

Model reduction, coherence, and information transfer in stochastic biochemical systems

Juan David Marmolejo-Lozano ¹, Nikola Popović ^{2,3} and Ramon Grima ^{1,*}

¹*School of Biological Sciences, University of Edinburgh, Edinburgh, United Kingdom*

²*School of Mathematics, University of Edinburgh, Edinburgh, United Kingdom*

³*Maxwell Institute for Mathematical Sciences, University of Edinburgh, Edinburgh, United Kingdom*



(Received 7 November 2025; accepted 6 April 2026; published 29 April 2026)

Simplified stochastic models are widely used in the study of frequency-resolved noise propagation in biochemical reaction networks, a common measure being the coherence between random fluctuations in molecule number trajectories. Such models have also found widespread application in the quantification of how information is transmitted in reaction networks via the mutual information (MI) rate. A common assumption is that under timescale separation, estimates for the coherence and MI rate obtained from simplified (reduced) models closely approximate those in the underlying full models. Here, we challenge that assumption by showing that while reduced models can faithfully reproduce low-order statistics of molecular counts, they frequently incur substantial discrepancies in the coherence spectrum, especially at intermediate and high frequencies. These errors, in turn, lead to significant inaccuracies in the resulting estimates for the MI rates. We show that the observed discrepancies are due to the interplay between the structure of the underlying reaction networks, the specific model reduction method that is applied, and the asymptotic limits relating the full and reduced models. We illustrate our results in canonical models of enzyme catalysis and gene expression, highlighting practical implications for quantifying information flow in cells.

DOI: [10.1103/9h19-b6ht](https://doi.org/10.1103/9h19-b6ht)

I. INTRODUCTION

Realistic models of stochastic biochemical systems are typically difficult to study analytically, or even through stochastic simulation, due to the large numbers of interacting species and the myriad interactions involved. A much more common approach relies on the construction and study of reduced models that capture the essential dynamics of the underlying “full” systems, but that involve reduced numbers of interacting species or simplified reaction kinetics. Several methodologies have been devised to achieve model reduction, of which three are commonly used: (i) reduction through separation of timescales [1–6], (ii) reduction through abundance-scale separation [7–11], and (iii) reduction by generation of delay distributions [12–15]. Some of these methodologies are equivalent to minimizing the Kullback-Leibler divergence between the full and reduced models, defined on the space of Monte Carlo simulation trajectories [16]. Of the three types of model reduction, the most widely used are methods based on timescale separation; henceforth, we thus exclusively focus on that type.

Reduced models of timescale separation type can be formally proven to accurately estimate molecule number statistics, such as moments and distributions, in subregions

of parameter space. A prominent example is the slow-scale linear-noise approximation (ssLNA), the unique reduced form of the standard linear-noise approximation (LNA) of a complex biochemical system under timescale separation conditions, which can be rigorously derived using projection operator techniques [1]. Specifically, the ssLNA leads to an accurate reduced model when perturbations about the steady-state concentrations of a subset of the species—the slow species—decay on a much slower timescale than those about the steady-state concentrations of the remaining (fast) species; thus, the ssLNA is the stochastic equivalent of the deterministic quasi-steady-state approximation (QSSA). The large-molecule-number limits of several common stochastic reduction methods which assume that the set of governing reactions can be grouped into subsets of slow and fast reactions [17–19] can be shown to be special cases of the ssLNA [1]. The accuracy of the ssLNA in calculating the first and second moments of molecular counts, and their power spectra, has been verified by comparison with stochastic simulation using the stochastic simulation algorithm (SSA) [20] for a large variety of reaction networks; examples include reactions catalyzed by enzymes, genetic feedback loops, and biologically detailed models of stochastic gene expression [21,22]. An alternative heuristic type of model reduction, which is also commonly found in the literature [23–25], assumes the form of the probabilities for the effective reactions by comparison with the well-known effective macroscopic rate laws, such as in the assumption of Hill-type propensities in the chemical master equation (CME). The heuristic CME and the associated heuristic LNA, derived from such propensities, are accurate only in a subregion of the parameter space where timescale separation holds [21,26].

*Contact author: ramon.grima@ed.ac.uk

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

Reduced models have also been used to calculate the mutual information (MI) rate between molecule number trajectories of a pair of species in a biochemical reaction network [27]. Note that the MI rate is the time-averaged mutual information between a pair of stationary stochastic processes, which is different from instantaneous mutual information; for biochemical networks, the former is more relevant, as the latter cannot correctly quantify the amount of information transmitted per unit time [28–30]. A related concept is the coherence between molecule number trajectories of a pair of species, which measures how strongly the two trajectories are related at each frequency—within the linear-noise (or Gaussian) approximation, the coherence completely determines the MI rate via a simple integral formula [27,28,31,32]. For non-Gaussian fluctuations, as is typically the case, that formula, while not exact, is still useful as a lower bound for the true MI rate [33]. However, the relationship between the coherence and the MI rate predicted by full and reduced models has not been significantly explored, in contrast to the relationship between the moments of the molecular counts, as previously discussed. Since expressions for these quantities have been calculated exclusively for very simple (reduced) models of biological systems, it has, in effect, been implicitly assumed that the coherence and information transfer measures associated to the reduced models converge to those in the underlying full models in some kinetic parameter limit. For instance, the study of the MI rate between species X and Y interacting via the reaction scheme $\emptyset \rightarrow X \rightarrow \emptyset$, $X \rightarrow X + Y$, $Y \rightarrow \emptyset$ in [27] can be interpreted as a simple model for the transcription of mRNA X and its subsequent translation into protein Y ; however, that is clearly a reduced model, as many of the biologically important intermediate steps—such as gene state switching, nuclear export, and binding of mRNA to ribosomes—are omitted with the implicit understanding that these details are not important to the calculation of the MI rate, provided the intermediate steps are very fast.

The ssLNA lends itself perfectly to the task of understanding the relationship between the coherence and the MI rate in complex biochemical systems and their corresponding reduced models, as it allows for the straightforward calculation of these measures. Furthermore, the parameter limits relating the full and reduced models can be derived precisely, and the convergence of molecular moments is guaranteed in these limits. In this paper, we demonstrate that for many reaction networks, model reduction can produce markedly different coherence and MI rate estimates compared to the underlying full models, even when kinetic parameters are chosen so that the full and reduced models yield almost indistinguishable molecular moments at all times.

The paper is organized as follows. In Sec. II, we provide a concise summary of the general matrix formulation of the LNA, and we define the coherence and the MI rate in that context. In Sec. III, we summarize the two main types of timescale-separation-based model reduction: a popular heuristic LNA approach and the rigorous ssLNA. In Sec. IV, we prove that the coherence at zero frequency is exactly predicted by the ssLNA, but not by the heuristic LNA; we also show that the stoichiometry of the reaction networks described by the full and reduced models determines whether the coherence spectra of the two models obey similar

or different scaling laws at large frequencies. An important implication is that certain reduced models exhibit infinite MI rates, while those in the underlying full models are finite. In Sec. V, we illustrate our results in three reaction networks, including in models of enzyme catalysis and gene expression; furthermore, we identify special systems where model reduction perfectly preserves the coherence and information transfer measures. We conclude with a discussion of our results in Sec. VI.

II. CALCULATION OF COHERENCE AND MI RATE FROM THE LNA

Consider a general chemical reaction network involving N species that are reacting via R reactions,

$$\sum_{i=1}^N s_{ij} X_i \xrightarrow{k_j} \sum_{i=1}^N r_{ij} X_i, \quad j = 1, \dots, R. \quad (1)$$

Here, j is an index identifying the j th reaction; X_i denotes species i ; s_{ij} and r_{ij} are the integer stoichiometric coefficients; and k_j is the macroscopic rate constant of reaction j . By the Law of Mass Action, the rate at which the j th reaction occurs is given by $f_j = k_j \prod_{i=1}^N \phi_i^{s_{ij}}$, where ϕ_i is the concentration of species X_i . The net change in the number of molecules of X_i when reaction j occurs is $S_{ij} = r_{ij} - s_{ij}$.

It is well known that in the limit of large volumes, the dynamics of the mean molecule numbers are given by the deterministic rate equations,

$$\frac{d\vec{\phi}}{dt} = \underline{S}\vec{f}(\vec{\phi}, t), \quad (2)$$

where $\underline{S}$ is the stoichiometric matrix with elements S_{ij} , with $i = 1, \dots, N$ and $j = 1, \dots, R$; $\vec{f}$ is the rate function vector with elements f_i , where $i = 1, \dots, N$; and $\vec{\phi}$ is the mean concentration vector with elements ϕ_i , with $i = 1, \dots, N$.

We define the Jacobian matrix $\underline{J}$ of the rate equations with elements

$$J_{ij} = \frac{\partial}{\partial \phi_j} \sum_{r=1}^R S_{ir} f_r, \quad i, j = 1, \dots, N, \quad (3)$$

and the diffusion matrix $\underline{D}$ with elements

$$D_{ij} = \sum_{r=1}^R S_{ir} S_{jr} f_r, \quad i, j = 1, \dots, N, \quad (4)$$

which can be compactly written as $\underline{D} = \underline{S}\underline{F}\underline{S}^T$, where $\underline{F}$ is a diagonal matrix whose elements are f_i ($i = 1, \dots, N$).

Note that $D_{ii} > 0$, but that D_{ij} ($i \neq j$) is not sign definite and that it can also equal zero. Note, also, that the diffusion matrix D is positive semidefinite, which means that it is symmetric ($D_{ij} = D_{ji}$) and that its elements obey the inequality $D_{ij}^2 \leq D_{ii} D_{jj}$ [34,35]. In what follows, we will assume steady-state conditions: the time derivative in Eq. (2) is set to zero, which implies that the concentrations and the Jacobian and diffusion matrices are time-independent functions of the rate constants.

In the limit of large volume Ω and under the assumption that the system of rate equations in Eq. (2) has a

unique steady state, fluctuations in molecule numbers are described by an Ornstein-Uhlenbeck process that is commonly called the linear-noise approximation (LNA) [34,36]. Under that approximation, the variance-covariance matrix $\underline{C}$ of the concentration fluctuations is given by the solution of the Lyapunov equation,

$$\frac{d\underline{C}}{dt} = \underline{J}\underline{C} + \underline{C}\underline{J}^T + \frac{\underline{D}}{\Omega}; \quad (5)$$

this matrix formulation of the LNA is originally due to Elf and Ehrenberg [23]. The power spectrum matrix for the fluctuations in concentration is given by

$$\underline{P}(\omega) = \frac{1}{\Omega}(\underline{J} + i\underline{I}_N\omega)^{-1}\underline{D}(\underline{J}^T - i\underline{I}_N\omega)^{-1}, \quad (6)$$

where $\underline{I}_N$ is the identity matrix of rank N and ω is the frequency [21,37]. The (magnitude-squared) coherence ξ_{ij} between the time series of the fluctuations in the numbers of molecules of species X_i and X_j is given by

$$\xi_{ij}(\omega) = \frac{|P_{ij}(\omega)|^2}{P_{ii}(\omega)P_{jj}(\omega)}. \quad (7)$$

Since both factors in the denominator of Eq. (7) are positive and since the numerator is non-negative for any ω , it follows that the coherence is non-negative. By the Cauchy-Schwarz inequality, one can also show that $\xi_{ij}(\omega) \leq 1$.

Note that the coherence can be interpreted as a correlation analysis in the frequency domain since $\xi_{ij}(\omega)$ equals the square of the correlation coefficient of the frequency component ω of the time series of X_i and X_j . Hence, $\xi_{ij}(\omega) = 0$ when the fluctuations in X_i and X_j at frequency ω are uncorrelated, whereas $\xi_{ij}(\omega) = 1$ when they are perfectly correlated. That interpretation is due to the Wiener-Khinchin theorem: the variance is the integral of the power spectrum over ω , while the covariance is the integral of the cross spectrum over ω ,

$$\xi_{ij}(\omega) = \frac{|P_{ij}(\omega)\Delta\omega|^2}{(P_{ii}(\omega)\Delta\omega)(P_{jj}(\omega)\Delta\omega)} = \frac{C_{ij}(\omega)^2}{C_{ii}(\omega)C_{jj}(\omega)}, \quad (8)$$

where $C_{ij}(\omega)$ is the covariance of the two signals after they are filtered by passing through a bandpass filter centered at frequency ω and infinitesimally small bandwidth $\Delta\omega$; similarly, $C_{ii}(\omega)$ and $C_{jj}(\omega)$ are the variances of the signals X_i and X_j after filtering. In what follows, we shall refer to a plot of the coherence as a function of frequency as the coherence spectrum.

Since the fluctuations in the LNA are Gaussian distributed, it also follows that the MI rate between the two time series is given by

$$M_{ij} = -\frac{1}{4\pi \ln 2} \int_{-\infty}^{\infty} \ln[1 - \xi_{ij}(\omega)] d\omega. \quad (9)$$

Because of the factor of $\ln 2$, the MI rate is expressed in units of bits per unit time. Note that the formula for the MI rate in Eq. (9) stems from its formal definition as the limit of the time-averaged mutual information between corresponding segments of two stationary stochastic processes, as the segment length approaches infinity [27,32]. It can also be proven analytically that Eq. (9) provides a lower bound on the MI rate when the fluctuations are non-Gaussian [33], which might not be tight [38,39]. Since, in reality, molecular fluctuations in

biochemical systems are discrete, they are never truly Gaussian as assumed by the LNA; hence, we shall purely interpret Eq. (9) as an analytically tractable lower bound for the actual MI rate, which can be estimated, in some cases, by Monte Carlo approaches based on path weight sampling [38].

III. REDUCED LNA UNDER TIMESCALE SEPARATION CONDITIONS: HEURISTIC AND RIGOROUS APPROACHES

Consider the case where a subset of the N species interacting via the reactions in Eq. (1) are slow, while the remainder are fast. Let N_s and N_f be non-negative integers with $N_s + N_f = N$, where N_s and N_f are the numbers of slow and fast species, respectively. Specifically, fast species exhibit rapid relaxation of small deviations from their steady-state concentrations, while slow species relax on much longer timescales.

A. Heuristic model reduction

A heuristic reduction method is based on the deterministic QSSA: in the rate equations, the time derivatives of the concentrations of the fast species are set to zero, and expressions are derived for the concentrations of the fast species in terms of the slow species. These expressions are then substituted into the rate equations for the slow species to obtain a set of new rate equations that are specifically functions of the slow species only.

A common example of this type of reduction occurs in the modeling of the enzyme-catalyzed reaction network $S + E \rightleftharpoons C \rightarrow E + P$, where S is the substrate, E is the enzyme, C is the complex, and P is the product. Under the assumption that E and C are the fast species and S and P are the slow species, the QSSA allows for the rate equations to be reduced to those describing the simpler system $S \rightarrow P$ with an effective nonlinear rate proportional to $\phi_S/(K_M + \phi_S)$, where ϕ_S is the substrate concentration and K_M denotes the Michaelis-Menten constant.

For a general system that is reduced via the QSSA, to the reduced rate equations there corresponds a coarse-grained reaction network composed of $R^* < R$ effective reactions between the slow species,

$$\sum_{i=1}^{N_s} s_{ij}^* X_i^s \rightarrow \sum_{i=1}^{N_s} r_{ij}^* X_i^s, \quad j = 1, \dots, R^*. \quad (10)$$

Here, X_i^s denotes the i th slow species, and s_{ij}^* and r_{ij}^* are the integer stoichiometric coefficients. The rate at which reaction j occurs is some function $f_j^*(\phi_1^s, \dots, \phi_{N_s}^s)$, where ϕ_i^s is the concentration of the slow species X_i^s . Note that the function f_j^* is not generally consistent with the Law of Mass Action, an example being the Michaelis-Menten rate law discussed above, as it describes the rate of an effective reaction, not an elementary one. The net change in the number of molecules of X_i^s when reaction j occurs is $S_{ij}^* = r_{ij}^* - s_{ij}^*$.

The LNA of this effective reduced system of reactions is given by Eqs. (3)–(6), with S_{ij} and f_j replaced by S_{ij}^* and f_j^* , respectively. The coherence then follows directly from Eq. (7). The resulting method for constructing a reduced LNA is heuristic because (i) it is not rigorously derived from the

LNA of the full system under timescale separation conditions, and (ii) it assumes that the reduced microscopic kinetics is similar to the reduced deterministic and macroscopic kinetics. The heuristic reduction is one the methods we consider in this paper, as it has been widely applied, including in recent studies of the MI rate [39,40]. We note that despite its heuristic nature, this method does lead to accurate predictions in a subregion of the parameter space where timescale separation holds—specifically, it applies for parameter sets satisfying both the quasiequilibrium and quasi-steady-state approximations [21,36,41,42].

B. Rigorous model reduction

In [1,21], a rigorous adiabatic elimination method was developed for the LNA. Starting from the Fokker-Planck equation describing the LNA of the full system, the method, which is called the ssLNA, allows for the derivation of a simplified Fokker-Planck equation that describes the time evolution of the joint probability density function of the slow variables only. In what follows, we briefly summarize the method—we shall refer to the system of fast and slow interacting species as the full model, while we denote the reduced system of slow species interacting via effective reactions as the reduced model.

First, we relabel the species in Eq. (1) so that $X_1, \dots, X_{N_s}$ are the slow ones and $X_{N_s+1}, \dots, X_N$ are the fast ones. We partition the $N \times N$ square matrix $\underline{J}$ whose elements are given by Eq. (3) and the $N \times R$ matrix $\underline{S}$ as

$$\underline{J} = \begin{pmatrix} \underline{J}_{ss} & \underline{J}_{sf} \\ \underline{J}_{fs} & \underline{J}_{ff} \end{pmatrix} \quad \text{and} \quad \underline{S} = \begin{pmatrix} \underline{S}_s \\ \underline{S}_f \end{pmatrix},$$

respectively. Here, the dimensions of the rectangular (or square) submatrices $\underline{J}_{ss}$, $\underline{J}_{sf}$, $\underline{J}_{fs}$, and $\underline{J}_{ff}$ are $N_s \times N_s$, $N_s \times N_f$, $N_f \times N_s$, and $N_f \times N_f$, respectively, while $\underline{S}_s$ and $\underline{S}_f$ have dimensions $N_s \times R$ and $N_f \times R$, respectively.

The effective Jacobian and diffusion matrices, valid under timescale separation conditions, are given by

$$\underline{J}_S = \underline{J}_{ss} - \underline{J}_{sf} \underline{J}_{ff}^{-1} \underline{J}_{fs}, \quad (11)$$

$$\underline{D}_S = (\underline{A} - \underline{B})(\underline{A} - \underline{B})^T, \quad (12)$$

where

$$\underline{A} = \underline{S}_s \sqrt{\underline{F}} \quad \text{and} \quad \underline{B} = \underline{J}_{sf} \underline{J}_{ff}^{-1} \underline{S}_f \sqrt{\underline{F}}.$$

Note that $\underline{J}_S$ is the Schur complement of the submatrix $\underline{J}_{ff}$ of the matrix $\underline{J}$, which is defined provided $\underline{J}_{ff}$ is invertible. Note, also, that $\underline{F}$ is the $R \times R$ diagonal matrix with elements of $\tilde{f}$, i.e., of the rate function vector in the full model.

Next, Eq. (12) can be written in the form $\underline{D} = \underline{S}' \underline{F} \underline{S}'^T$, where $\underline{S}'$ is an effective stoichiometric matrix of the reduced system given by

$$\underline{S}' = \underline{S}_s - \underline{J}_{sf} \underline{J}_{ff}^{-1} \underline{S}_f. \quad (13)$$

Notably, the elements of $\underline{S}'$ are generally not integers, which reflects the fact that reduction of a full model described by the CME does not typically result in a model that can be described by a different CME, as the reduced dynamics may be non-Markovian [43]—a Langevin formulation of the reduced ssLNA dynamics always exists, however.

In analogy to Eqs. (5) and (6), the equations for the variance-covariance matrix $\underline{C}_S$ of concentration fluctuations of the slow species obtained from the ssLNA read

$$\frac{d\underline{C}_S}{dt} = \underline{J}_S \underline{C}_S + \underline{C}_S \underline{J}_S^T + \frac{\underline{D}_S}{\Omega}, \quad (14)$$

while the power spectrum matrix $\underline{P}_S$ for the fluctuations in concentration is given by

$$\underline{P}_S(\omega) = \frac{1}{\Omega} (\underline{J}_S + i \underline{I}_{N_s} \omega)^{-1} \underline{D}_S (\underline{J}_S^T - i \underline{I}_{N_s} \omega)^{-1}, \quad (15)$$

where $\underline{I}_{N_s}$ is the identity matrix of rank N_s and ω is the frequency. The coherence between any two slow species can then be calculated from Eq. (7), with $P_{ij}(\omega)$ replaced by $[\underline{P}_S]_{ij}(\omega)$.

We note that, generally, the ssLNA does not correspond to the heuristic LNA discussed earlier. The heuristic approach has been shown to be equivalent to the implicit assumption that under timescale separation conditions, reversible elementary reactions involving fast species do not contribute to the intrinsic noise in the concentrations of slow species—an assumption that is generally incorrect, with the exception of parameter sets satisfying both the quasi-equilibrium and quasi-steady-state approximations [21,36]. It can be shown that heuristic model reduction and the ssLNA always result in the same set of reduced deterministic rate equations and, hence, in the same Jacobian $\underline{J}_S$; however, they differ in their diffusion matrices.

IV. GENERAL RELATIONSHIPS BETWEEN COHERENCE SPECTRA AND MI RATES OF THE FULL AND REDUCED MODELS

A. Coherence at zero frequency

In this section, we investigate the relationship between the coherences predicted by the full and reduced models for any pair of slow species at $\omega = 0$.

From Eqs. (6) and (15), it immediately follows that the spectra at $\omega = 0$ of the full LNA and the ssLNA are given by

$$P_{ij}(0) = [\underline{J}^{-1} \underline{D} \underline{J}^{-T}]_{ij} \quad \text{and} \quad [\underline{P}_S]_{ij}(0) = [\underline{J}_S^{-1} \underline{D}_S \underline{J}_S^{-T}]_{ij},$$

respectively, where $i, j \in \{1, \dots, N_s\}$. Note that for the full model, we merely consider the upper-left $N_s \times N_s$ block of $\underline{J}^{-1} \underline{D} \underline{J}^{-T}$, which we denote as $\underline{P}_{ss}$, the reason being that we are only concerned with the spectra of the slow species.

In Appendix A, we prove that

$$\underline{P}_{ss} = \underline{J}_S^{-1} \underline{D}_S \underline{J}_S^{-T},$$

from which it immediately follows that the power spectra and the cross spectra of the slow species, as predicted by the full LNA and the ssLNA, coincide at zero frequency,

$$P_{ij}(0) = [\underline{P}_S]_{ij}(0), \quad (16)$$

where $i, j \in \{1, \dots, N_s\}$.

Finally, by Eq. (7), it follows that *the coherences in the full model and the reduced model obtained from the ssLNA are precisely equal at $\omega = 0$* . Since the heuristic model reduction leads to an LNA with the same Jacobian, but a different diffusion matrix than the ssLNA, it follows that the coherence

of the heuristic reduced model cannot generally be guaranteed to be correct at $\omega = 0$.

B. Coherence spectrum in the large-frequency limit

We start by considering the leading-order asymptotics of the spectra P_{ii} and cross spectra P_{ij} of the full model in the limit of large ω .

Let $\underline{\Lambda} = \underline{J} + i\omega\underline{L}_N$ and $\underline{\Delta} = \underline{J}^T - i\omega\underline{L}_N$. Note that these matrices are key components in the expression for the power spectrum given by Eq. (6), with

$$\begin{aligned}\underline{\Lambda}^{-1} &= (i\omega)^{-1} \left(\underline{L}_N - i \frac{\underline{J}}{\omega} \right)^{-1} \\ &= (i\omega)^{-1} \sum_{k=0}^{\infty} i^k \frac{\underline{J}^k}{\omega^k} = \sum_{k=0}^{\infty} i^{k-1} \frac{\underline{J}^k}{\omega^{k+1}},\end{aligned}\quad (17)$$

where we used the Neumann series representation, which is valid in the limit of large ω , in the last two steps. Similarly, we can write

$$\begin{aligned}\underline{\Delta}^{-1} &= (-i\omega)^{-1} \left(\underline{L}_N + i \frac{\underline{J}^T}{\omega} \right)^{-1} \\ &= (-i\omega)^{-1} \sum_{k=0}^{\infty} (-i)^k \frac{(\underline{J}^T)^k}{\omega^k} = \sum_{k=0}^{\infty} (-i)^{k-1} \frac{(\underline{J}^T)^k}{\omega^{k+1}}.\end{aligned}\quad (18)$$

Now, by Eq. (6), we have

$$\underline{P} = \Omega^{-1} \underline{\Lambda}^{-1} \underline{D} \underline{\Delta}^{-1}, \quad (19)$$

$$= \Omega^{-1} \sum_{k=0}^{\infty} \sum_{m=0}^{\infty} i^{k-1} (-i)^{m-1} \frac{\underline{J}^k}{\omega^{k+1}} \underline{D} \frac{(\underline{J}^T)^m}{\omega^{m+1}}, \quad (20)$$

$$= \Omega^{-1} \sum_{k=0}^{\infty} \sum_{m=0}^{\infty} \frac{i^{k+3m-4}}{\omega^{k+m+2}} \underline{J}^k \underline{D} (\underline{J}^T)^m. \quad (21)$$

In the limit of large ω , the leading-order series expansion (in ω^{-1}) for $\underline{P}$ is hence given by

$$\underline{P} = \frac{\Omega^{-1} \underline{D}}{\omega^2} - \frac{i\Omega^{-1}}{\omega^3} (\underline{D} \underline{J}^T - \underline{J} \underline{D}) + O(\omega^{-4}), \quad (22)$$

which is obtained for $(k, m) = (0, 0)$, $(k, m) = (0, 1)$, and $(k, m) = (1, 0)$ in Eq. (21). For a general system of interacting chemical species, $\underline{D} \underline{J}^T \neq \underline{J} \underline{D}$, which implies that the second term in Eq. (22) is nonzero. For a system of noninteracting species, it is straightforward to show from the definitions of the Jacobian and diffusion matrices in Eqs. (3) and (4) that both $\underline{J}$ and $\underline{D}$ are diagonal matrices and, hence, that $\underline{D} \underline{J}^T = \underline{J} \underline{D}$.

Using Eqs. (7), (9), and (22), we find that if $D_{ij} \neq 0$, then

$$\xi_{ij}(\omega) \rightarrow \frac{D_{ij}^2}{D_{ii} D_{jj}} \quad \text{as } \omega \rightarrow \infty, \quad (23)$$

and, hence, that $M_{ij} \rightarrow \infty$. Recall that because the diffusion matrix $\underline{D}$ is symmetric and positive semidefinite [34], $D_{ij}^2 \leq D_{ii} D_{jj}$ holds [35], which implies $0 \leq \xi_{ij}(\omega) \leq 1$.

By contrast, if $D_{ij} = 0$, then we find

$$\xi_{ij}(\omega) \sim \frac{([\underline{D} \underline{J}^T]_{ij} - [\underline{J} \underline{D}]_{ij})^2}{D_{ii} D_{jj}} \frac{1}{\omega^2} \quad \text{as } \omega \rightarrow \infty, \quad (24)$$

which implies that

$$M_{ij} \geq 0 \quad \text{is finite.} \quad (25)$$

Note that although $\underline{D} \underline{J}^T \neq \underline{J} \underline{D}$, in general, for a system of interacting species, as previously mentioned, it is possible for specific parameter values to yield $[\underline{D} \underline{J}^T]_{ij} = [\underline{J} \underline{D}]_{ij}$ —in that case, $\xi_{ij}(\omega) \propto \omega^{-4}$ as $\omega \rightarrow \infty$; see below for an example. A proof of Eq. (25) can be found in Appendix B.

Next, we give a physical interpretation of these results. The case where $D_{ij} = \sum_{p=1}^R S_{ip} S_{jp} f_p(\vec{\phi}) \neq 0$ occurs when *in the set of reactions defining the biochemical system, there exists at least one reaction in which the numbers of molecules of species X_i and X_j change simultaneously when the reaction occurs.* For such reaction networks, the LNA predicts an infinite MI rate between species X_i and X_j , which reflects the fact that the change in the numbers of molecules of X_i and X_j is *perfectly coupled* in that case; for example, the number of molecules of species X_i decreases by -1 when the reaction $X_i \rightarrow X_j$ occurs, while the number of molecules of species X_j increases by $+1$.

By contrast, *when the species X_i and X_j are either not involved in a common reaction or are involved in one or more common reactions such that the net change in the number of molecules of one of the species is zero*, then $D_{ij} = 0$. The reason is that in these cases, for any reaction p , either S_{ip} or S_{jp} will be zero. Then, the LNA predicts a finite MI rate, which reflects the imperfect coupling of the fluctuations in the numbers of molecules of species X_i and X_j . We note that these conclusions hold for any pair of species—slow or fast, or a combination thereof—described by the LNA of the full model.

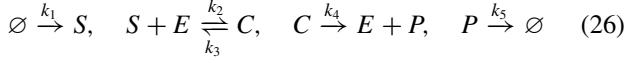
Starting from Eq. (15), similar results can be derived for the power spectra and cross spectra of slow species from the ssLNA: Eqs. (23) and (24) hold, with $\underline{D}$ and $\underline{J}$ replaced by $\underline{D}_S$ and $\underline{J}_S$, respectively. In that case, $[\underline{D}_S]_{ij} = \sum_{p=1}^R S'_{ip} S'_{jp} f_p(\vec{\phi})$, where $\underline{S}'$ is the effective stoichiometric matrix of the reduced system; see Eq. (13). Under timescale separation, the condition $[\underline{D}_S]_{ij} \neq 0$ holds when there exists at least one reaction in which the numbers of molecules of the slow species X_i and X_j effectively change simultaneously when the reaction occurs.

Conclusion. The coherence spectra of two slow species X_i and X_j in the full model and the reduced model obtained from the ssLNA obey different scaling laws for large ω if, in only one of these two models, there exists at least one reaction in which the numbers of molecules of species X_i and X_j change simultaneously when the reaction occurs—one of the models then predicts an infinite MI rate, while the MI rate predicted by the other is finite. If both the full model and the ssLNA obey the aforementioned reaction constraint or if neither obeys it, then the coherence spectra satisfy the same scaling laws for large ω ; nevertheless, these might still not agree exactly if the corresponding prefactors differ. In that case, both models will predict either an infinite MI rate or a finite MI rate. In summary, model reduction via the ssLNA generally does not preserve the coherence spectrum, except for very small frequencies. The same general conclusions hold for a comparison of the coherence spectra of the slow species in the full model and the heuristic reduced model, except that there is then also lack of agreement at $\omega = 0$.

V. APPLICATIONS

A. Michaelis-Menten model of enzyme catalysis with substrate synthesis

Consider the set of reactions



describing a conventional enzyme-catalyzed reaction network in which substrate S is produced, binds enzyme E to form complex C , and then disassociates into the original unbound enzyme form E and product P . The latter is subsequently degraded. Note that the total concentration ϕ_T of enzyme molecules in unbound and bound form is a time-independent constant since the enzyme only promotes the reaction but is itself neither produced nor destroyed. We do not explicitly model substrate decay as the corresponding rate, which, for most proteins, is typically due to dilution, is much smaller than the effective removal rate of the substrate by the enzyme reaction.

In the large-system-size limit, the mean concentrations of all chemical species in Eq. (26) are found from the deterministic rate equations in Eq. (2); solution of these equations at the steady state leads to the mean concentrations for the substrate, enzyme, and product,

$$\phi_S = \frac{K_M x}{1 - x}, \quad \phi_E = \phi_T - \frac{k_1}{k_4}, \quad \text{and} \quad \phi_P = \frac{k_1}{k_5}, \quad (27)$$

where $K_M = (k_3 + k_4)/k_2$ is the Michaelis-Menten constant and $x = k_1/(k_4\phi_T)$. Note that steady-state conditions only ensue provided that the substrate production rate is lower than the maximum rate at which the enzyme can catalyze the substrate into the product, i.e., for $x < 1$. The free enzyme and complex concentrations are related by a conservation law, $\phi_E + \phi_C = \phi_T$, where ϕ_T is the total (constant) enzyme concentration.

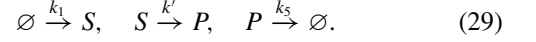
From the rate equations, it can be shown (Appendix C) that the timescales associated with the various species are given by

$$\tau_S = \frac{1}{k_2\phi_T} \frac{1}{1 - x}, \quad \tau_P = \frac{1}{k_5}, \quad \text{and} \quad \tau_C = \frac{1 - x}{k_3 + k_4}. \quad (28)$$

We shall consider the case where the substrate and product are the slow species and the complex (and free enzyme) are the fast ones. The associated timescale separation condition is $\tau_S, \tau_P \gg \tau_C$. Note that the timescale τ_X of species X ($X = S, P, C$) is defined as the time it takes for an initial perturbation about the steady-state concentration of X to decay, while the concentrations of the remaining species are assumed to remain at the steady state. This definition is heuristic, as it neglects the coupling between species in Eq. (2), but allows for a simple identification of timescales due to the fact that individual perturbations decay exponentially with time. If simultaneous perturbations about the steady-state concentrations of all species are considered, the total perturbation in each species is typically given by a weighted sum of exponentials from which it is difficult to extract timescales analytically, even though a numerical identification is achievable. In Appendix C, we show that the ratios of the timescales obtained from the two approaches—individual versus simultaneous perturbation—are consistent with each other. Hence, for the remainder of the

paper, we shall exclusively rely on timescales derived from our heuristic approach.

Using the timescale separation condition $\tau_S, \tau_P \gg \tau_C$, one can reduce the rate equations for the full model to those of a simpler system of reactions by means of the QSSA,



Here, the enzyme and complex dynamics are not explicitly modeled; rather, they are implicitly considered through a judicious choice of the effective rate constant k' . The rate of the effective reaction $S \rightarrow P$ is of the Michaelis-Menten form, $k'\phi_S = k_4\phi_T\phi_S/(K_M + \phi_S)$.

We refer to the LNA corresponding to the full circuit in Eq. (26) as the full LNA. Similarly, the heuristic LNA is the LNA corresponding to the reduced circuit in Eq. (29). A second, more rigorous way to obtain a reduced LNA when the complex is fast relies on the ssLNA. The corresponding calculations can be found in Appendix C. Therein, we show that while all three approaches predict the same means for all species, the second moments differ. In particular, the ssLNA agrees with the full LNA whenever the timescale separation condition $\tau_S, \tau_P \gg \tau_C$ is satisfied. However, the heuristic LNA requires an additional constraint, $k_3 \gg k_4$, which we call the “infrequent catalysis constraint,” as it implies that when complex C is formed, it is much more probable that it decays to E and S rather than to E and P . Similar results have previously been reported for simpler metabolic systems [41].

It can also be shown that the coherences between the substrate and product fluctuations obtained from the LNA of the full model, the ssLNA, and the heuristic LNA are given by

$$\xi_{S,P}^{\text{full}}(\omega) = \frac{c_0 + c_1\omega^2 + c_2\omega^4}{c_3 + c_4\omega^2 + c_5\omega^4 + c_6\omega^6}, \quad (30)$$

$$\xi_{S,P}^{\text{ssLNA}}(\omega) = \frac{d_0 + d_1\omega^2}{d_2 + d_3\omega^2}, \quad \text{and} \quad (31)$$

$$\xi_{S,P}^{\text{heur}}(\omega) = \frac{1}{4}, \quad (32)$$

respectively. Here, c_i ($i = 0, 1, 3, 4, 5, 6$) and d_i ($i = 0, \dots, 3$) are positive constants, while c_2 is a non-negative constant, all of which are functions of the reaction rates k_i and the total enzyme concentration ϕ_T ; for precise formulas, see Appendix C.

Our main question now is: given a choice of parameters such that the ssLNA and the heuristic LNA accurately predict the statistics of the substrate and product, does it follow that the corresponding predictions for the coherence spectrum are accurate?

We test our hypothesis as follows. We fix the parameter set to $k_1 = 1$, $k_2 = 0.1$, $k_3 = 2$, $k_4 = 20$, $k_5 = 1$, and $\phi_T = 1$, which enforces timescale separation conditions since $\tau_S = 10.53$, $\tau_P = 1$, and $\tau_C = 0.043$. The steady-state predictions from the ssLNA and the full LNA for the variances and covariances of the concentration fluctuations of the substrate and product are very similar; to six significant figures, these read $C_S^{\text{full}} = 11.6617$, $C_S^{\text{ssLNA}} = 11.6620$; $C_P^{\text{full}} = 0.961920$, $C_P^{\text{ssLNA}} = 0.960092$; and $C_{S,P}^{\text{full}} = 0.0399845$, $C_{S,P}^{\text{ssLNA}} = 0.0399076$, which implies that the relative errors are less than 0.2%. However, the ssLNA yields very poor estimates for the coherence spectrum, as seen in

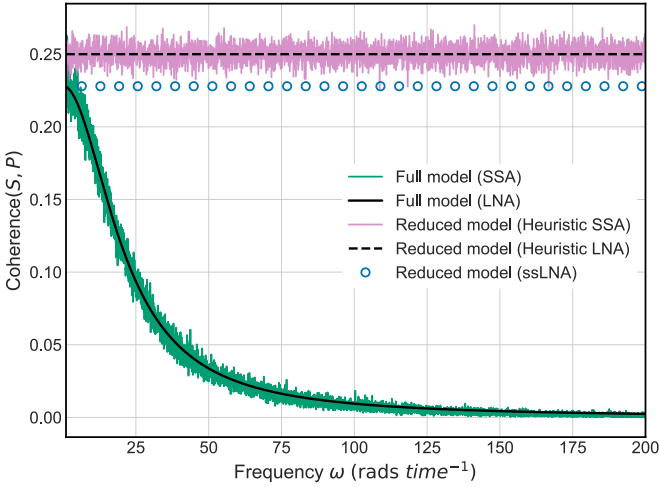


FIG. 1. Comparison of the coherence spectrum of the full and reduced models for the enzyme-catalyzed reaction. The parameter set is chosen so that timescale separation is enforced; see the main text for the corresponding parameter values. The solid and dashed black lines represent the predictions of the full LNA [Eq. (30)] and the heuristic LNA [Eq. (32)], respectively; the prediction of the ssLNA [Eq. (31)] is indicated by open blue circles. Estimates for the coherence spectra computed from SSA simulations of the full model [Eq. (26)] and the heuristic reduced model [Eq. (29)], processed with Welch's method, are shown in green and pink, respectively. The SSA output was sampled every 0.01 time units for a total duration of $T_{\max} = 10^6$ units; the Nyquist frequency is $\omega_{\max} \approx 314$.

Fig. 1; in fact, its prediction is similar to the coherence spectrum predicted by the heuristic LNA. We also estimate the coherence spectra directly by applying Welch's method [44], as implemented in the `scipy.signal` package, to trajectories generated by the SSA [20] of the full and reduced reaction networks in Eqs. (26) and (29), respectively. For details on the implementation of the SSA for the reduced model, which we henceforth refer to as the heuristic SSA, see [19,45]. Since the coherence spectra from SSA simulations agree very well with the predictions of the full LNA and the heuristic LNA, it follows that the differences between these two types of LNA are not related to the continuum approximation that is inherent in the LNA.

In particular, we note that the coherence spectrum of the full LNA always decays to zero, while both the ssLNA and the heuristic LNA predict a constant coherence in the large- ω limit—this discrepancy is evident from Eqs. (30)–(32), in fact, which show that the coherence decays as ω^{-2} except for the special case when $x = 1/2$, where it decays as ω^{-4} ; see Appendix C. The resulting differences in the tails of the coherence spectra agree with the general derivation in Sec. IV B—they stem from the fact that there are no reactions in Eq. (26) which cause a simultaneous change in the net numbers of substrate and product molecules, whereas such a reaction is present in the reduced model, given by Eq. (29). The different scaling laws for large ω predicted by the full and reduced models mirror those in the cross spectrum; see Appendix C. By contrast, note that the differences between the predicted coherences are smallest for low ω , which is also

borne out by theory,

$$\xi_{S,P}^{\text{full}}(0) = \xi_{S,P}^{\text{ssLNA}}(0) = \frac{1}{4} \frac{1}{1 + Kx(x-1)} \neq \xi_{S,P}^{\text{heur}}(0), \quad (33)$$

where $K = k_4/(k_3 + k_4)$. In the example of Fig. 1, $x = 1/20$ is very small; therefore, all three predicted coherences are in good agreement when $\omega = 0$. The exactness of the ssLNA prediction at $\omega = 0$ is a special case of the general derivation in Sec. IV A.

Finally, we seek to understand the underlying reason for the inaccuracy of the coherences predicted by the reduced LNAs. It can be shown that the limit of large k_3 in Eqs. (30) and (31) at fixed ω results in convergence of the second-order moments of the full LNA and the heuristic LNA: $C_S^{\text{full}}/C_S^{\text{heur}} \rightarrow 1$, $C_P^{\text{full}}/C_P^{\text{heur}} \rightarrow 1$, and $C_{S,P}^{\text{full}} - C_{S,P}^{\text{heur}} \rightarrow 0$. The same is found for the ssLNA. In addition, the coherences predicted by the three LNAs also converge to the same value,

$$\xi_{S,P}^{\text{full}}(\omega), \xi_{S,P}^{\text{ssLNA}}(\omega) \rightarrow \frac{1}{4} = \xi_{S,P}^{\text{heur}}(\omega). \quad (34)$$

In contrast, the full LNA, given by Eq. (30), predicts that if $x \neq 1/2$, then the coherence spectrum decays as

$$\xi_{S,P}^{\text{full}}(\omega) \sim \frac{1}{4} \frac{(2x-1)^2}{(x-1)^2} \frac{k_3 k_4}{\omega^2} \quad (35)$$

for large frequencies in the timescale separation limit of large k_3 , whereas for $x = 1/2$, the spectrum decays as

$$\xi_{S,P}^{\text{full}}(\omega) \sim \frac{4k_3^3 k_4}{\omega^4}. \quad (36)$$

The discrepancies between the expressions in Eqs. (34)–(36) are due to the timescale separation limit and the large- ω limit *not commuting*. In Eq. (34), we took the limit of large k_3 at fixed ω , while in Eqs. (35) and (36), we first took the limit of large ω in Eq. (30), followed by the large- k_3 limit. We note that the limit taken at fixed ω implicitly assumes that the largest (Nyquist) frequency $\omega_{\max}$ at which the coherence spectrum can be estimated is always much smaller than the value of k_3 . By contrast, in stochastic simulations, for any finitely large value of k_3 , it is possible to measure the coherence spectrum for any ω since $\omega_{\max}$ can be made arbitrarily large by increasing the sampling rate of the time series of substrate and product molecule numbers output by the SSA. In a real-world scenario, the Nyquist frequency is set by the sampling rate of the device used to make molecule number measurements; if $k_3 \gg \omega_{\max}$, then the large-frequency tail of the measured coherence spectrum would be consistent with Eq. (34) in that case, whereas if $k_3 \ll \omega_{\max}$, then the tails would be described by Eqs. (35) and (36). In other words, the full LNA can accurately describe both cases, while the reduced LNAs can only approximate the low-Nyquist-frequency case.

Hence, we conclude that the coherence spectra predicted by reduced models of enzyme catalysis are generally not accurate, even if parameter sets are chosen so that molecule numbers are predicted accurately. These results have implications for the MI rate between substrate and product, which is given by Eq. (9). Since $\xi_{S,P}^{\text{heur}}$ is a constant and $\xi_{S,P}^{\text{ssLNA}}$ tends to a constant for large ω , the MI rate is infinite according to the heuristic LNA and the ssLNA. Since the full LNA predicts $\xi_{S,P}^{\text{full}} \propto \omega^{-2}$ for large ω , it follows from Appendix B that the MI rate is always finite and positive, even when timescale

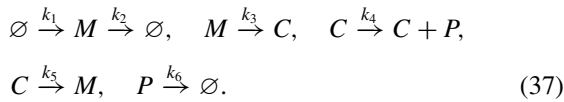
separation holds. In fact, the MI rate can be quite small, in contradiction to the infinite rate predicted by the reduced LNAs. For example, for the parameter set used in Fig. 1, the MI rate computed by substituting the coherence predicted by the full LNA, given by Eq. (30), into Eq. (9) is merely 1.83 bits per unit time.

B. Models of protein translation

Next, we consider two different models of protein translation and calculate, for each of those, the coherence and MI rate between mRNA and protein fluctuations. We hence quantify the information flow from mRNA to the protein, which is a topic that has attracted significant attention in the past two decades, as the correlation between the two is not as high as expected [46–49].

1. Simple model of translation

Consider the following system of reactions:



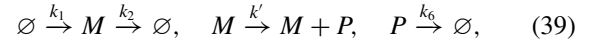
We model translation using a minimal reaction scheme in which cytoplasmic mRNA M is transcribed at rate k_1 and degraded at rate k_2 . Free mRNA can bind ribosomes, which are assumed to be in large excess, forming a translating complex C at rate k_3 . From the complex, protein P is synthesized at rate k_4 without the complex being consumed. Ribosome unbinding from the mRNA occurs at rate k_5 , releasing the mRNA. Finally, protein molecules are degraded (or diluted) at rate k_6 . We refer to the LNA corresponding to this circuit as the full LNA.

From the deterministic rate equations, one finds that the timescales of the three species are given by

$$\tau_M = \frac{1}{k_2 + k_3}, \quad \tau_P = \frac{1}{k_6}, \quad \text{and} \quad \tau_C = \frac{1}{k_5}. \quad (38)$$

When $\tau_C \ll \tau_M, \tau_P$, perturbations about the steady-state concentration of the complex decay much faster than those about

the steady-state concentrations of the mRNA and protein. The complex is then in a quasisteady state, with the deterministic rate equations simplifying to those of an effective reduced circuit,



where $k' = k_3 k_4 / k_5$. We refer to the LNA corresponding to this circuit as the heuristic LNA. A second, more rigorous way to obtain a reduced LNA when the complex is fast relies on the ssLNA. The corresponding calculations can be found in Appendix D.

The predictions obtained from the three LNAs for the steady-state mean concentrations coincide exactly. However, the predicted second moments at the steady state differ. The timescale separation condition corresponds to $\tau_M, \tau_P \gg \tau_C$. There are different options for enforcing that condition, with varying consequences for the relationships between the second moments resulting from the three LNAs. For example, if we set $k_2 = z$, $k_3 = a_1 z$, and $k_6 = a_2 z$, with a_1 and a_2 some positive constants, and take the limit of $z \rightarrow 0$, then these relationships simplify to

$$\frac{C_M^{\text{full}}}{C_M^{\text{heur}}} = 1, \quad \frac{C_{M,P}^{\text{full}}}{C_{M,P}^{\text{heur}}} = 1, \quad \frac{C_P^{\text{full}}}{C_P^{\text{heur}}} = 1 + \frac{(a_2 + 1)\lambda}{a_1 \lambda + a_2 + 1}, \quad (40)$$

$$\frac{C_M^{\text{full}}}{C_M^{\text{ssLNA}}} = 1, \quad \frac{C_{M,P}^{\text{full}}}{C_{M,P}^{\text{ssLNA}}} = 1, \quad \text{and} \quad \frac{C_P^{\text{full}}}{C_P^{\text{ssLNA}}} = 1, \quad (41)$$

where $\lambda = k_4 / k_5$. Note that while the ssLNA perfectly agrees with the full LNA in this limit, the heuristic LNA does not agree with the full LNA unless the further constraint of $\lambda \rightarrow 0$ is enforced, which we shall call the “low translation constraint.” A different way to impose a timescale separation is to take $k_5 \rightarrow \infty$, in which case all three LNAs agree: in this limit, one automatically enforces both the timescale separation condition $\tau_M, \tau_P \gg \tau_C$ and the low translation constraint ($\lambda \rightarrow 0$).

The coherence spectra of the mRNA and protein, as predicted by the full LNA, the heuristic LNA, and the ssLNA, are given by

$$\xi_{M,P}^{\text{full}}(\omega) = \frac{\alpha \lambda (\gamma - \omega^2 \tau_C^2)^2}{[\gamma + \omega^2 (\alpha + \gamma) \tau_C^2] \{ \gamma [\lambda (\alpha + \gamma) + \gamma] + \omega^2 \tau_C^2 [\alpha^2 + 2\alpha(\gamma + 1) + \gamma^2 + \lambda + 1] + \omega^4 \tau_C^4 \}}, \quad (42)$$

$$\xi_{M,P}^{\text{ssLNA}}(\omega) = \frac{\alpha \gamma \lambda}{\gamma [\lambda (\alpha + \gamma) + \gamma] + (\lambda + 1) \omega^2 \tau_C^2}, \quad \text{and} \quad (43)$$

$$\xi_{M,P}^{\text{heur}}(\omega) = \frac{\alpha \gamma \lambda}{\gamma (\alpha \lambda + \gamma) + \omega^2 \tau_C^2}, \quad (44)$$

respectively, where $\alpha = k_3 / k_5$, $\beta = k_6 / k_5$, $\lambda = k_4 / k_5$, and $\gamma = k_2 / k_5$. Note that $\alpha + \gamma = \tau_C / \tau_M$ and $\beta = \tau_C / \tau_P$.

The coherences predicted by the full LNA and the ssLNA agree exactly at $\omega = 0$, while the coherence of the heuristic LNA differs,

$$\xi_{M,P}^{\text{full}}(0) = \xi_{M,P}^{\text{ssLNA}}(0) = \frac{\alpha \lambda}{\lambda (\alpha + \gamma) + \gamma} \neq \xi_{M,P}^{\text{heur}}(0) = \frac{\alpha \lambda}{\alpha \lambda + \gamma}. \quad (45)$$

These relationships constitute a special case of the general derivation in Sec. IV A. We also note that the large-frequency tails of the coherence spectra of the three LNAs scale as ω^{-2} , which follows directly from the general results in Sec. IV B—the agreement stems from the fact that in both the full model, given by Eq. (37), and the reduced model, given by Eq. (39), no reactions occur that cause a simultaneous change in the net numbers of mRNA and product molecules.

However, we note that for intermediate frequencies, there is a major difference between the coherence spectra. While both the ssLNA and the heuristic LNA predict a monotonically decreasing coherence spectrum with increasing ω , the full LNA does not. Rather, it predicts that as ω increases from 0, the coherence first decreases and attains the value of 0 at the critical frequency $\omega_c = \sqrt{\gamma/\tau_C} = \sqrt{k_2 k_5}$; it then increases again, reaching a peak before decreasing as ω^{-2} . Note that this nonmonotonic behavior is present for all finite values of the rate constants, and hence it is generally impossible to recover the monotonicity predicted by the reduced LNAs.

As an example, we reconsider one of the limits applied previously to enforce timescale separation: setting $k_2 = z$, $k_3 = a_1 z$, and $k_6 = a_2 z$, where a_1 and a_2 are some positive constants, and taking the limit of $z \rightarrow 0$ at fixed ω , we find $\omega_c \rightarrow 0$, which implies that the coherence spectrum of the full LNA now exhibits a coherence of 0 at $\omega = 0$, increases to reach a peak, and then decreases as ω^{-2} , which still does not agree with the monotonically decreasing behavior predicted by the ssLNA and the heuristic LNA. (Still, in this limit, the first and second moments of mRNA and protein noise predicted by the ssLNA agree exactly with those obtained from the full LNA.) We can also consider the second limit introduced before to enforce timescale separation, with $k_5 \rightarrow \infty$ at fixed ω . We first note that the latter is impossible to achieve practically, as any chosen value of k_5 will still be finite despite being very large—hence, ω_c is large but finite, which implies that if the SSA output is sampled on a sufficiently fine time grid, then the coherence spectrum will always display a peak before decaying as ω^{-2} , in contradiction to the monotonic predictions of the ssLNA and the heuristic LNA.

The nonvanishing differences between the full LNA and the reduced LNAs are also apparent if we consider the large-frequency tails of the coherence spectra,

$$\xi_{M,P}^{\text{full}}(\omega) \sim \frac{\alpha\lambda}{(\alpha + \gamma)\tau_C^2} \frac{1}{\omega^2}, \quad (46)$$

$$\xi_{M,P}^{\text{ssLNA}}(\omega) \sim \frac{\alpha\gamma\lambda}{(\lambda + 1)\tau_C^2} \frac{1}{\omega^2}, \quad \text{and} \quad (47)$$

$$\xi_{M,P}^{\text{heur}}(\omega) \sim \frac{\alpha\gamma\lambda}{\tau_C^2} \frac{1}{\omega^2}. \quad (48)$$

If we again impose the low translation constraint ($\lambda \rightarrow 0$), then the ssLNA and the heuristic LNA agree; however, the full LNA only agrees with those if we apply the additional constraint that $\gamma = (\alpha + \gamma)^{-1}$. That constraint is impossible to satisfy under timescale separation conditions, however, since it requires α and γ to be very small due to $\alpha + \gamma = \tau_C/\tau_M$.

We tested some of these analytical conclusions by means of stochastic simulation; the results are shown in Fig. 2(a). We chose a parameter set that enforces timescale separation conditions, with $\tau_C/\tau_M = 0.02$ and $\tau_C/\tau_P = 0.005$, and then compared the coherence spectra predicted by the full LNA, the ssLNA, the heuristic LNA, the SSA of the full model, given by Eq. (37), and the SSA of the reduced model, given by Eq. (39). As expected, the full LNA and the SSA of the full model agreed exactly, as did the heuristic LNA and the

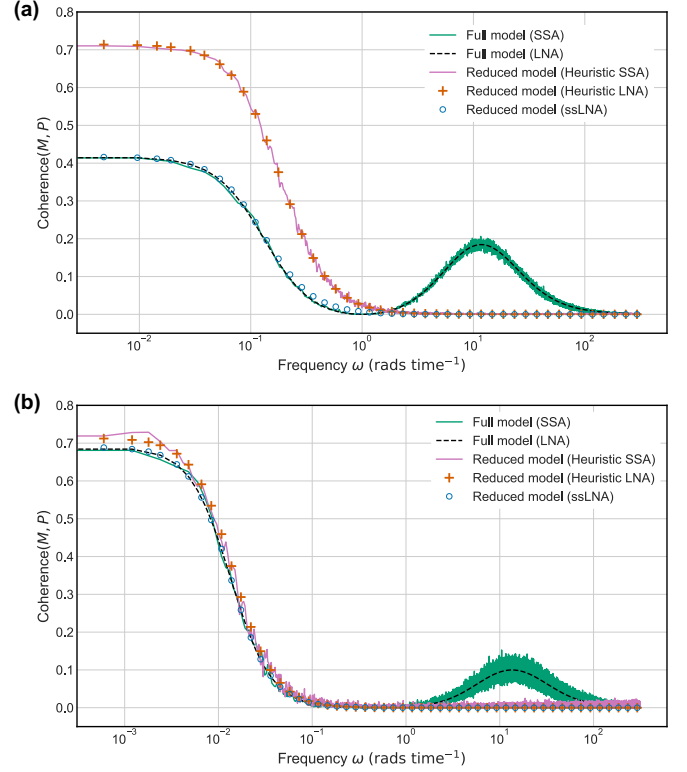


FIG. 2. Comparison of the coherence spectrum of the full and reduced models for the simple translation model. We consider two parameter sets: (a) $k_1 = 10$, $k_2 = 0.1$, $k_3 = 0.1$, $k_4 = 25$, $k_5 = 10$, and $k_6 = 0.05$; (b) $k_1 = 10$, $k_2 = 0.0075$, $k_3 = 0.1$, $k_4 = 5$, $k_5 = 25$, and $k_6 = 0.005$. Both enforce timescale separation conditions, with $\tau_C/\tau_M = 0.02$ and $\tau_C/\tau_P = 0.005$ for (a), and $\tau_C/\tau_M = 0.0043$ and $\tau_C/\tau_P = 0.0002$ for (b). While the first parameter set does not obey the low translation constraint, the second set does since $k_4 \ll k_5$. The coherence was estimated numerically with Welch's method from SSA simulations of the full model [Eq. (37)] and the reduced model [Eq. (39)]; the SSA output was sampled every 0.01 time units for a total duration of a few million time units. The Nyquist frequency is $\omega_{\max} \approx 314$. We also show the coherence predicted by the theory in Eqs. (42)–(44) for the full LNA, the ssLNA, and the heuristic LNA.

SSA of the reduced model. The coherence at $\omega = 0$ of the full LNA was perfectly predicted by the ssLNA, which accurately predicts the entire LNA spectrum up to $\omega \approx \omega_c = 1$, but becomes inaccurate for higher frequencies—in particular, since the ssLNA prediction is monotonically decreasing with frequency, it completely misses the peak at $\omega \approx 10$. Note that the second moments of the full LNA and the ssLNA are very close, as claimed above: $\frac{(\sigma_M^2)^{\text{full}}}{(\sigma_M^2)^{\text{ssLNA}}} = 1$, $\frac{(\sigma_P^2)^{\text{full}}}{(\sigma_P^2)^{\text{ssLNA}}} = 0.992$, and $\frac{(\sigma_{MP})^{\text{full}}}{(\sigma_{MP})^{\text{ssLNA}}} = 0.993$. By contrast, the heuristic LNA prediction is inaccurate for all values of ω , which stems from the fact that the chosen parameter set enforces timescale separation conditions, but not the additional low translation constraint of small λ . When this additional constraint is applied, the ssLNA and the heuristic LNA are of similar accuracy [Fig. 2(b)].

These results have significant implications for the MI rate. For the example in Fig. 2(b), we apply Eq. (9) to compute the MI rate in bits per unit time using the full LNA, the ssLNA,

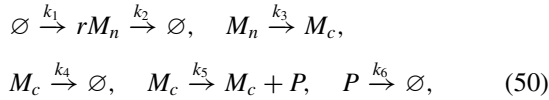
and the heuristic LNA,

$$M_{M,P}^{\text{full}} = 1.26, \quad M_{M,P}^{\text{ssLNA}} = 4.3 \times 10^{-3}, \quad \text{and} \\ M_{M,P}^{\text{heur}} = 4.9 \times 10^{-3}. \quad (49)$$

Hence, the inaccuracies in the coherence spectrum predicted by the reduced LNAs for intermediate and large frequencies lead to a massive underestimation, by a factor of about 300, of the MI rate.

2. Model of translation with compartmentalization and transcriptional bursting

Consider the system of reactions



in which basal transcription produces nuclear mRNA (M_n) from DNA in bursts of size r . We assume that each burst size r is independently drawn from a geometric distribution with mean b , which models transcriptional bursting [50,51]. Newly synthesized M_n can be degraded in the nucleus or exported to the cytoplasm, where it is labeled as cytoplasmic mRNA (M_c). Cytoplasmic mRNA is translated by ribosomes to produce protein P , with both M_c and P subject to first-order degradation. For simplicity, we do not explicitly model the mRNA-ribosome complex C , as we did in the reaction scheme in Eq. (37); however, qualitatively similar results are obtained if we add that additional species. In what follows, we refer to the LNA corresponding to the circuit in Eq. (50) as the full LNA.

From the deterministic rate equations, one finds that the timescales of the three species are given by

$$\tau_{M_n} = \frac{1}{k_2 + k_3}, \quad \tau_{M_c} = \frac{1}{k_4}, \quad \text{and} \quad \tau_P = \frac{1}{k_6}. \quad (51)$$

When $\tau_{M_n} \ll \tau_{M_c}, \tau_P$, perturbations about the steady-state concentration of nuclear mRNA decay much faster than those about the steady-state concentrations of cytoplasmic mRNA and protein. Nuclear mRNA is then in a quasisteady state, with the deterministic rate equations simplifying to those of an effective reduced circuit,



where $k' = k_1 k_3 / (k_2 + k_3)$. The properties of the random variable r are as for the full circuit. We refer to the LNA corresponding to Eq. (52) as the heuristic LNA. A second, more rigorous way to obtain a reduced LNA when the complex is fast relies on the ssLNA. The underlying calculations can be found in Appendix E.

In the limit of $k_2 \rightarrow \infty$, nuclear mRNA is very short lived compared to the other two species [see Eq. (51)] and the second moments at the steady state predicted by the ssLNA match those in the full LNA; however, those predicted by the heuristic LNA do not:

$$\frac{C_{M_c}^{\text{full}}}{C_{M_c}^{\text{heur}}} = \frac{1}{1+b}, \quad \frac{C_{M_c,P}^{\text{full}}}{C_{M_c,P}^{\text{heur}}} = \frac{1}{1+b}, \\ \frac{C_P^{\text{full}}}{C_P^{\text{heur}}} = \frac{k_4 + k_5 + k_6}{k_4 + k_5(1+b) + k_6}, \\ \frac{C_{M_c}^{\text{full}}}{C_{M_c}^{\text{ssLNA}}} = 1, \quad \frac{C_{M_c,P}^{\text{full}}}{C_{M_c,P}^{\text{ssLNA}}} = 1, \quad \text{and} \quad \frac{C_P^{\text{full}}}{C_P^{\text{ssLNA}}} = 1. \quad (53)$$

For the heuristic LNA to agree with the full LNA and the ssLNA, the mean burst size b must be much smaller than 1; that is an unrealistic constraint since typical burst sizes for genes in eukaryotic cells are approximately in the range of 2–140; see Table I of [52].

The coherence spectra of the cytoplasmic mRNA and protein predicted by the full LNA, the heuristic LNA, and the ssLNA are given by

$$\xi_{M_c,P}^{\text{full}}(\omega) = \frac{k_4 k_5 \{ (k_2 + k_3) [k_2 + k_3(1+b)] + \omega^2 \}}{(k_2 + k_3) k_4 \{ k_2(k_4 + k_5) + k_3[k_4 + k_5(1+b)] \} + [(k_2 + k_3)^2 + k_4(k_4 + k_5)] \omega^2 + \omega^4}, \quad (54)$$

$$\xi_{M_c,P}^{\text{ssLNA}}(\omega) = \frac{[k_2 + k_3(1+b)] k_4 k_5}{k_2 k_4(k_4 + k_5) + k_3 k_4[k_4 + k_5(1+b)] + (k_2 + k_3) \omega^2}, \quad \text{and} \quad (55)$$

$$\xi_{M_c,P}^{\text{heur}}(\omega) = \frac{(1+b) k_4 k_5}{k_4[k_4 + k_5(1+b)] + \omega^2}, \quad (56)$$

respectively. As expected, the full LNA and the ssLNA agree exactly at zero frequency, but the heuristic LNA does not. In the limit of large frequencies, the coherence spectra reduce to

$$\xi_{M_c,P}^{\text{full}}(\omega) \sim k_4 k_5 \frac{1}{\omega^2}, \quad (57)$$

$$\xi_{M_c,P}^{\text{ssLNA}}(\omega) \sim \frac{(k_2 + k_3 + b k_3) k_4 k_5}{k_2 + k_3} \frac{1}{\omega^2}, \quad \text{and} \quad (58)$$

$$\xi_{M_c,P}^{\text{heur}}(\omega) \sim (1+b) k_4 k_5 \frac{1}{\omega^2}. \quad (59)$$

Clearly, under timescale separation conditions, with $k_2 \rightarrow \infty$, the tails of the coherence spectra of the ssLNA and the full

LNA agree, which is the first example in this paper where such agreement is found: for the metabolic and other translation models that we previously studied, the tails of the coherence spectra did not match in the timescale separation limit.

In fact, we find that in this limit, it is not merely the tails, but also the full coherence spectra of the full LNA and the ssLNA which converge to the same expression,

$$\xi_{M_c,P}^{\text{full}}(\omega), \quad \xi_{M_c,P}^{\text{ssLNA}}(\omega) \rightarrow \frac{k_4 k_5}{k_4(k_4 + k_5) + \omega^2}, \quad (60)$$

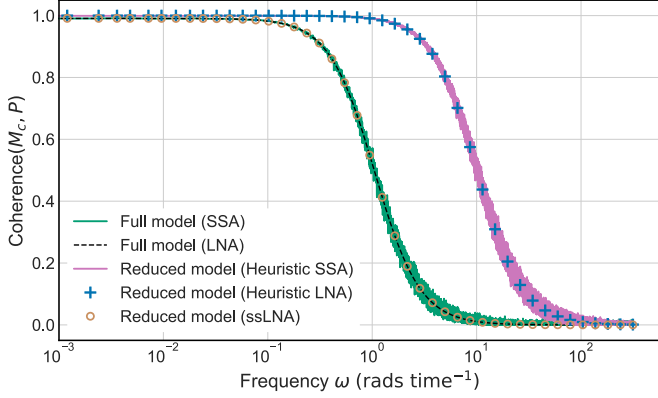


FIG. 3. Comparison of the coherence spectrum of the full and reduced models for a model of translation that includes transcriptional bursting and nuclear export. We choose the parameter set $b = 100$, $k_1 = 5$, $k_2 = 1000$, $k_3 = 1$, $k_4 = 0.1$, $k_5 = 10$, and $k_6 = 0.05$, which enforces timescale separation conditions due to $\tau_{M_n}/\tau_{M_c} = 10^{-4}$ and $\tau_{M_n}/\tau_P = 5 \times 10^{-5}$. The coherence was estimated numerically with Welch's method from SSA simulations of the full model [Eq. (50)] and the reduced model [Eq. (52)]; the SSA output was sampled every 0.01 time units for a total duration of a few million time units. The Nyquist frequency is $\omega_{\max} \approx 314$. We also show the coherence predicted by the theory in Eqs. (54) through (56) for the full LNA, the ssLNA, and the heuristic LNA.

while the coherence spectrum of the heuristic LNA converges to a different limit,

$$\xi_{M_c, P}^{\text{heur}}(\omega) \rightarrow \frac{(1+b)k_4k_5}{k_4[k_4 + k_5(1+b)] + \omega^2}. \quad (61)$$

Correspondingly, the MI rates calculated from Eq. (9) are given by

$$M_{M_c, P}^{\text{full}}, M_{M_c, P}^{\text{ssLNA}} \rightarrow \frac{\sqrt{k_4(k_4 + k_5)} - k_4}{\ln 4} \quad \text{and} \quad (62)$$

$$M_{M_c, P}^{\text{heur}} \rightarrow \frac{\sqrt{k_4[k_4 + (1+b)k_5]} - k_4}{\ln 4}, \quad (63)$$

respectively. Hence, we have shown that it is possible for the coherence spectrum and the MI rate to be correctly preserved when a model is reduced through the ssLNA. The underlying reason is likely that here, the fast species (nuclear mRNA) affects the slow species (cytoplasmic mRNA and protein), but that it itself is not affected by them, in contrast to previous models. However, our results also show that use of the commonly applied heuristic method of model reduction does not give the correct coherence spectrum and MI rate in the limit of timescale separation; in particular, the MI rate is always overestimated by the heuristic method. In Fig. 3, we verify our analytical results for the coherence spectra using stochastic simulation.

Note that in the aforementioned discussion, we have enforced timescale separation by taking the limit as $k_2 \rightarrow \infty$. A different way to enforce the timescale separation condition ($\tau_{M_n} \ll \tau_{M_c}, \tau_P$) involves taking the limit of $k_3 \rightarrow \infty$ [see Eq. (51)], which leads to different results than previously reported: (i) the second moments of the full LNA, the ssLNA, and the heuristic LNA agree exactly in this limit [recall

Eq. (53)]; and (ii) the tails of the coherence spectra of the full LNA and the ssLNA do not agree in the timescale separation limit. The latter can be directly deduced by inspection of Eqs. (57)–(59) which show that in the limit as $k_3 \rightarrow \infty$, the coherence spectra predicted by the heuristic LNA and the ssLNA converge to a single common expression that is different from the one obtained from the full LNA. Hence, we have shown that it is not simply the method of model reduction which affects these predictions, but also *the particular parameter limit by which timescale separation is enforced*.

VI. DISCUSSION

In this paper, we have demonstrated that although model reduction methodologies may reliably estimate the molecule number statistics of the underlying full models under timescale separation conditions, they do not necessarily provide accurate estimates for coherence spectra or MI rates. We have shown that even when reduced models are derived using rigorous methods such as the ssLNA, they can substantially misestimate coherence, particularly at intermediate and high frequencies, consequently leading to either significantly lower or significantly higher values for the MI rates than are predicted by the corresponding full models. Moreover, we have shown that commonly used heuristic reduction methods can further exacerbate these inaccuracies by neglecting noise from fast reactions.

One of the main issues we identified is that reduced models may incorrectly predict the scaling laws characterizing the tails of coherence spectra, with considerable repercussions for the calculation of the associated MI rates. That issue inherently results from the fact that model reduction implicitly involves a timescale separation limit, with one or more parameter values in the full model tending to either zero or infinity at *constant* frequency, a procedure that generally does not faithfully preserve the large-frequency tails of coherence spectra. For example, in the enzyme-catalyzed system [Eq. (26)], the LNA predicts that for all parameter values, including those enforcing timescale separation, the coherence scales as ω^{-2} for large frequencies, whereas the LNA of the reduced model [Eq. (29)] predicts a constant that is independent of the frequency and the parameter set. Another example is provided by the simple protein translation model, given by Eq. (37), and its reduced version, given by Eq. (39), where the tails of the coherence spectra scale as ω^{-2} , but with different prefactors that do not converge to each other in the timescale separation limit. In that same example, the full model predicts that for all parameter sets, there exists a peak in the coherence spectrum at intermediate frequencies, whereas the reduced models predict a monotonically decreasing coherence spectrum.

Furthermore, we have shown that while model reduction may not generally preserve coherence spectra and MI rates, there are specific cases where it does; for example, when modeling compartmentalization, transcriptional bursting, and protein translation in Sec. VB 2, the full model and the rigorously reduced model from the ssLNA predicted the same coherence spectra for *all frequencies* when timescale separation was enforced by taking the nuclear mRNA degradation rate to infinity. The precise conditions under which such perfect agreement arises remain unclear; however, it is likely in part

attributable to the absence of feedback from the slow species to the fast species in the full model. That cannot be the only reason, though, as we have also shown that if timescale separation is enforced differently, by taking the nuclear export rate to infinity, the coherence spectrum of the ssLNA no longer converges to that of the full LNA. Hence, our results indicate that the coherence spectrum and the MI rate are, in general, not uniquely defined under timescale separation conditions.

We note that constant coherence at large frequencies, and the associated infinite MI rate, were previously identified for the simple biochemical system $\emptyset \rightarrow A \rightarrow \emptyset$, $A \rightarrow B \rightarrow \emptyset$; see Motif II in [27,28] and Motif b in [53]. Our results in Sec. IV B extend this special case: we show that any biochemical system to which the LNA is applicable and which contains at least one chemical reaction that simultaneously changes the copy numbers of species A and B yields constant high-frequency coherence between these species under the LNA or its reduced forms. In [53], it was further shown that for the circuit mentioned above, a finite MI rate can be recovered if an alternative formulation of the LNA is used which retains reaction-specific information in the trajectories and which does not rely on coherence as an intermediate step in the MI rate calculation. However, from the model reduction point of view in the present paper, a more fundamental issue is that the circuit $\emptyset \rightarrow A \rightarrow \emptyset$, $A \rightarrow B \rightarrow \emptyset$ is simply unrealistic: in reality, it represents the reduced form of a more complex circuit such as $\emptyset \rightarrow A \rightarrow \emptyset$, $A \rightarrow C_1 \rightarrow C_2 \rightarrow \dots \rightarrow C_N \rightarrow B \rightarrow \emptyset$, where the species C_i ($i = 1, \dots, N$) are fast intermediates. By the results of Sec. IV B, for such a circuit, the coherence decreases as ω^{-2} in the limit of large frequencies, with finite MI rate, since there is no reaction in which the molecule numbers of A and B change simultaneously when it occurs. In the future, it would be interesting to understand how the calculation of the MI rate via the alternative formulation of the LNA in [53] is affected by model reduction. Based on the example discussed above, our expectation is that the discrepancies between the MI rates predicted by the full LNA and the ssLNA are reduced in comparison to those observed here.

In conclusion, our results highlight the need for caution when employing reduced models to infer coherence or information-theoretic properties. Given that all models are approximations of reality, the question arises of how to design reduced descriptions that accurately capture not only molecular number statistics, but also the coherences and MI rates of the underlying systems, which can be achieved by systematically constructing models of increasing complexity until further refinements yield negligible changes in their predictions.

ACKNOWLEDGMENTS

J.D.M.-L. acknowledges financial support from a Ph.D. scholarship provided by the Darwin Trust. R.G. acknowledges financial support from the Leverhulme Trust (Grant No. RPG-2024-082). The authors thank Peter Swain for valuable discussions.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available. The data are available from the authors upon reasonable request.

APPENDIX A: PROOF OF SPECTRAL CONSISTENCY OF ssLNA AT $\omega = 0$

Here we show that the power spectra and the cross spectra of the slow species, as predicted by the full LNA and the ssLNA, coincide at zero frequency, i.e.,

$$P_{ij}(0) = [P_S]_{ij}(0), \quad (\text{A1})$$

where $i, j \in \{1, \dots, N_s\}$ and

$$P_{ij}(0) = [J^{-1} D_S J^{-T}]_{ij} \quad \text{and} \quad [P_S]_{ij}(0) = [J_S^{-1} D_S J_S^{-T}]_{ij}.$$

We proceed as follows: since $P_{ij}(0)$ is a function of J^{-1} , we first obtain an expression for that matrix. Subsequently, we evaluate the upper-left $N_s \times N_s$ submatrix of $J^{-1} D_S J^{-T}$ since we are only concerned with the slow species. We then simplify the resulting expression to show that it coincides exactly with $J_S^{-1} D_S J_S^{-T}$.

Let the (unknown) inverse of J in block form be given by

$$J^{-1} = \begin{pmatrix} \underline{X} & \underline{Y} \\ \underline{Z} & \underline{W} \end{pmatrix},$$

where the blocks $\underline{X}$, $\underline{Y}$, $\underline{Z}$, and $\underline{W}$ have dimensions $N_s \times N_s$, $N_s \times N_f$, $N_f \times N_s$, and $N_f \times N_f$, respectively. Note that from Sec. III B, we have

$$J = \begin{pmatrix} J_{ss} & J_{sf} \\ J_{fs} & J_{ff} \end{pmatrix}.$$

The matrix equation $J J^{-1} = I_M$, where I_M is the identity matrix of rank M , yields the block equations,

$$J_{ss} \underline{X} + J_{sf} \underline{Z} = I_{N_s}, \quad (\text{A2})$$

$$J_{ss} \underline{Y} + J_{sf} \underline{W} = \underline{0}, \quad (\text{A3})$$

$$J_{fs} \underline{X} + J_{ff} \underline{Z} = \underline{0}, \quad (\text{A4})$$

$$J_{fs} \underline{Y} + J_{ff} \underline{W} = I_{N_f}. \quad (\text{A5})$$

Assuming that J_{ff} is invertible, we solve Eq. (A4) for $\underline{Z}$ to find

$$\underline{Z} = -J_{ff}^{-1} J_{fs} \underline{X}. \quad (\text{A6})$$

Substituting Eq. (A6) into Eq. (A2), we can solve for $\underline{X}$,

$$\underline{X} = J_S^{-1}, \quad (\text{A7})$$

where we assumed that J_S , defined in Eq. (11), is invertible. Substitution of Eq. (A7) into Eq. (A6) yields

$$\underline{Z} = -J_{ff}^{-1} J_{fs} J_S^{-1}. \quad (\text{A8})$$

Next, we solve for $\underline{Y}$ and $\underline{W}$: from Eq. (A5), we have

$$\underline{W} = J_{ff}^{-1} (I_{N_f} - J_{fs} \underline{Y}). \quad (\text{A9})$$

Substituting Eq. (A9) into Eq. (A3) and solving for $\underline{Y}$, we find

$$\underline{Y} = -J_S^{-1} J_{sf} J_{ff}^{-1}. \quad (\text{A10})$$

Finally, we substitute Eq. (A10) into Eq. (A9), which gives the following expression for $\underline{W}$:

$$\underline{W} = J_{ff}^{-1} + J_{ff}^{-1} J_{fs} J_S^{-1} J_{sf} J_{ff}^{-1}.$$

Thus, the block-inverse formula for the Jacobian $\underline{J}$ of the full model is

$$\underline{J}^{-1} = \begin{pmatrix} \underline{J}_S^{-1} & -\underline{J}_S^{-1} \underline{J}_{sf} \underline{J}_{ff}^{-1} \\ -\underline{J}_{ff}^{-1} \underline{J}_{fs} \underline{J}_S^{-1} & \underline{J}_{ff}^{-1} + \underline{J}_{ff}^{-1} \underline{J}_{fs} \underline{J}_S^{-1} \underline{J}_{sf} \underline{J}_{ff}^{-1} \end{pmatrix}.$$

Let the upper-left $N_s \times N_s$ submatrix of $\underline{J}^{-1} \underline{D} \underline{J}^{-T}$ be denoted by $\underline{P}_{ss}$, and write the matrices in the triple product as

$$\underline{J}^{-1} = \begin{pmatrix} \underline{X} & \underline{Y} \\ \underline{Z} & \underline{W} \end{pmatrix}, \quad \underline{D} = \begin{pmatrix} \underline{D}_{ss} & \underline{D}_{sf} \\ \underline{D}_{fs} & \underline{D}_{ff} \end{pmatrix}, \quad \text{and} \\ \underline{J}^{-T} = \begin{pmatrix} \underline{X}^T & \underline{Z}^T \\ \underline{Y}^T & \underline{W}^T \end{pmatrix}.$$

It then follows that only the following terms contribute to the upper-left block of the product of these three matrices:

$$\underline{P}_{ss} = \underline{X} \underline{D}_{ss} \underline{X}^T + \underline{X} \underline{D}_{sf} \underline{Y}^T + \underline{Y} \underline{D}_{fs} \underline{X}^T + \underline{Y} \underline{D}_{ff} \underline{Y}^T. \quad (\text{A11})$$

Hence, to evaluate $\underline{P}_{ss}$, we only need the upper-left and upper-right blocks from the block-inverse formula for $\underline{J}^{-1}$, i.e., $\underline{X} = \underline{J}_S^{-1}$ and $\underline{Y} = -\underline{J}_S^{-1} \underline{J}_{sf} \underline{J}_{ff}^{-1}$. We also require the block elements of $\underline{D}$ —since $\underline{D} = \underline{S} \underline{F} \underline{S}^T = (\underline{S} \sqrt{\underline{F}})(\underline{S} \sqrt{\underline{F}})^T$, its block form is

$$\underline{D} = \begin{pmatrix} \underline{D}_{ss} & \underline{D}_{sf} \\ \underline{D}_{fs} & \underline{D}_{ff} \end{pmatrix} = \begin{pmatrix} \underline{A} \underline{A}^T & \underline{A} \underline{E}_f^T \\ \underline{E}_f \underline{A}^T & \underline{E}_f \underline{E}_f^T \end{pmatrix},$$

where $\underline{A} = \underline{S}_s \sqrt{\underline{F}}$ and $\underline{E}_f = \underline{S}_f \sqrt{\underline{F}}$. Substituting these expressions and the block formulas for $\underline{D}$ into Eq. (A11), we obtain

$$\begin{aligned} \underline{P}_{ss} &= \underline{J}_S^{-1} (\underline{A} \underline{A}^T) \underline{J}_S^{-T} + \underline{J}_S^{-1} (\underline{A} \underline{E}_f^T) (-\underline{M}^T \underline{J}_S^{-T}) \\ &\quad + (-\underline{J}_S^{-1} \underline{M}) (\underline{E}_f \underline{A}^T) \underline{J}_S^{-T} \\ &\quad + (-\underline{J}_S^{-1} \underline{M}) (\underline{E}_f \underline{E}_f^T) (-\underline{M}^T \underline{J}_S^{-T}), \end{aligned}$$

where $\underline{M} = \underline{J}_{sf} \underline{J}_{ff}^{-1}$. Factoring $\underline{J}_S^{-1}$ on the left and $\underline{J}_S^{-T}$ on the right, we find

$$\underline{P}_{ss} = \underline{J}_S^{-1} [\underline{A} \underline{A}^T - \underline{A} \underline{E}_f^T \underline{M}^T - \underline{M} \underline{E}_f \underline{A}^T + \underline{M} \underline{E}_f \underline{E}_f^T \underline{M}^T] \underline{J}_S^{-T}.$$

Now, recall that $\underline{B} = \underline{M} \underline{E}_f = \underline{J}_{sf} \underline{J}_{ff}^{-1} \underline{S}_f \sqrt{\underline{F}}$. It can also be shown from Eq. (12) that the diffusion matrix in the ssLNA is given by

$$\begin{aligned} \underline{D}_S &= (\underline{A} - \underline{B})(\underline{A} - \underline{B})^T = \underline{A} \underline{A}^T - \underline{A} \underline{B}^T - \underline{B} \underline{A}^T + \underline{B} \underline{B}^T \\ &= \underline{A} \underline{A}^T - \underline{A} \underline{E}_f^T \underline{M}^T - \underline{M} \underline{E}_f \underline{A}^T + \underline{M} \underline{E}_f \underline{E}_f^T \underline{M}^T. \end{aligned}$$

Therefore, we have

$$\underline{P}_{ss} = \underline{J}_S^{-1} \underline{D}_S \underline{J}_S^{-T}.$$

Hence, by Eqs. (6) and (15), it follows that the power spectra and the cross spectra of the slow species, as predicted by the full LNA and the ssLNA, coincide at zero frequency,

$$P_{ij}(0) = [\underline{P}_S]_{ij}(0), \quad (\text{A12})$$

where $i, j \in \{1, \dots, N_s\}$.

APPENDIX B: PROOF OF FINITENESS OF MI RATE WHEN $D_{ij} = 0$

We start by summarizing the well-known properties of the coherence ξ_{ij} between two signals which, in our case, are the

real-valued time series of the fluctuations in the numbers of molecules of the two species X_i and X_j : (i) $0 \leq \xi_{ij} \leq 1$ and (ii) $\xi_{ij}(\omega) = \xi_{ij}(-\omega)$.

Hence, it follows by Eq. (9) that the MI rate can be written as

$$\begin{aligned} M_{ij} &= -\frac{1}{4\pi \ln 2} \int_{-\infty}^{\infty} \ln[1 - \xi_{ij}(\omega)] d\omega \\ &= -\frac{1}{2\pi \ln 2} \int_0^{\infty} \ln[1 - \xi_{ij}(\omega)] d\omega. \end{aligned} \quad (\text{B1})$$

We divide the integral in Eq. (B1) into two parts,

$$\begin{aligned} M_{ij} &= -\frac{1}{2\pi \ln 2} \left\{ \int_0^R \ln[1 - \xi_{ij}(\omega)] d\omega \right. \\ &\quad \left. + \int_R^{\infty} \ln[1 - \xi_{ij}(\omega)] d\omega \right\}, \end{aligned} \quad (\text{B2})$$

where R is some positive number. The integral from 0 to R in Eq. (B2) is positive and finite since $\xi_{ij}(\omega)$ is a continuous function of ω that is bounded between 0 and 1. [In particular, $\xi_{ij}(\omega)$ is strictly below 1, as we are considering open systems in which fluctuations between any two different species can never be perfectly correlated due to synthesis occurring at random points in time.] If $D_{ij} = 0$, then we have proven in Sec. IV B that $\xi_{ij}(\omega) \propto A\omega^{-2}$ as $\omega \rightarrow \infty$, where A is a positive number; see Eq. (24). We choose R to be sufficiently large so that this scaling law applies. It then follows that

$$\int_R^{\infty} \ln[1 - \xi_{ij}(\omega)] d\omega \sim A \int_R^{\infty} \omega^{-2} d\omega = A/R > 0.$$

Hence, Eq. (B2) implies that if $D_{ij} = 0$, then the MI rate is a finite positive quantity.

APPENDIX C: CALCULATIONS FOR ENZYME REACTION NETWORKS

1. Full LNA and heuristic LNA

We make use of the LNA frameworks described in Secs. II and III A to calculate the coherence spectrum for the numbers of substrate and product molecules in the full and reduced enzyme systems under steady-state conditions.

The stoichiometric matrix and rate function vector for the full system, given by Eq. (26), are defined by

$$\begin{aligned} \underline{S}_{\text{full}} &= \begin{bmatrix} 1 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 1 & -1 & -1 & 0 \end{bmatrix} \quad \text{and} \\ \vec{f}_{\text{full}} &= [k_1, \quad k_2(\phi_T - \phi_C)\phi_S, \quad k_3\phi_C, \quad k_4\phi_C, \quad k_5\phi_P]^T, \end{aligned}$$

where the first, second, and third rows of $\underline{S}_{\text{full}}$ account for the net changes in the numbers of molecules of species S , P , and C , respectively. Note that free enzyme E does not need to be explicitly accounted for due to the conservation law between the free enzyme and complex.

Similarly, for the heuristically reduced system, given by Eq. (29), we can write

$$\underline{S}_{\text{heur}} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \quad \text{and} \quad \vec{f}_{\text{heur}} = \left[k_1, \frac{k_4 \phi_T \phi_S}{K_M + \phi_S}, k_5 \phi_P \right]^T,$$

where the first and second rows of the stoichiometric matrix $\underline{S}_{\text{heur}}$ account for the net changes in the numbers of molecules of species S and P , respectively.

The rate equations for the full system are given by Eq. (2), with $\underline{S}$ and $\vec{f}$ replaced by $\underline{S}_{\text{full}}$ and $\vec{f}_{\text{full}}$, respectively,

$$\frac{d\phi_S}{dt} = k_1 - k_2 \phi_S (\phi_T - \phi_C) + k_3 \phi_C, \quad (\text{C1})$$

$$\frac{d\phi_P}{dt} = k_4 \phi_C - k_5 \phi_P, \quad (\text{C2})$$

$$\frac{d\phi_C}{dt} = k_2 \phi_S (\phi_T - \phi_C) - (k_3 + k_4) \phi_C. \quad (\text{C3})$$

Their solution at the steady state yields the mean concentrations shown in Eq. (27).

The timescale of species X with concentration ϕ_X is derived via the substitution $\phi_X \mapsto \phi_X^* + \epsilon_X(t)$ in these rate equations: assuming a small perturbation $\epsilon_X(t)$ from the steady state ϕ_X^* , we obtain a linear differential equation for

$\epsilon_X(t)$ with solution $\epsilon_X(t) = \epsilon_X(0)e^{-t/\tau_X}$; the initial perturbation about the steady-state concentration of X hence decays exponentially with time. The resulting expressions for the timescales τ_X ($X = S, P, C$) are given in Eq. (28). Recall that in our heuristic derivation, we have assumed an initial perturbation in the concentration of species X only, neglecting the coupling between species in Eqs. (C1)–(C3); the general case of a simultaneous perturbation in the concentrations of all species is considered in Appendix C3.

The rate equations for the reduced system, which are obtained from Eq. (2) with $\underline{S}$ and $\vec{f}$ replaced by $\underline{S}_{\text{heur}}$ and $\vec{f}_{\text{heur}}$, respectively, read

$$\frac{d\phi_S}{dt} = k_1 - \frac{k_4 \phi_T \phi_S}{K_M + \phi_S}, \quad (\text{C4})$$

$$\frac{d\phi_P}{dt} = \frac{k_4 \phi_T \phi_S}{K_M + \phi_S} - k_5 \phi_P. \quad (\text{C5})$$

Their solution at the steady state is equally given by Eq. (27). Note that the reduced system of rate equations, given by Eqs. (C4) and (C5), can be obtained from the rate equations for the full system by setting $d\phi_C/dt = 0$ in Eq. (C3), solving for ϕ_C , and then substituting the resulting expression into Eqs. (C1) and (C2).

Given Eqs. (3) and (4), the corresponding Jacobian and diffusion matrices are given by

$$\underline{J}_{\text{full}} = \begin{bmatrix} -k_2(\phi_T - \frac{k_1}{k_4}) & 0 & \frac{(k_1 + \phi_T k_3)k_4}{\phi_T k_4 - k_1} \\ 0 & -k_5 & k_4 \\ k_2(\phi_T - \frac{k_1}{k_4}) & 0 & -\frac{\phi_T k_4 (k_3 + k_4)}{\phi_T k_4 - k_1} \end{bmatrix} \quad \text{and} \quad \underline{J}_{\text{heur}} = \begin{bmatrix} -\frac{(k_1 - \phi_T k_4)^2}{\phi_T k_4 K_M} & 0 \\ \frac{(k_1 - \phi_T k_4)^2}{\phi_T k_4 K_M} & -k_5 \end{bmatrix},$$

as well as by

$$\underline{D}_{\text{full}} = \begin{bmatrix} \frac{2k_1(k_3 + k_4)}{k_4} & 0 & -\frac{k_1(2k_3 + k_4)}{k_4} \\ 0 & 2k_1 & -k_1 \\ -\frac{k_1(2k_3 + k_4)}{k_4} & -k_1 & \frac{2k_1(k_3 + k_4)}{k_4} \end{bmatrix} \quad \text{and} \quad \underline{D}_{\text{heur}} = \begin{bmatrix} 2k_1 & -k_1 \\ m - k_1 & 2k_1 \end{bmatrix},$$

respectively. The covariances between the concentration fluctuations of the substrate and product, and the variances of the concentrations of each of these species, are found by solving Eq. (5) at the steady state,

$$C_S^{\text{full}} = \Omega^{-1} \frac{\phi_T x \{1 + \Gamma - x\Gamma + \lambda[1 + x^2 + \Gamma - x(1 + \Gamma)]\}}{(1 + \lambda)\Gamma(1 + \Gamma)}, \quad (\text{C6})$$

$$C_S^{\text{heur}} = \Omega^{-1} \frac{\phi_T x}{\Gamma}, \quad (\text{C7})$$

$$C_P^{\text{full}} = \Omega^{-1} \frac{\phi_T \lambda (x - 1)x[\lambda(x - 1)\Gamma(1 + \Gamma) - \{(1 + \Gamma)^2 + \lambda[(x - 1)x + (1 + \Gamma)^2]\}\Lambda - (1 + \lambda)(1 + \Gamma)\Lambda^2]}{(1 + \lambda)(1 + \Gamma)\Lambda[\Lambda(1 + \Gamma + \Lambda) + \lambda(\Gamma - x\Gamma + \Lambda + \Gamma\Lambda + \Lambda^2)]}, \quad (\text{C8})$$

$$C_P^{\text{heur}} = \Omega^{-1} \frac{\phi_T \lambda (1 - x)x}{(1 + \lambda)\Lambda}, \quad (\text{C9})$$

$$C_{S,P}^{\text{full}} = -\Omega^{-1} \frac{\phi_T \lambda (x - 1)x^2[\lambda(x - 1) - (1 + \lambda)\Lambda]}{(1 + \lambda)(1 + \Gamma)[\lambda(x - 1)\Gamma - (1 + \lambda)(1 + \Gamma)\Lambda - (1 + \lambda)\Lambda^2]}, \quad \text{and} \quad (\text{C10})$$

$$C_{S,P}^{\text{heur}} = 0, \quad (\text{C11})$$

where $x = k_1/(k_4 \phi_T)$, $\lambda = k_4/k_3$, $\Lambda = \tau_C/\tau_P$, and $\Gamma = \tau_C/\tau_S$. Note that $0 < x < 1$ (see the main text) and that the timescales τ_X are defined in Eq. (28).

In the limit of timescale separation, i.e., for $\Lambda \rightarrow 0$ and $\Gamma \rightarrow 0$, and of infrequent catalysis ($\lambda \rightarrow 0$), taken at constant x and ϕ_T , we find that the variance and covariance predictions of the full and heuristic LNAs agree: $C_S^{\text{full}}/C_S^{\text{heur}} \rightarrow 1$, $C_P^{\text{full}}/C_P^{\text{heur}} \rightarrow 1$, and $C_{S,P}^{\text{full}} - C_{S,P}^{\text{heur}} \rightarrow 0$, for all x .

The power spectrum matrices for the full and reduced systems are given by Eq. (6),

$$\underline{P}^{\text{full}}(\omega) = \Omega^{-1}(\underline{J}_{\text{full}} + i\underline{L}_3\omega)^{-1}\underline{D}_{\text{full}}(\underline{J}_{\text{full}}^T - i\underline{L}_3\omega)^{-1}, \quad (\text{C12})$$

$$\underline{P}^{\text{heur}}(\omega) = \Omega^{-1}(\underline{J}_{\text{heur}} + i\underline{L}_2\omega)^{-1}\underline{D}_{\text{heur}}(\underline{J}_{\text{heur}}^T - i\underline{L}_2\omega)^{-1}. \quad (\text{C13})$$

By Eq. (7), the coherences between the substrate and product timeseries for the full and reduced systems read

$$\xi_{S,P}^{\text{full}}(\omega) = \frac{|\underline{P}_{12}^{\text{full}}(\omega)|^2}{\underline{P}_{11}^{\text{full}}(\omega)\underline{P}_{22}^{\text{full}}(\omega)} = \frac{c_0 + c_1\omega^2 + c_2\omega^4}{c_3 + c_4\omega^2 + c_5\omega^4 + c_6\omega^6} \quad \text{and} \quad (\text{C14})$$

$$\xi_{S,P}^{\text{heur}}(\omega) = \frac{|\underline{P}_{12}^{\text{heur}}(\omega)|^2}{\underline{P}_{11}^{\text{heur}}(\omega)\underline{P}_{22}^{\text{heur}}(\omega)} = \frac{1}{4}, \quad (\text{C15})$$

where c_i are constants that are defined as

$$\begin{aligned} c_0 &= k_3^6 \lambda^3 (1 + \lambda)^3 \Gamma^2, \\ c_1 &= k_3^4 \lambda (1 + \lambda) \{ [1 + \lambda + 2\lambda(x - 1)x]^2 + 2(1 + \lambda)(1 + \lambda x)\Gamma + (1 + \lambda)^2 \Gamma^2 \}, \\ c_2 &= k_3^2 \lambda (1 + \lambda) (1 - 3x + 2x^2)^2, \\ c_3 &= 4k_3^6 \lambda^3 (1 + \lambda)^2 [1 + \lambda + \lambda(x - 1)x]\Gamma^2, \\ c_4 &= 4k_3^4 \lambda (1 + \lambda) \{ [1 + \lambda + \lambda(x - 1)x]^2 + 2(1 + \lambda x)[1 + \lambda + \lambda(x - 1)x]\Gamma + (1 + \lambda)\{1 + \lambda[2 + x(2x - 3)]\}\Gamma^2 \}, \\ c_5 &= 4k_3^2 (x - 1)^2 \{ (1 + 2\lambda)[1 + \lambda + \lambda(x - 1)x] + 2(1 + \lambda)(1 + \lambda x)\Gamma + (1 + \lambda)^2 \Gamma^2 \}, \quad \text{and} \\ c_6 &= 4(x - 1)^4. \end{aligned} \quad (\text{C16})$$

Since $0 < x < 1$, $\lambda > 0$, $k_3 > 0$, and $\Gamma > 0$, it follows that $c_i > 0$ for $i = 0, 1, 3, 4, 5, 6$, while c_2 is positive when $x \neq 1/2$ and equal to 0 when $x = 1/2$. An important implication is that for the full LNA, in the limit of large ω , $\xi_{S,P}^{\text{full}}(\omega) \propto \omega^{-2}$ when $x \neq 1/2$, whereas $\xi_{S,P}^{\text{full}}(\omega) \propto \omega^{-4}$ for $x = 1/2$. By contrast, $\xi_{S,P}^{\text{heur}}$ is independent of ω .

To understand the large discrepancy between the predictions of the full and heuristic LNAs, we consider the predicted power spectra of substrate and product fluctuations and their cross spectrum since these compose the coherence; see Eqs. (C14) and (C15). We find that both the full LNA and the heuristic LNAs predict that the tails of the power spectra of the substrate and product decay as ω^{-2} . However, the full LNA predicts that the tail of the squared modulus of the cross spectrum of the substrate and product decays as ω^{-6} , while the heuristic LNA predicts a slower decay with ω^{-4} if $k_1 \neq k_4\phi_T/2$. For the special case where $x = 1/2$, the squared modulus of the cross spectrum in the full LNA decays as ω^{-8} , while the heuristic LNA still predicts a slower decay at rate ω^{-4} . Hence, the discrepancy between the coherences predicted by the two LNAs is due to discrepancies in the tails of the cross spectra.

2. ssLNA

We apply the ssLNA framework described in Sec. III B to calculate the coherence spectrum for the substrate and

product, assuming that they are the slow species, while the enzyme and complex are the fast species.

The partitioned stoichiometric matrices of the slow and fast species are given by

$$\underline{S}_s = \begin{bmatrix} 1 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix} \quad \text{and} \quad \underline{S}_f = [0 \quad 1 \quad -1 \quad -1 \quad 0].$$

Note that $\underline{S}_s$ corresponds to the elements in the first and second rows of $\underline{S}_{\text{full}}$, while $\underline{S}_f$ equals the last row of $\underline{S}_{\text{full}}$ in Appendix C 1.

The partitioned Jacobian matrices are given by

$$\underline{J}_{ss} = \begin{bmatrix} -k_2(\phi_T - \frac{k_1}{k_4}) & 0 \\ 0 & -k_5 \end{bmatrix}, \quad \underline{J}_{sf} = \begin{bmatrix} \frac{(k_1 + \phi_T k_3)k_4}{\phi_T k_4 - k_1} \\ k_4 \end{bmatrix},$$

$$\underline{J}_{fs} = [k_2(\phi_T - \frac{k_1}{k_4}) \quad 0], \quad \text{and} \quad \underline{J}_{ff} = -\frac{\phi_T k_4 (k_3 + k_4)}{\phi_T k_4 - k_1}.$$

These matrices can be extracted from the Jacobian matrix $\underline{J}_{\text{full}}$ in Appendix C 1: $\underline{J}_{ss}$ is the upper-left 2×2 submatrix; $\underline{J}_{sf}$ is the upper-right 2×1 submatrix; $\underline{J}_{fs}$ is the lower-left 1×2 submatrix; and $\underline{J}_{ff}$ is the lower-right 1×1 submatrix, or scalar. Then, we can calculate the effective diffusion and Jacobian matrices as

$$\underline{A} = \underline{S}_s \sqrt{\underline{F}} = \begin{bmatrix} \sqrt{k_1} & -\sqrt{\frac{k_1(k_3 + k_4)}{k_4}} & \sqrt{\frac{k_1 k_3}{k_4}} & 0 & 0 \\ 0 & 0 & 0 & \sqrt{k_1} & -\sqrt{k_1} \end{bmatrix},$$

$$\underline{B} = \underline{J}_{sf} \underline{J}_{ff}^{-1} \underline{J}_f \sqrt{\underline{F}} = \begin{bmatrix} 0 & -\frac{(k_1 + \phi_T k_3) \sqrt{\frac{k_1(k_3+k_4)}{k_4}}}{\phi_T(k_3+k_4)} & \frac{(k_1 + \phi_T k_3) \sqrt{\frac{k_1 k_3}{k_4}}}{\phi_T(k_3+k_4)} & \frac{\sqrt{k_1}(k_1 + \phi_T k_3)}{\phi_T(k_3+k_4)} & 0 \\ 0 & \frac{\sqrt{\frac{k_1(k_3+k_4)}{k_4}}(k_1 - \phi_T k_4)}{\phi_T(k_3+k_4)} & \frac{\sqrt{\frac{k_1 k_3}{k_4}}(\phi_T k_4 - k_1)}{\phi_T(k_3+k_4)} & \frac{\sqrt{k_1}(\phi_T k_4 - k_1)}{\phi_T(k_3+k_4)} & 0 \end{bmatrix},$$

$$\underline{D}_S = (\underline{A} - \underline{B})(\underline{A} - \underline{B})^T = \begin{bmatrix} \frac{2k_1(k_1^2 - \phi_T k_1 k_4 + \phi_T^2 k_4(k_3+k_4))}{\phi_T^2 k_4(k_3+k_4)} & -\frac{k_1(2k_1^2 - 2\phi_T k_1 k_4 + \phi_T^2 k_4(k_3+k_4))}{\phi_T^2 k_4(k_3+k_4)} \\ -\frac{k_1(2k_1^2 - 2\phi_T k_1 k_4 + \phi_T^2 k_4(k_3+k_4))}{\phi_T^2 k_4(k_3+k_4)} & \frac{2k_1(k_1^2 - \phi_T k_1 k_4 + \phi_T^2 k_4(k_3+k_4))}{\phi_T^2 k_4(k_3+k_4)} \end{bmatrix},$$

and

$$\underline{J}_S = \underline{J}_{ss} - \underline{J}_{sf} \underline{J}_{ff}^{-1} \underline{J}_{fs} = \begin{bmatrix} -\frac{(k_1 - \phi_T k_4)^2}{\phi_T k_4 K_M} & 0 \\ \frac{(k_1 - \phi_T k_4)^2}{\phi_T k_4 K_M} & -k_5 \end{bmatrix}.$$

Note that $\underline{F}$ is a diagonal matrix with elements corresponding to the components of the vector $\vec{f}_{\text{full}}$; recall Appendix C 1. The covariance between the fluctuations of the substrate and product concentrations, and the variances of the concentrations of each of these species, are found from Eq. (5) with $\underline{J} = \underline{J}_S$ and $\underline{D} = \underline{D}_S$ at the steady state,

$$C_S^{\text{ssLNA}} = \frac{\phi_T x [1 + \lambda + \lambda(x-1)x]}{(1+\lambda)\Gamma}, \quad (\text{C17})$$

$$C_{S,P}^{\text{ssLNA}} = \frac{\phi_T \lambda^2 (x-1)^2 x^2}{(1+\lambda)[\Lambda + \lambda(\Gamma - x\Gamma + \Lambda)]}, \quad \text{and} \quad (\text{C18})$$

$$C_P^{\text{ssLNA}} = \frac{\phi_T \lambda (1-x)x \{\Lambda + \lambda[\Gamma - x\Gamma + \Lambda + (x-1)x\Lambda]\}}{(1+\lambda)\Lambda[\Lambda + \lambda(\Gamma - x\Gamma + \Lambda)]}. \quad (\text{C19})$$

In the limit of timescale separation, i.e., for $\Lambda \rightarrow 0$ and $\Gamma \rightarrow 0$, taken at constant x , ϕ_T , and λ , we find that the variances and covariances predicted by the full LNA and the ssLNA agree: $C_S^{\text{full}}/C_S^{\text{ssLNA}} \rightarrow 1$, $C_P^{\text{full}}/C_P^{\text{ssLNA}} \rightarrow 1$, and $C_{S,P}^{\text{full}}/C_{S,P}^{\text{ssLNA}} \rightarrow 1$, for all x .

The power spectrum matrix is found from Eq. (6),

$$\underline{P}^{\text{ssLNA}}(\omega) = \Omega^{-1}(\underline{J}_S + i\underline{L}_2\omega)^{-1} \underline{D}_S (\underline{J}_S^T - i\underline{L}_2\omega)^{-1}. \quad (\text{C20})$$

The coherence between the substrate and product timeseries is then given by Eq. (7),

$$\xi_{S,P}^{\text{ssLNA}}(\omega) = \frac{|\underline{P}_{12}^{\text{ssLNA}}(\omega)|^2}{\underline{P}_{11}^{\text{ssLNA}}(\omega) \underline{P}_{22}^{\text{ssLNA}}(\omega)} = \frac{d_0 + d_1 \omega^2}{d_2 + d_3 \omega^2}, \quad (\text{C21})$$

where

$$\begin{aligned} d_0 &= \phi_T^6 k_3^6 \lambda^4 (1+\lambda)^2 (x-1)^4, \\ d_1 &= \phi_T^4 k_3^4 K_M^2 \lambda^2 [1 + \lambda + 2\lambda(x-1)x]^2, \\ d_2 &= 4\phi_T^6 k_3^6 \lambda^4 (1+\lambda)(x-1)^4 \{1 + \lambda[1 + (x-1)x]\}, \quad \text{and} \\ d_3 &= 4\phi_T^4 k_3^4 K_M^2 \lambda^2 [1 + \lambda + \lambda(x-1)x]^2. \end{aligned} \quad (\text{C22})$$

Note that $d_i > 0$ for $i = 0, \dots, 3$, as well as that $d_0/d_2 = c_0/c_3$ in Eq. (C16); hence, the coherences predicted by the ssLNA and the full LNA agree exactly at $\omega = 0$. In the limit of large ω , the coherence of the ssLNA tends to the constant d_1/d_3 . Applying the limit of inefficient catalysis, with $\lambda \rightarrow 0$,

we obtain $d_1/d_3 \rightarrow 1/4$, which is the value predicted by the heuristic LNA in Eq. (C15). Hence, the ssLNA and the heuristic LNA agree for large frequencies provided that timescale separation conditions hold, as assumed by the ssLNA, with the additional assumption of inefficient catalysis.

3. General timescale separation conditions

Substituting $\vec{\phi} \mapsto \vec{\phi}^* + \vec{\epsilon}(t)$ into the general rate equations, given by Eq. (2), where $\vec{\phi}^*$ and $\vec{\epsilon}(t)$ are the steady-state concentration and the perturbation vector, respectively, and Taylor expanding, we obtain

$$\frac{d\vec{\epsilon}}{dt} = \underline{J}\vec{\epsilon},$$

where $\underline{J}$ is the constant Jacobian matrix obtained by evaluating Eq. (3) at the steady-state concentration $\vec{\phi}^*$. The corresponding solution in matrix exponential form is given by

$$\vec{\epsilon}(t) = \vec{\epsilon}_0 \exp(\underline{J}t), \quad (\text{C23})$$

where $\vec{\epsilon}_0$ is the vector of initial perturbations at time $t = 0$. It then follows that the total perturbation in each species will typically be a weighted sum of exponentials in time. Hence, for general initial conditions, it is difficult to derive analytical expressions for the corresponding timescales from Eq. (C23).

Instead, we adopt a numerical approach and define the timescale of species X as the last threshold-crossing time,

$$\tau_X := \max\{t \geq 0 : |\epsilon_X(t)| = |\epsilon_X(0)|/e\}. \quad (\text{C24})$$

Equivalently, τ_X is the maximum time it takes for the perturbation $\epsilon_X(t)$ to reach the value $|\epsilon_X(0)|/e$. Equation (C24) is equally valid for monotonically and nonmonotonically (oscillatorily) decreasing transients; for the latter, it is particularly useful, as it gives the time at which the threshold is last reached before the perturbation decays monotonically to zero.

To account for a range of perturbations given a fixed parameter set, we sample random initial conditions whose components are independent (i.i.d.) Gaussian random variables with zero mean and a standard deviation of 0.01. We then evaluate Eq. (C23) numerically to obtain, for each species X , the timescale τ_X , as defined by Eq. (C24). This procedure is repeated for 10^3 random initial conditions from which we calculate the distributions of the timescale ratios.

Applying the above procedure to the Michaelis-Menten model of enzyme catalysis with substrate synthesis, given by Eq. (26), with parameters $k_1 = 1$, $k_2 = 0.1$, $k_3 = 2$, $k_4 = 20$, $k_5 = 1$, and $\phi_T = 1$, as in Fig. 1, we find that the mean ratios

of the timescales are

$$\left\langle \frac{\tau_S}{\tau_C} \right\rangle = 287.5 \quad \text{and} \quad \left\langle \frac{\tau_P}{\tau_C} \right\rangle = 64.6. \quad (\text{C25})$$

The median of the distribution of the ratio of the substrate to complex timescales is 282.1; for the distribution of the ratio of product to complex timescales, it is 31.1.

Note that the alternative, heuristic identification of timescales via the expressions in Eq. (28), which assumes an

individual perturbation in one species only, results in the ratios $\tau_S/\tau_C = 244.9$ and $\tau_P/\tau_C = 23.3$, which are not far from the medians of the distributions obtained by our numerical procedure. Importantly, both predict that for the chosen parameter set, the substrate and product are the slow species, while the complex is the fast species. This example hence validates our use of Eq. (28) as a simple heuristic for choosing parameter sets with clearly separated timescales.

APPENDIX D: CALCULATIONS FOR SIMPLE MODEL OF TRANSLATION

The stoichiometric matrix and rate function vector for the reaction scheme in Eq. (37) are given by

$$\underline{S}_{\text{full}} = \begin{bmatrix} 1 & -1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 & -1 & 0 \end{bmatrix} \quad \text{and} \quad \vec{f}_{\text{full}} = [k_1, \quad k_2 \phi_M, \quad k_3 \phi_M, \quad k_4 C, \quad k_5 C, \quad k_6 \phi_P]^T.$$

Note that the element in the first row and j th column of the matrix $\underline{S}_{\text{full}}$ represents the net change in the number of molecules of M when the j th reaction, associated with reaction rate constant k_j , occurs. Similarly, the second and third rows describe the changes in the numbers of molecules of P and C , respectively.

For the heuristically reduced reaction scheme, given by Eq. (39), we can write

$$\underline{S}_{\text{heur}} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix} \quad \text{and} \quad \vec{f}_{\text{heur}} = [k_1, \quad k_2 \phi_M, \quad k' \phi_M, \quad k_6 \phi_P]^T.$$

The first and second rows of the stoichiometric matrix $\underline{S}_{\text{heur}}$ describe the changes in the numbers of molecules of M and P , respectively. The i th column is associated with the reaction whose rate is given by the i th element of $\vec{f}_{\text{heur}}$.

The rate equations, Jacobian and diffusion matrices, Lyapunov equations, power spectra, and coherence spectra for the full LNA can be constructed by replacing $\underline{S}$ and $\vec{f}$ by $\underline{S}_{\text{full}}$ and $\vec{f}_{\text{full}}$, respectively, in the equations of Sec. II. Similarly, the heuristic LNA can be constructed by replacing $\underline{S}$ by $\underline{S}_{\text{heur}}$ and $\vec{f}$ by $\vec{f}_{\text{heur}}$ in Sec. II.

For the ssLNA, the relevant matrices read

$$\underline{S}_s = \begin{bmatrix} 1 & -1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 \end{bmatrix} \quad \text{and} \quad \underline{S}_f = [0 \quad 0 \quad 1 \quad 0 \quad -1 \quad 0],$$

as well as

$$\underline{J}_{ss} = \begin{bmatrix} -k_2 - k_3 & 0 \\ 0 & -k_6 \end{bmatrix}, \quad \underline{J}_{sf} = \begin{bmatrix} k_5 \\ k_4 \end{bmatrix}, \quad \underline{J}_{fs} = [k_3 \quad 0], \quad \text{and} \quad \underline{J}_{ff} = -k_5.$$

Note that $\underline{S}_s$ and $\underline{S}_f$ correspond to the first two rows and to the last row, respectively, of the matrix $\underline{S}_{\text{full}}$. The matrices $\underline{J}_{ss}$, $\underline{J}_{sf}$, $\underline{J}_{fs}$, and $\underline{J}_{ff}$ are the block submatrices of the Jacobian matrix of the full system, given by Eq. (3), with $\underline{S}$ and $\vec{f}$ replaced by $\underline{S}_{\text{full}}$ and $\vec{f}_{\text{full}}$, respectively. Substituting the matrices $\underline{S}_s$, $\underline{S}_f$, $\underline{J}_{ss}$, $\underline{J}_{sf}$, $\underline{J}_{fs}$, and $\underline{J}_{ff}$ into the relevant equations of Sec. III B, one obtains the effective Jacobian and diffusion matrices from which one can construct and solve the equations for the covariance matrix, power spectra, and coherence spectra of the ssLNA.

Solving the Lyapunov equations of the full LNA, the heuristic LNA, and the ssLNA at the steady state, one obtains the following relationships between the corresponding second moments:

$$\frac{C_M^{\text{full}}}{C_M^{\text{heur}}} = 1, \quad \frac{C_{M,P}^{\text{full}}}{C_{M,P}^{\text{heur}}} = \frac{\beta + \gamma}{\beta(\alpha + \beta + \gamma + 1) + \gamma}, \quad (\text{D1})$$

$$\frac{C_P^{\text{full}}}{C_P^{\text{heur}}} = \frac{(\beta + \gamma)[\lambda(\alpha + \beta + \gamma) + \beta(\alpha + \beta + \gamma + 1) + \gamma]}{[\beta(\alpha + \beta + \gamma + 1) + \gamma](\alpha\lambda + \beta + \gamma)}, \quad (\text{D2})$$

$$\frac{C_M^{\text{full}}}{C_M^{\text{ssLNA}}} = 1, \quad \frac{C_{M,P}^{\text{full}}}{C_{M,P}^{\text{ssLNA}}} = \frac{\beta + \gamma}{\beta(\alpha + \beta + \gamma + 1) + \gamma}, \quad \text{and} \quad (\text{D3})$$

$$\frac{C_P^{\text{full}}}{C_P^{\text{ssLNA}}} = \frac{(\beta + \gamma)[\lambda(\alpha + \beta + \gamma) + \beta(\alpha + \beta + \gamma + 1) + \gamma]}{[\beta(\alpha + \beta + \gamma + 1) + \gamma][\lambda(\alpha + \beta + \gamma) + \beta + \gamma]}, \quad (\text{D4})$$

where $\alpha = k_3/k_5$, $\beta = k_6/k_5$, $\lambda = k_4/k_5$, and $\gamma = k_2/k_5$. Note that $\alpha + \gamma = \tau_C/\tau_M$ and $\beta = \tau_C/\tau_P$ represent the ratios of the timescale of the complex to the timescales of the mRNA and protein, respectively.

APPENDIX E: CALCULATIONS FOR MODEL OF TRANSLATION WITH COMPARTMENTALIZATION AND TRANSCRIPTIONAL BURSTING

The stoichiometric matrix and rate function vector for the reaction scheme in Eq. (50) are given by

$$\underline{S}_{\text{full}} = \begin{bmatrix} 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ r & -1 & -1 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \vec{f}_{\text{full}} = [k_1, \quad k_2 M_n, \quad k_3 M_n, \quad k_4 M_c, \quad k_5 M_c, \quad k_6 P]^T,$$

respectively. The element in the first row and j th column of $\underline{S}_{\text{full}}$ represents the net change in the number of molecules of M_c when the j th reaction, associated with reaction rate constant k_j , occurs. Similarly, the second and third rows describe the changes in the numbers of molecules of P and M_n , respectively. The integer r is sampled randomly from the geometric distribution with support on the non-negative integers and mean b .

For the reduced reaction scheme, given by Eq. (52), we have

$$\underline{S}_{\text{heur}} = \begin{bmatrix} r & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix} \quad \text{and} \quad \vec{f}_{\text{heur}} = [k', \quad k_4 M_c, \quad k_5 M_c, \quad k_6 P]^T,$$

where $k' = k_1 k_3 / (k_2 + k_3)$. The first and second rows of $\underline{S}_{\text{heur}}$ describe the changes in the numbers of molecules of M_c and P , respectively. The i th column is associated with the reaction whose rate is given by the i th element of $\vec{f}_{\text{heur}}$.

The rate equations, Jacobian and diffusion matrices, Lyapunov equations, power spectra, and coherence spectra for the full LNA can be constructed by replacing $\underline{S}$ by $\underline{S}_{\text{full}}$ and $\vec{f}$ by $\vec{f}_{\text{full}}$ in the equations of Sec. II. Similarly, the heuristic LNA can be constructed by replacing $\underline{S}$ and $\vec{f}$ by $\underline{S}_{\text{heur}}$ and $\vec{f}_{\text{heur}}$, respectively, in Sec. II. Note that wherever a factor of r^2 appears in the diffusion matrices, one should replace it by $\langle r^2 \rangle$, which is the mean square burst size calculated from the geometric distribution with mean burst size b : $\langle r^2 \rangle = \sum_{k=0}^{\infty} k^2 (1-p)^k p = b(1+2b)$, where $p = (1+b)^{-1}$; note also that $\langle r \rangle = b$. This averaging over powers of r in the diffusion matrices is necessary, given that r is not fixed but, rather, a randomly sampled integer [54].

For the ssLNA, the relevant matrices are given by

$$\underline{S}_s = \begin{bmatrix} 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix} \quad \text{and} \quad \underline{S}_f = [r \quad -1 \quad -1 \quad 0 \quad 0 \quad 0],$$

as well as by

$$\underline{J}_{ss} = \begin{bmatrix} -k_4 & 0 \\ k_5 & -k_6 \end{bmatrix}, \quad \underline{J}_{sf} = \begin{bmatrix} k_3 \\ 0 \end{bmatrix}, \quad \underline{J}_{fs} = [0 \quad 0], \quad \text{and} \quad \underline{J}_{ff} = -k_2 - k_3.$$

Note that $\underline{S}_s$ and $\underline{S}_f$ correspond to the first two rows and to the last row, respectively, of the matrix $\underline{S}_{\text{full}}$. The matrices $\underline{J}_{ss}$, $\underline{J}_{sf}$, $\underline{J}_{fs}$, and $\underline{J}_{ff}$ are the block submatrices of the Jacobian matrix of the full system, given by Eq. (3), with $\underline{S}$ and $\vec{f}$ replaced by $\underline{S}_{\text{full}}$ and $\vec{f}_{\text{full}}$, respectively. Substituting the matrices $\underline{S}_s$, $\underline{S}_f$, $\underline{J}_{ss}$, $\underline{J}_{sf}$, $\underline{J}_{fs}$, and $\underline{J}_{ff}$ into the relevant equations of Sec. III B, one obtains the effective Jacobian and diffusion matrices from which one can construct and solve the equations for the covariance matrix, power spectra, and coherence spectra of the ssLNA. As before, wherever a factor of r^2 appears in the diffusion matrices, one should replace it by $\langle r^2 \rangle$.

-
- [1] P. Thomas, R. Grima, and A. V. Straube, Rigorous elimination of fast stochastic variables from the linear noise approximation using projection operators, *Phys. Rev. E* **86**, 041110 (2012).
 - [2] J. K. Kim, K. Josić, and M. R. Bennett, The validity of quasi-steady-state approximations in discrete stochastic simulations, *Biophys. J.* **107**, 783 (2014).
 - [3] Y. M. Song, H. Hong, and J. K. Kim, Universally valid reduction of multiscale stochastic biochemical systems using simple nonelementary propensities, *PLoS Comput. Biol.* **17**, e1008952 (2021).
 - [4] N. Popović, C. Marr, and P. S. Swain, A geometric analysis of fast-slow models for stochastic gene expression, *J. Math. Biol.* **72**, 87 (2016).
 - [5] H.-W. Kang, W. R. KhudaBukhsh, H. Koepl, and G. A. Rempala, Quasi-steady-state approximations derived from the stochastic model of enzyme kinetics, *Bull. Math. Biol.* **81**, 1303 (2019).
 - [6] P. Thomas, N. Popović, and R. Grima, Phenotypic switching in gene regulatory networks, *Proc. Natl. Acad. Sci. USA* **111**, 6994 (2014).
 - [7] T. Jahnke, On reduced models for the chemical master equation, *Multiscale Model. Simul.* **9**, 1646 (2011).
 - [8] S. Smith, C. Cianci, and R. Grima, Model reduction for stochastic chemical systems with abundant species, *J. Chem. Phys.* **143**, 214105 (2015).
 - [9] A. Hellander and P. Lötstedt, Hybrid method for the chemical master equation, *J. Comput. Phys.* **227**, 100 (2007).
 - [10] Z. Cao and R. Grima, Linear mapping approximation of gene regulatory networks with stochastic dynamics, *Nat. Commun.* **9**, 3305 (2018).
 - [11] C. Jia and R. Grima, Holimap: An accurate and efficient method for solving stochastic gene network dynamics, *Nat. Commun.* **15**, 6557 (2024).

- [12] M. Barrio, A. Leier, and T. T. Marquez-Lago, Reduction of chemical reaction networks through delay distributions, *J. Chem. Phys.* **138**, 104114 (2013).
- [13] A. Leier, M. Barrio, and T. T. Marquez-Lago, Exact model reduction with delays: Closed-form distributions and extensions to fully bi-directional monomolecular reactions, *J. R. Soc. Interface* **11**, 20140108 (2014).
- [14] J. Zhang and T. Zhou, Markovian approaches to modeling intracellular reaction processes with molecular memory, *Proc. Natl. Acad. Sci. USA* **116**, 23542 (2019).
- [15] J. Szavits-Nossan and R. Grima, Solving stochastic gene-expression models using queueing theory: A tutorial review, *Biophys. J.* **123**, 1034 (2024).
- [16] K. Öcal, G. Sanguinetti, and R. Grima, Model reduction for the chemical master equation: An information-theoretic approach, *J. Chem. Phys.* **158**, 114113 (2023).
- [17] Y. Cao, D. T. Gillespie, and L. R. Petzold, The slow-scale stochastic simulation algorithm, *J. Chem. Phys.* **122**, 014116 (2005).
- [18] J. Goutsias, Quasiequilibrium approximation of fast reaction kinetics in stochastic biochemical systems, *J. Chem. Phys.* **122**, 184102 (2005).
- [19] C. V. Rao and A. P. Arkin, Stochastic chemical kinetics and the quasi-steady-state assumption: Application to the Gillespie algorithm, *J. Chem. Phys.* **118**, 4999 (2003).
- [20] D. T. Gillespie, Exact stochastic simulation of coupled chemical reactions, *J. Phys. Chem.* **81**, 2340 (1977).
- [21] P. Thomas, A. V. Straube, and R. Grima, The slow-scale linear noise approximation: An accurate, reduced stochastic description of biochemical networks under timescale separation conditions, *BMC Syst. Biol.* **6**, 39 (2012).
- [22] M. Ma, J. Szavits-Nossan, A. Singh, and R. Grima, Analysis of a detailed multistage model of stochastic gene expression using queueing theory and model reduction, *Math. Biosci.* **373**, 109204 (2024).
- [23] J. Elf and M. Ehrenberg, Fast evaluation of fluctuations in biochemical networks with the linear noise approximation, *Genome Res.* **13**, 2475 (2003).
- [24] E. Ziv, I. Nemenman, and C. H. Wiggins, Optimal signal processing in small stochastic biochemical networks, *PLoS ONE* **2**, e1077 (2007).
- [25] M. Rodríguez Martínez, J. Soriano, T. Tlustý, Y. Pilpel, and I. Furman, Messenger RNA fluctuations and regulatory RNAs shape the dynamics of a negative feedback loop, *Phys. Rev. E* **81**, 031924 (2010).
- [26] J. Holehouse and R. Grima, Revisiting the reduction of stochastic models of genetic feedback loops with fast promoter switching, *Biophys. J.* **117**, 1311 (2019).
- [27] F. Tostevin and P. R. ten Wolde, Mutual information between input and output trajectories of biochemical networks, *Phys. Rev. Lett.* **102**, 218101 (2009).
- [28] F. Tostevin and P. R. ten Wolde, Mutual information in time-varying biochemical systems, *Phys. Rev. E* **81**, 061917 (2010).
- [29] M. Meijers, S. Ito, and P. R. ten Wolde, Behavior of information flow near criticality, *Phys. Rev. E* **103**, L010102 (2021).
- [30] R. Fan and A. Hilfinger, Characterizing the nonmonotonic behavior of mutual information along biochemical reaction cascades, *Phys. Rev. E* **110**, 034309 (2024).
- [31] T. Munakata and M. Kamiyabu, Stochastic resonance in the Fitzhugh-Nagumo model from a dynamic mutual information point of view, *Eur. Phys. J. B* **53**, 239 (2006).
- [32] A. Komae, Mutual information rate between stationary Gaussian processes, *Results Appl. Math.* **7**, 100107 (2020).
- [33] P. P. Mitra and J. B. Stark, Nonlinear limits to the information capacity of optical fibre communications, *Nature (London)* **411**, 1027 (2001).
- [34] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier, Amsterdam, 1992), Vol. 1.
- [35] R. A. Horn and C. R. Johnson, *Matrix Analysis* (Cambridge University Press, Cambridge, 2012).
- [36] D. Schnoerr, G. Sanguinetti, and R. Grima, Approximation and inference methods for stochastic biochemical kinetics—A tutorial review, *J. Phys. A: Math. Theor.* **50**, 093001 (2017).
- [37] C. W. Gardiner *et al.*, *Handbook of Stochastic Methods* (Springer-Verlag, Berlin, 1985), Vol. 3.
- [38] M. Reinhardt, G. Tkačik, and P. R. ten Wolde, Path weight sampling: Exact Monte Carlo computation of the mutual information between stochastic trajectories, *Phys. Rev. X* **13**, 041017 (2023).
- [39] A.-L. Moor and C. Zechner, Dynamic information transfer in stochastic biochemical networks, *Phys. Rev. Res.* **5**, 013032 (2023).
- [40] M. Reinhardt, A. J. Tjalma, A.-L. Moor, C. Zechner, and P. R. Ten Wolde, Mutual information rate—linear noise approximation and exact computation, [arXiv:2508.21220](https://arxiv.org/abs/2508.21220).
- [41] P. Thomas, A. V. Straube, and R. Grima, Communication: limitations of the stochastic quasi-steady-state approximation in open biochemical reaction networks, *J. Chem. Phys.* **135**, 181103 (2011).
- [42] K. R. Sanft, D. T. Gillespie, and L. R. Petzold, Legitimacy of the stochastic Michaelis-Menten approximation, *IET Syst. Biol.* **5**, 58 (2011).
- [43] J. Janssen, The elimination of fast variables in complex chemical reactions. III. Mesoscopic level (irreducible case), *J. Stat. Phys.* **57**, 187 (1989).
- [44] P. D. Welch, The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms, *IEEE Trans. Audio Electroacoust.* **15**, 70 (1967).
- [45] Gillespie's SSA samples paths from the CME of the reduced system, given by Eq. (29), which has propensities $k_1\Omega$ for the reaction $\emptyset \xrightarrow{k_1} S$, $k'_S n_S = \frac{k_4\phi_T n_S}{\Omega K_M + n_S}$ for the reaction $S \xrightarrow{k'_S} P$, and $k_5 n_P$ for the reaction $P \xrightarrow{k_5} \emptyset$; here, Ω is the volume of the system and n_i denotes the number of molecules of species i .
- [46] T. Maier, M. Güell, and L. Serrano, Correlation of mRNA and protein in complex biological samples, *FEBS Lett.* **583**, 3966 (2009).
- [47] S. H. Payne, The utility of protein and mRNA correlation, *Trends Biochem. Sci.* **40**, 1 (2015).
- [48] C. Buccielli and M. Selbach, mRNAs, proteins and the emerging principles of gene expression control, *Nat. Rev. Genet.* **21**, 630 (2020).

- [49] D. Popovic, B. Koch, M. Kueblbeck, J. Ellenberg, and L. Pelkmans, Multivariate control of transcript to protein variability in single mammalian cells, *Cell Syst.* **7**, 398 (2018).
- [50] C. Jia, Simplification of Markov chains with infinite state space and the mathematical theory of random gene expression bursts, *Phys. Rev. E* **96**, 032402 (2017).
- [51] I. Golding, J. Paulsson, S. M. Zawilski, and E. C. Cox, Real-time kinetics of gene activity in individual bacteria, *Cell* **123**, 1025 (2005).
- [52] Z. Cao and R. Grima, Analytical distributions for detailed models of stochastic gene expression in eukaryotic cells, *Proc. Natl. Acad. Sci. USA* **117**, 4682 (2020).
- [53] A.-L. Moor, A. Tjalma, M. Reinhardt, P. R. Ten Wolde, and C. Zechner, State-versus reaction-based information processing in biochemical networks, [arXiv:2505.13373](https://arxiv.org/abs/2505.13373).
- [54] R. Ramaswamy, N. González-Segredo, I. F. Sbalzarini, and R. Grima, Discreteness-induced concentration inversion in mesoscopic chemical systems, *Nat. Commun.* **3**, 779 (2012).