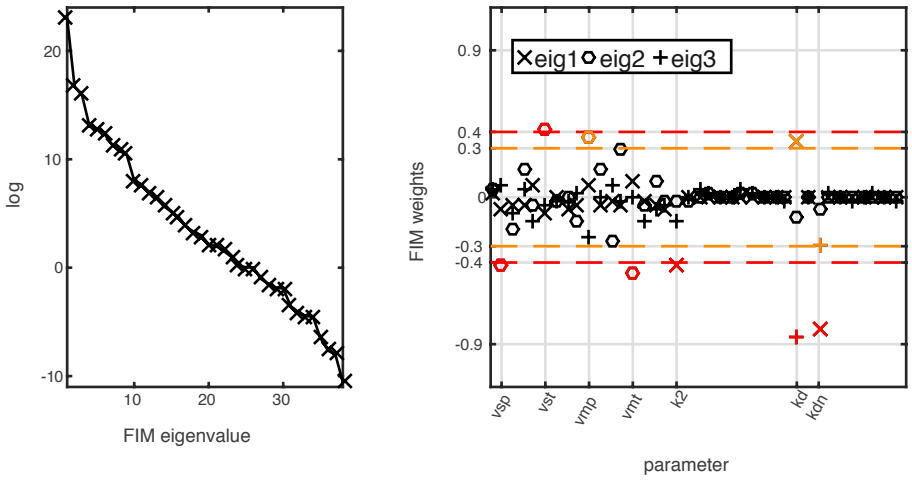
$$\ell(\theta) \approx \ell(\theta^*) - \sum_i \sigma_i^2 \, {\theta'_i}^2 .$$

for θ near θ^*. From these facts it follows that the σ_i^2 are the eigenvalues of the FIM and that the matrix U diagonalises it.

If we assume that the σ_i are ordered so that $\sigma_1^2 \geq \cdots \geq \sigma_s^2$ then it follows that near the maximum the likelihood is most sensitive when θ_1' is varied and least sensitive when θ_s' is. Moreover, σ_i is a measure of this sensitivity. Since $\theta_i' = \sum_j U_{ij}(\theta_j - \theta_j^*)$ we can regard $S_{ij} = \sigma_i U_{ij}$ as the contribution of the parameter θ_j to varying θ_i' and thus $S_j^2 = \sum_i S_{ij}^2$ can be regarded as a measure of the sensitivity of the system to θ_j. It is sometimes appropriate to normalise this and instead consider $\hat{S}_j = S_j^2 / \sum_i S_i^2$.

The theory of optimal experimental design is based on the idea of trying to make the σ_i decrease as slowly as possible so that the likelihood is as peaked as possible around the maximum, thus maximising the information content of the experimental sampling methods. Various criteria have been proposed such as D-optimality which maximises the determinant of the FIM and A-optimality that minimise the trace of the inverse of the FIM [5]. Diagonal elements of the inverse of FIM constitute a lower-bound for variance of any unbiased estimator of elements of θ (Cramer-Rao inequality). However, for the systems we consider the σ_i typically decrease very fast and there are many of them. Thus, in general, criteria based on a single number are less likely to be of less use than consideration of the set of σ_i as a whole.

Calculation of the FIM for stochastic systems using the LNA has been carried out in [11] but only for small systems and not for long times. It is notable that the pcLNA enables one to do this for large systems and large times. As an example, we analyse the stochastic behavior of the circadian clock based on the limit distribution $P(\underline{Q}|Q_0)$ when $\underline{Q} = Q_{x_0}^{(1)}, Q_{x_1}^{(1)}, Q_{x_0}^{(2)}, Q_{x_1}^{(2)}, \ldots, Q_{x_0}^{(m)}, Q_{x_1}^{(m)}$ where $x_0 = g(t_0)$ and $x_1 = g(t_1)$ are chosen so that t_0 is the time of the peak of *per* mRNA M_P and t_1 the peak of the nuclear complex of PER and TIM proteins C_N. We compute the Fisher Information of the latter distribution which has a closed form. As we can see in figure 9(a) the eigenvalues of the Fisher Information matrix decay exponentially, with a sharp decline followed by a slower decrease. This reveals that the influential directions in the parameter space of the system are much less than its total dimension. Furthermore, a few parameters appear to be most influential. The eigenvectors associated with the three largest eigenvalues of Fisher Information matrix have large entries only for the parameters *kdn* (PER-TIM complex nuclear degradation), *kd* (*per* mRNA linear degradation), *k2* (PER-TIM complex transportation to cytosol), *vst* (*tim* mRNA transcription), *vmt* (*tim* mRNA degradation), *vsp* (*per* mRNA transcription) and *vmp* (*per* mRNA degradation).



(a) FIM eigenvalues (b) FIM eigenvectors

Figure 9: Fisher Information Analysis of the circadian clock system. (a) The logarithm of the eigenvalues of the Fisher Information Matrix (FIM). (b) The entries/weights of the eigenvectors corresponding to the 3 largest eigenvalues of FIM. Large weights are marked with red color, moderate to large weights with orange color.

The exponential decrease of the eigenvalues is typical of tightly coupled deterministic systems [12, 13, 14, 15, 16, 17, 18], but has to our knowledge not been demonstrated before for stochastic systems. It has important consequences. For example, it tells us that only a few parameters can be estimated efficiently from time-series data unless the system is perturbed in some way to get complementary data and that there will be identifiability problems that can be analysed using the FIM. It also can be used to design experiments that will give data so that the FIM of the combined models (old and new experimental data) will alleviate the decline of the eigenvalues.

3.5 Calculating likelihoods via a pcLNA Kalman Filter

Although there is no elegant formula for $P(\underline{X}|X(t_0)) = P(X(t_1), \ldots, X(t_m)|X(t_0))$ similar to $P(\underline{Q}|X(t_0))$ above, we can efficiently calculate it. To do this we derive a Kalman Filter for the pcLNA that is a modification of the Kalman Filter associated with the LNA [10]. This can be used to compute the likelihood function $L(\theta; \underline{\hat{X}})$ of the system parameters θ with respect to observations of the species concentrations, $\underline{\hat{X}} = \left(\hat{X}(t_0), \hat{X}(t_1), \ldots, \hat{X}(t_N)\right)$ recorded at N times $t_0, t_1, \ldots, t_N$. This is slightly more general than just calculating $P(\underline{X}|X(t_0))$ because we allow for a measurement equation. The Kalman filter can also be used for forward prediction.

We assume the measurement equation,

$$\hat{X}(t) = BX(t) + \epsilon, \tag{8}$$

relating the observations $\hat{X}(t)$ to the state variables, $X(t)$. Here B is a transformation matrix (often simply removing unobserved species or introducing unknown scalings) and $\epsilon = (\epsilon_1, \ldots, \epsilon_n) \sim MVN(0, \Sigma_\epsilon)$ the observational error. The pcLNA likelihood can be decomposed as

$$L\left(\theta; \underline{\hat{X}}\right) = P\left(\hat{X}(t_0); \theta\right) \prod_{i=1}^{n} P\left(\hat{X}(t_i) | \hat{X}(t_{i-1}); \theta\right)$$

Let $m(t)$ and $S(t)$ be the mean and covariance matrix of the stochastic process $\xi(t) = \sqrt{\Omega}(X(t) - g(t))$ under the LNA. Then

$$\mu(t) = g(t) + \Omega^{-1/2} m(t), \quad \Sigma(t) = \Omega^{-1} S(t) \tag{9}$$

are the mean and covariance matrix of $X(t)$ and

$$\hat{\mu}(t) = B\mu(t), \quad \hat{\Sigma}(t) = B\Sigma(t)B^T + \Sigma_\epsilon \tag{10}$$

the mean and covariance of $\hat{X}(t)$. Denote $\hat{e}(t) = \hat{X}(t) - \hat{\mu}(t)$ the prediction error at time t. The posterior estimates[2] of $(\mu(t), \Sigma(t))$ derived using Bayes rule are

$$\mu^*(t) = \mu(t) + \Sigma(t)B^T\hat{\Sigma}(t)^{-1}\hat{e}(t), \quad \Sigma^*(t) = \Sigma(t) - \Sigma(t)B^T\hat{\Sigma}(t)^{-1}B\Sigma(t). \tag{11}$$

The pcLNA Kalman Filter algorithm uses the following recursive algorithm for computing the terms in $L\left(\theta; \underline{\hat{X}}\right)$.

1. Input $\left(\hat{X}(t_0), \hat{X}(t_1), \ldots, \hat{X}(t_N)\right)$, $\mu(t_0)$, $\Sigma(t_0)$, B and Σ_ϵ.
2. Compute $P\left(\hat{X}(t_0); \theta\right)$ from $MVN(\hat{\mu}(t_0), \hat{\Sigma}(t_0))$ using equation(10).
3. **For iteration** $i = 1, 2, \ldots$
   (a) Set $(X(t_{i-1})|\hat{X}(t_{i-1})) \sim MVN(\mu^*(t_{i-1}), \Sigma^*(t_{i-1}))$ using equation (11);
   (b) Derive $g(t'_{i-1}) = G(\mu^*(t_{i-1}))$ and $(\kappa(t'_{i-1})|\hat{X}(t_{i-1})) \sim MVN(m^*(t'_{i-1}), S^*(t'_{i-1}))$ using equation (12);
   (c) Set $(X(t_i)|\hat{X}(t_{i-1})) \sim MVN(\mu(t_i), \Sigma(t_i))$ using equations (13) and (14);
   (d) Compute $P\left(\hat{X}(t_i)|\hat{X}(t_{i-1}); \theta\right) \sim MVN(\hat{\mu}(t_i), \hat{\Sigma}(t_i))$.

In the for loop, the posterior mean and covariance matrix of the corrected stochastic process κ are

$$m^*(t'_{i-1}) = \Omega^{1/2}(\mu^*(t_{i-1}) - g(t'_{i-1})), \; S^*(t'_{i-1}) = \Omega E_2(t'_{i-1}) \left(\Sigma^*_{22}(t'_{i-1}) - \frac{\Sigma^*_{21}(t'_{i-1})\Sigma^*_{12}(t'_{i-1})}{\Sigma^*_{11}(t'_{i-1})}\right) E_2(t'_{i-1})^T \tag{12}$$

where the matrix $E_2(t) = [e_2(t) \ldots e_n(t)]$, the projections (in cartesian coordinates)

$$\Sigma^*_{11}(t) = e_1(t)^T \Sigma^*(t) e_1(t), \quad \Sigma^*_{12}(t) = e_1(t)^T \Sigma^*(t) E_2(t),$$

[2] conditional on the observed measurement at the current time

$$\Sigma^*_{21}(t) = \Sigma^*_{12}(t)^T, \quad \Sigma^*_{22}(t) = E_2(t)^T \Sigma^*(t) E_2(t)$$

and the prior[3] mean and covariance matrix of the species population X at time t_i,

$$\mu(t_i) = g(t_i) + \Omega^{-1/2} m(t_i), \quad \Sigma(t_i) = \Omega^{-1} S(t_i), \tag{13}$$

with

$$m(t_i) = C(t'_{i-1}, t_i) m^*(t'_{i-1}), \quad S(t_i) = V(t'_{i-1}, t_i) \tag{14}$$

the prior mean and covariance of the stochastic process ξ.

Equation (12) is derived by restricting the posterior distribution of the noise process on the transversal section $\mathcal{S}_{g(t'_{i-1})}$ normal to $e_1(t'_{i-1})$ using the Schur complement. If this correction is omitted, step 3(b) of the above algorithm is replaced by the standard LNA step,

$$(\xi(t_{i-1}) | \hat{X}(t_{i-1})) \sim N(m^*(t_{i-1}), S^*(t_{i-1}))$$

where

$$m^*(t_{i-1}) = \sqrt{\Omega}(\mu^*(t_{i-1}) - g(t_{i-1})), \quad S^*(t_{i-1}) = \Omega \Sigma^*(t_{i-1}).$$

4 Discussion

We present a comprehensive treatment of stochastic modelling for large stochastic oscillatory systems. Practical algorithms for fast long-term simulation and likelihood-based statistical inference are provided along with the essential tools for a more analytical study of such systems.

There is considerable scope for future work in various directions. We expect that these results can be extended to a broader class of systems including those that are chaotic in the $\Omega \to \infty$ limit. Our approach should provide the opportunity to develop new methodology for parameter estimation, likelihood-based inference and experimental design in such systems. Finally, there is currently much interest in information transfer and decision-making in signaling systems and our methods provide new tools with which to tackle problems in this area.

If system biologists are to reliably use complex stochastic models to provide robust understanding it is crucial that there are analytical tools to enable a rigorous assessment of the quality and selection of these models and their fit to current biological knowledge and data. Our aim in this paper is to contribute to that but the results should be of much broader interest.

Acknowledgments: This research was funded by the BBSRC Grant BB/K003097/1 (Systems Biology Analysis of Biological Timers and Inflammation). DAR was also supported by funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement nÂ° 305564.

[3]before the observation of the measurement at the current time

References

[1] D.T. Gillespie, *Exact stochastic simulation of coupled chemical reactions*, The journal of physical chemistry, 25 (1977), pp. 2340–2361.

[2] Daniel T. Gillespie, *Approximate accelerated stochastic simulation of chemically reacting systems.* The Journal of Chemical Physics. 115 (4) (2001) 00219606 p. 1716.

[3] Daniel T. Gillespie, *The Chemical Langevin Equation.* The Journal of Chemical Physics. 113:1 (2000) pp. 297â306.

[4] N.G. van Kampen, *Stochastic Processes in Physics and Chemistry. Third Edition*, Amsterdam: Elsevier (2007).

[5] T.G. Kurtz, *Limit Theorems for Sequences of Jump Markov Processes Approximating Ordinary Differential Processes*, Journal of Applied Probability, 8:2 (1971), pp. 344-356.

[6] R.P. Boland, T. Galla and A.J. McKane, *How limit cycles and quasi-cycles are related in systems with intrinsic noise*, Journal of Statistical Mechanics: Theory and Experiment, 9 (2008).

[7] D. Gonze, J. Halloy, J-C. Leloup and A. Goldbeter, *Stochastic model for circadian rhythms: effect of molecular noise on periodic and chaotic behaviour.*, Comptes Rendus Biologies 326:2 (2003), pp. 189â203.

[8] D.A. Potoyan and P.G. Wolynes, *On the dephasing of genetic oscillators.*, Proceedings of the National Academy of Sciences of the United States of America 111 (2014), pp. 2391-6.

[9] P. Hartman, *Ordinary differential equations*, New York: Wiley (1964).

[10] B. Finkenstädt, D.J. Woodcock, M. Komorowski, C.V. Harper, J.R.E. Davis, M.R.H. White and D.A. Rand, *Quantifying intrinsic and extrinsic noise in gene transcription using the linear noise approximation: An application to single cell data.*, The Annals of Applied Statistics 7 (2013), pp. 1960-1982.

[11] M. Komorowski, M.J. Costa, D.A. Rand and M.P.H. Stumpf, *Sensitivity, robustness, and identifiability in stochastic chemical kinetics models.*, Proceedings of the National Academy of Sciences of the United States of America 21 (2011), pp. 8645-8650.

[12] K. Brown and J. Sethna, *Statistical mechanical approaches to models with many poorly known parameters.*, Phys. Rev. E 68:2 (2003), 021904.

[13] K.S. Brown, C.C. Hill, G.A. Calero, C.R. Myers, K.H. Lee, J. Sethna, R.A. Cerione *The statistical mechanics of complex signaling networks: nerve growth factor signaling.*, Physical biology 1:3 (2004), pp. 184-195.

[14] R. N. Gutenkunst, et al. *Universally sloppy parameter sensitivities in systems biology models.* PLoS Computational Biology, 3:10 (2007), pp. 1871-1878.

[15] D.A. Rand, B.V. Shulgin, D. Salazar and A.J. Millar, *Design principles underlying circadian clocks.* Journal of The Royal Society Interface 1:1 (2004), pp. 119-130.

[16] D.A. Rand, B.V. Shulgin, D. Salazar and A.J. Millar, *Uncovering the design principles of circadian clocks: Mathematical analysis of flexibility and evolutionary goals.* Journal of Theoretical Biology 238:3 (2006), pp. 616-635.

[17] D.A. Rand, *Mapping the global sensitivity of cellular network dynamics: Sensitivity heat maps and a global summation law.* Journal of The Royal Society Interface 5 (2008), pp. S59-S69.

[18] J. Waterfall, F. Casey, R. Gutenkunst, K. Brown, C. Myers, P. Brouwer, V. Elser, J. Sethna, *Sloppy-Model Universality Class and the Vandermonde Matrix.* Phys. Rev. Lett. 97:15 (2006) 150601 1-4.