

# Estimations of intrinsic and extrinsic noise in models of nonlinear genetic networks

Matthew Scott<sup>a)</sup>

Center for Theoretical Biological Physics, Department of Physics, University of California, San Diego, La Jolla, California, 92093-0374

Brian Ingalls<sup>b)</sup>

Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1

Mads Kærn<sup>c)</sup>

Ottawa Institute of Systems Biology, Department of Cellular and Molecular Medicine, Department of Physics, University of Ottawa, 451 Smyth Road, Ottawa, Ontario, Canada K1H 8M5

(Received 15 February 2006; accepted 17 May 2006; published online 30 June 2006)

We discuss two methods that can be used to estimate the impact of internal and external variability on nonlinear systems, and demonstrate their utility by comparing two experimentally implemented oscillatory genetic networks with different designs. The methods allow for rapid estimations of intrinsic and extrinsic noise and should prove useful in the analysis of natural genetic networks and when constructing synthetic gene regulatory systems. © 2006 American Institute of Physics.

[DOI: [10.1063/1.2211787](https://doi.org/10.1063/1.2211787)]

**The expression of genes is a highly controlled process that involves elaborate transcriptional regulatory programs encoded in extensive protein-protein and protein-DNA interaction networks. It is commonly held that the relationship between genotype, environmental cues, and cellular phenotype is deterministic and that transcriptional regulatory programs operate with high precision to ensure gene activation and repression in the appropriate temporal order. However, measurements of gene expression at the single cell level have demonstrated significant variability among isogenic cells experiencing identical environments, challenging the general validity of these assumptions. Indeed, it is becoming increasingly clear that gene expression is quite a noisy process, and that variation among supposedly identical individuals may play a more prominent role in key areas like development and disease than previously anticipated. Brute-force stochastic simulations may not be feasible for large-scale systems or in systematic analysis, which call for more efficient methods for estimating the impact of intrinsic and extrinsic variability on system function.**

## I. INTRODUCTION

Recent years have seen remarkable progress in our understanding of the origins and consequences of stochasticity in gene expression.<sup>1-4</sup> These advances have been driven by theoretical advances enabling biological hypotheses formulation using stochastic process and dynamical systems theory, and experimental advances enabling measurements of gene expression at the single cell level. Noticeably, it has been confirmed experimentally that the precision of gene expres-

sion is limited by the fundamental statistical-mechanical  $1/N$  scaling of relative fluctuation amplitude (i.e., the noise) and abundance  $N$ ,<sup>5</sup> and that population variability in the expression of single genes depends on DNA-encoded parameters in accordance with relationships predicted by stochastic theory.<sup>6-8</sup> Additionally, research strategies combining experimental and theoretical approaches are beginning to illuminate how noisy gene regulatory signals propagate through signaling pathways, transcriptional cascades, and feedback loops, and how gene expression noise can lead to the emergence and switching between distinct expression states.<sup>7,9-18</sup>

The impact of stochasticity on the functioning of gene networks has also been the subject of considerable theoretical investigations, particularly in the context of design principles that can ensure the maintenance of network function despite internal or external perturbations.<sup>19,20</sup> Some of the most intriguing subjects of this research are networks able to adapt, as in bacterial chemotaxis, or support oscillatory gene expression (see, e.g., Refs. 21-43). For oscillatory systems, a comprehensive analysis recently demonstrated that some network designs are more robust than others.<sup>39</sup> Additionally, two synthetic oscillatory networks with different architectures, the Repressilator<sup>25</sup> and the Atkinson oscillator,<sup>31</sup> show marked differences in their ability to generate and maintain oscillations when implemented experimentally. *Escherichia coli* cells harboring the Repressilator network show a few noisy oscillations that desynchronize on a short time scale, while *E. coli* cells harboring the Atkinson network maintain population-synchronous (damped) oscillations over a long time scale. While the design of the Atkinson oscillator is related to that of a noise resilient minimal model of circadian oscillations,<sup>26,44</sup> it is unclear if the differences in apparent robustness observed experimentally are due to differences in their network design.

<sup>a)</sup>Electronic mail: [mScott@ctbp.ucsd.edu](mailto:mScott@ctbp.ucsd.edu)

<sup>b)</sup>Electronic mail: [bingalls@math.uWaterloo.ca](mailto:bingalls@math.uWaterloo.ca)

<sup>c)</sup>Electronic mail: [mkaern@uOttawa.ca](mailto:mkaern@uOttawa.ca)

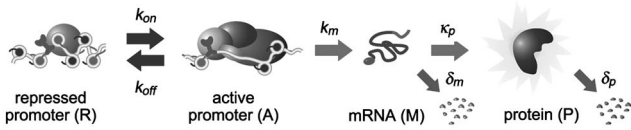


FIG. 1. A simple model for the expression of a single gene. All steps correspond to first-order reactions with the indicated rate constants (units of inverse time).

Despite significant progress made within the past few years, there is still a clear need for the development of methodologies that can be used to determine how different network designs behave in the presence of substantial internal and external perturbations. It is in general not possible to obtain exact solutions for nonlinear stochastic systems, and relying solely on brute force numerical simulation is in most cases prohibitively time consuming. Here, we discuss two methods that can be used to estimate the impact of internal and external variability on nonlinear oscillatory systems, and demonstrate their utility in a limited comparative investigation of the Repressilator and Atkinson networks.

This paper is organized as follows: In Sec. II, we review and discuss how internal and external perturbations and variability impact the expression of a single gene, and introduce the basic terminologies and methods that will be used in the subsequent sections. In Sec. III, we discuss the linear noise approximation (LNA),<sup>45,46</sup> which, as pointed out by others,<sup>47-49</sup> provides a method for rapid assessment of the impact of intrinsic noise on genetic networks. In Sec. IV, we extend the static noise approximation (SNA) introduced in Sec. II and the LNA introduced in Sec. III to a nonlinear oscillatory model system. The comparative analysis of the two genetic networks is done in Sec. V. While this analysis relies on approximate methods and is not exhaustive, it indicates that the Repressilator is more resilient to internal and external perturbations than the Atkinson oscillator when both systems exhibit a macroscopic limit cycle. The differences in apparent robustness observed experimentally may be due to differences in experimental strategies or in the excitability of the Atkinson oscillator.

## II. STOCHASTICITY IN THE EXPRESSION OF A SINGLE GENE

To start our discussion, we begin by considering how internal (intrinsic) and external (extrinsic) perturbations affect the expression of a single gene. A simple model of this process is given in Fig. 1. It involves the synthesis of mRNA from a single DNA-encoded gene template, the synthesis of protein from mRNA templates, and the decay of mRNA and protein molecules. Additionally, the DNA template switches between transcriptionally active and inactive states at rates that depend on the binding of transcriptional regulators to the promoter region of the gene.

In a deterministic description that ignores stochastic effects, the cellular mRNA and protein concentrations  $m$  and  $p$  are governed by the macroscopic rate equations

$$\frac{dm}{dt} = \kappa_M - \delta_M m, \quad \frac{dp}{dt} = \kappa_P m - \delta_P p, \tag{1}$$

where  $\kappa_M$  is the effective synthesis rate of mRNA molecules. This rate is proportional to the average probability of the promoter being in the active state ( $k_{on}/[k_{on}+k_{off}]$ ). The proportionality factor is the rate constant  $k_M$  associated with transcription from the active promoter (i.e.,  $\kappa_M = k_M k_{on}/[k_{on}+k_{off}]$ ). The steady state concentrations  $m^s$  and  $p^s$  are accordingly

$$m^s = \frac{\kappa_M}{\delta_M}, \quad p^s = \frac{\kappa_M \kappa_P}{\delta_M \delta_P}, \tag{2}$$

which, by definition, are related to the average steady state number of mRNA and protein molecules  $\langle \bar{m} \rangle^s$  and  $\langle \bar{p} \rangle^s$  through the system size  $\Omega$  (the volume) by the relationships  $m = \langle \bar{m} \rangle / \Omega$  and  $p = \langle \bar{p} \rangle / \Omega$ , respectively.

### A. Fluctuations in extrinsic factors

Extrinsic noise is generally defined as fluctuations and variability that arise in a system due to disturbances originating from its environment and thus depends on how the system of interest is defined. Examples of extrinsic sources of noise in the expression of a single gene include fluctuations in transcriptional regulatory signals, varying efficiencies of the transcriptional and translational machinery as well as variation in gene copy number and population dynamics. While these and other factors can be incorporated into more comprehensive descriptions of the gene expression process,<sup>50</sup> it is often the case that insufficient information is available to establish a meaningful level of quantitative detail. Consequently, the parametric variations that cause extrinsic noise are often modeled generically based on a minimal number of assumptions.

Here we focus on perhaps the simplest case of *static* external perturbations. To clearly separate noise due to variability in external factors from the noise that is intrinsic to the system, the impact of parametric variability is investigated in the macroscopic limit governed by deterministic rate equations. In this limit, we evaluate the steady state extrinsic noise by calculating how external perturbations in system parameters affect the steady state protein concentration. In later sections, this method, which we refer to as the static noise approximation (SNA), is extended to the more general case of concentrations that change in time. To our knowledge this method has not been previously used, but it is simply one of many ways to formalize the standard practise of investigating robustness by determining the effect of perturbations in individual parameters.

To employ the SNA, we begin by selecting a parameter  $\rho$  of interest and setting it to a fixed nominal value  $\rho = \rho_0$  where the steady state is characterized by  $p_{\rho_0}^s = p^s(\rho_0)$ . We then select random numbers  $\rho_i$  (corresponding, e.g., to individual cells) from a normal distribution with mean  $\rho_0$  and standard deviation  $\sigma_\rho$  such that the parametric noise perturbing the system is given by  $\eta_\rho = \sigma_\rho / \rho_0$ .

The steady state protein abundance corresponding to sample  $\rho_i$  is denoted by  $p_i^s = p^s(\rho_i)$  and gives rise to a distri-

bution with average  $\langle p \rangle^s$  and standard deviation  $\sigma_p$ . The relative strength of this induced variation is the extrinsic noise  $\eta_{\text{ext}} = \sigma_p / \langle p \rangle^s$ . This variation is related to the parametric noise through the extrinsic noise susceptibility  $\chi_\rho^s$ , which measures the amplification or attenuation by the system of parametric noise in  $\rho$ . In the case of protein abundances, the steady state extrinsic noise susceptibility  $\chi_\rho^s$  is given by

$$\chi_\rho^s = \frac{\eta_{\text{ext}}}{\eta_\rho} = \frac{\rho_0 \sigma_p}{\langle p \rangle^s \sigma_\rho} \tag{3}$$

This general result is only restricted by requiring a constrained width of the distribution in  $\rho$  to avoid sampling unphysical (e.g., negative) parameter values, or values for which the system no longer admits a steady state.

When the distribution in  $\rho$  is sufficiently narrow, the resulting perturbations  $\delta_i$  are typically *small*, and so the resulting change in steady state protein concentration can be approximated by a truncated Taylor series

$$p^s(\rho_0 + \delta_i) \approx p^s(\rho_0) + \delta_i S_{p,\rho}, \tag{4}$$

where  $S_{p,\rho}$  is the sensitivity coefficient of the steady state  $p^s$  with respect to  $\rho$  evaluated at  $\rho = \rho_0$

$$S_{p,\rho} = \left. \frac{\partial p^s(\rho)}{\partial \rho} \right|_{\rho=\rho_0} \tag{5}$$

In this *linear* regime of the SNA, the average steady state  $\langle p \rangle^s$  is equal to the macroscopic steady state  $p_{\rho_0}^s$  and it follows immediately that the extrinsic noise is given by

$$\eta_{\text{ext}}^2 = \frac{1}{(p_{\rho_0}^s)^2} \frac{1}{N} \sum_i [p_i^s - p_{\rho_0}^s]^2 = \eta_\rho^2 \left( \frac{\rho_0}{p_{\rho_0}^s} S_{p,\rho} \right)^2, \tag{6}$$

where the sum is taken over  $N$  samples. Consequently, the steady state extrinsic noise susceptibility takes the form

$$\chi_\rho^s = \frac{\rho_0}{p_{\rho_0}^s} |S_{p,\rho}| = \left| \frac{\partial \ln p_{\rho_0}^s}{\partial \ln \rho} \right|, \tag{7}$$

which is the relative sensitivity coefficient or logarithmic gain. The logarithmic gain is used commonly in biochemical control theory,<sup>51,52</sup> and also appears in several studies of genetic networks in the context of propagation of noise between distinct processes (see, e.g., Refs. 2 and 13).

As a simple example of the *linear* SNA method, consider the steady state protein concentration defined in Eq. (2). This steady state depends linearly on the rates of mRNA and protein synthesis,  $\kappa_M$  and  $\kappa_p$  and is associated with an extrinsic noise susceptibility of unity for each. Hence, parametric noise of magnitude  $\eta_\rho$  in each of these parameter yields a steady state variability characterized by  $\eta_{\text{ext}} = \eta_\rho$ , and is neither attenuated nor amplified by the system. The cases of parametric noise in the decay rates  $\delta_M$  and  $\delta_p$  are less trivial since the effect on the steady state is poorly approximated by Eq. (4). The use of the SNA in the nonlinear regime will be discussed in Sec. IV A 1.

Intuitive cases where amplification or attenuation of parametric noise arise, for example, when a regulatory signal modifies the effective rate of mRNA synthesis  $\kappa_M$  in a nonlinear fashion. In the genetic oscillators discussed in Sec. V,

the dependence of mRNA synthesis rate on an activating signal  $a$  and a repressing signal  $r$  will be modeled using Hill-type kinetics and the phenomenological relationships

$$\kappa_M(a) = \frac{k_M a^n}{1 + a^n}, \quad \kappa_M(r) = \frac{k_M}{1 + r^n}, \tag{8}$$

where  $n$  is the Hill coefficient. With these definitions, evaluation of Eq. (7) yields the steady state noise susceptibilities

$$\chi_a^s = \frac{n}{1 + a^n}, \quad \chi_r^s = \frac{n \cdot r^n}{1 + r^n}. \tag{9}$$

Hence, for fixed parametric variability, a weak activating signal amplifies the parametric noise ( $\chi_a^s > 1$  for  $a < \sqrt[n]{n-1}$ ) while a strong activating signal attenuates it ( $\chi_a^s < 1$  for  $a > \sqrt[n]{n-1}$ ). The reverse holds for a repressing signal. In both cases, the impact of parametric noise is minimized when the relative expression level is increased. For  $n=2$ , relative expression levels below 50% ( $a < 1$  or  $R > 1$ ) yield amplification while expression levels above 50% yield attenuation. For  $n=3$  and  $n=4$ , attenuation only occurs above 67% and 75% expression, respectively. Taken together with the dependence of  $\chi^s$  on the Hill coefficient  $n$ , these results indicate that highly nonlinear gene regulation are particularly susceptible to variability in transcription factor activities.

In reality, transcriptional regulatory signals, as well as the other system parameters, may be characterized poorly by the SNA. The only way to thoroughly investigate how extrinsic noise impacts a given system is to establish a sufficiently detailed model of the external process. In this case, however, the process becomes part of the system of interest and the investigation turns to the propagation of intrinsic noise. It is therefore necessary to employ tools that can be used to estimate how intrinsic noise affects nonlinear biochemical reaction systems (see also the contribution by Hasty *et al.*, in this issue of Chaos). Before discussing such a tool in Sec. III, we briefly review how an exact method for linear systems has been employed in the context of single gene expression.

### B. Intrinsic gene expression noise

The deterministic model in Eq. (1) is obtained from the qualitative picture in Fig. 1 by treating each step as an elementary chemical reaction and applying the law of mass action. The law of mass action was developed to describe chemical reactions under conditions where the number of each chemical species is so large that concentrations can be approximated as continuous variables without introducing significant error. In this macroscopic limit, the change in concentration following the addition or removal of a few molecules is imperceptible.

In reality, all molecular systems are composed of discrete states characterized by having  $n = n_1, n_2, \dots, n_d$  molecules of species  $s_1, s_2, \dots, s_d$ . A chemical reaction introduces a transition from the current state  $\mathbf{n}$  to a new state  $\mathbf{n}'$  that depends on the molecules generated and destroyed in that reaction. The probability that the  $\mathbf{n} \rightarrow \mathbf{n}'$  transition will take place at a given time is in turn proportional to the prob-

ability  $P(\mathbf{n}, t)$  that the system is in state  $\mathbf{n}$  at time  $t$ . This can be used to write an equation for the rate of change in the probability  $P(\mathbf{n}, t)$

$$\frac{\partial P(\mathbf{n}, t)}{\partial t} = \sum_{\mathbf{n}'} W_{\mathbf{n}' \rightarrow \mathbf{n}} P(\mathbf{n}', t) - W_{\mathbf{n} \rightarrow \mathbf{n}'} P(\mathbf{n}, t), \quad (10)$$

where  $W_{\mathbf{n}' \rightarrow \mathbf{n}}$  are the transition probabilities for reaction events moving the system into state  $\mathbf{n}$  from a neighboring state  $\mathbf{n}'$ , and  $W_{\mathbf{n} \rightarrow \mathbf{n}'}$  are the transition probabilities for reaction events moving the system out of the state  $\mathbf{n}$  to a neighboring state  $\mathbf{n}'$ . The discrete transition probabilities can in turn be expressed in terms of a step operator,  $\mathbf{E}_i^k$ . The step operator describes the addition or removal of  $k$  molecules of species  $i$  when a particular reaction occurs. For a function  $f(n_i, n_j)$  with two integer arguments,  $\mathbf{E}_i^k$  increments  $n_i$  by an integer  $k$

$$\mathbf{E}_i^k f(n_i, n_j) = f(n_i + k, n_j). \quad (11)$$

The rate equation for the probability density  $P(\mathbf{n}, t)$  in Eq. (10) is called the chemical Master equation.<sup>53</sup> The first and second moments of  $P(\mathbf{n}, t)$  with respect to a species  $i$  are the average number  $\langle n_i \rangle(t)$  of that molecule and variance  $\langle n_i^2 \rangle - \langle n_i \rangle^2 = \sigma_i^2(t)$ , respectively. The intrinsic noise, defined by  $\eta_{\text{int}}(t) = \sigma_i(t) / \langle n_i \rangle(t)$  for species  $i$ , is thus directly related to moments of  $P(\mathbf{n}, t)$ .

To construct the Master equation associated with the model of gene expression shown in Fig. 1, consider first the change in the system state due to degradation of mRNA. With  $\bar{a}$ ,  $\bar{m}$ , and  $\bar{p}$  denoting the number of active promoters, mRNA, and protein molecules, respectively, the change in the probability is given by

$$\begin{aligned} \frac{dP(\bar{a}, \bar{m}, \bar{p}, t)}{dt} &= \delta_M (\bar{m} + 1) P(\bar{a}, \bar{m} + 1, \bar{p}, t) \\ &\quad - \delta_M \bar{m} P(\bar{a}, \bar{m}, \bar{p}, t), \end{aligned} \quad (12)$$

where the first and second term describes the flux in and out of state  $\{\bar{a}, \bar{m}, \bar{p}, t\}$  due to the removal of one mRNA. With the use of the step operator  $\mathbf{E}_i^k$ , Eq. (12) can be presented in the compact form

$$\frac{dP}{dt} = \delta_M (\mathbf{E}_m^1 - 1) \bar{m} P, \quad (13)$$

where the probability density  $P = P(\bar{a}, \bar{m}, \bar{p}, t)$  is introduced for clarity. For a single promoter, the contributions to the probability flux due to promoter binding, transcription, protein synthesis, and protein degradation are obtained in a similar way yielding the Master equation

$$\begin{aligned} \frac{dP}{dt} &= k_{\text{on}} (\mathbf{E}_a^{-1} - 1) (1 - \bar{a}) P + k_{\text{off}} (\mathbf{E}_a^1 - 1) \bar{a} P \\ &\quad + k_M (\mathbf{E}_m^{-1} - 1) \bar{a} P + \kappa_P (\mathbf{E}_p^{-1} - 1) \bar{m} P \\ &\quad + \delta_M (\mathbf{E}_m^1 - 1) \bar{m} P + \delta_P (\mathbf{E}_p^1 - 1) \bar{p} P. \end{aligned} \quad (14)$$

Because the Master equation is linear in the state variables  $\bar{a}$ ,  $\bar{m}$ , and  $\bar{p}$ , it is relatively easy to calculate the moments of the probability distribution using the so-called mo-

ment functions.<sup>54</sup> With this approach, average steady state numbers of mRNA and protein molecules (the first moments of  $P$ ) are given by

$$\langle \bar{m} \rangle^s = \langle \bar{a} \rangle^s \frac{k_M}{\delta_M}, \quad \langle \bar{p} \rangle^s = \langle \bar{m} \rangle^s \frac{\kappa_P}{\delta_P}, \quad (15)$$

where  $\langle \bar{a} \rangle^s = k_{\text{on}} / (k_{\text{on}} + k_{\text{off}})$  is the average probability of the promoter being active. The analysis also yields the variances in  $\bar{a}$ ,  $\bar{m}$ , and  $\bar{p}$  as the second moments of  $P$  and an expression for the intrinsic noise in the steady state protein abundance containing terms arising from protein, mRNA, and promoter kinetics. Using the results obtained in Ref. 8 and assuming that protein decay is slow compared to the promoter kinetics ( $\delta_P \ll k_{\text{on}} + k_{\text{off}}$ ), the intrinsic noise in steady state protein concentration can be written as

$$\eta_{\text{int}}^2 = \underbrace{\frac{1}{\Omega} \left( \frac{1}{p^s} + \frac{1}{1 + \phi m^s} \right)}_{\text{mRNA and protein kinetics}} + \underbrace{\frac{\delta_M k_{\text{off}} / k_{\text{on}}}{(1 + \phi)(k_{\text{on}} + k_{\text{off}})}}_{\text{promoter kinetics}} \quad (16)$$

where  $\phi = \delta_M / \delta_P$ .

It is clear from Eq. (16) that there are three ways to suppress the impact of fluctuations arising from mRNA and protein kinetics; (1) large system size  $\Omega$ , (2) high mRNA and protein concentrations, and (3) rapid mRNA decay. Rapid mRNA decay ( $\delta_M \gg \delta_P$ ) suppresses the impact of low mRNA concentration on protein fluctuations through time averaging.<sup>2</sup> This suppression essentially comes about when the mRNA abundance fluctuates so rapidly about its mean that the slower downstream variable (i.e., protein abundance) respond only to changes in the mean.

Regardless of this time averaging, intrinsic noise can be reduced by increasing the expression level, which increases the concentrations of both mRNA and protein. When the system size  $\Omega$  is fixed, variation in expression level yields an inverse scaling relationship  $\eta_{\text{int}}^2 \sim 1/\bar{p}$  between the intrinsic noise and the abundance of the expressed protein. As mentioned in the Introduction, this relationship has been confirmed experimentally.<sup>25</sup>

The relationship between system size (i.e., volume) and intrinsic noise has been studied extensively in the context of chemical reaction systems. In such systems, for fixed concentrations the intrinsic noise generally becomes negligible when the volume in which the reactions take place is sufficiently large (the macroscopic or deterministic limit). This is because relative fluctuation amplitude generally scales with the number of molecules as  $\eta_{\text{int}} \sim 1/\sqrt{\Omega}$  regardless of the concentrations of the involved species, as captured by the contribution from mRNA and protein kinetics in Eq. (16).

In the sections to follow, we will use decreased system size as a means of evaluating the impact of intrinsic noise on the function of gene regulatory networks. However, it is worth noting that kinetics at promoters (whose copy number remain low regardless of system size) may generate additional variability in the processes (see, e.g., Refs. 3, 7, 8, and 55). Indeed, experiments in *Saccharomyces cerevisiae* have demonstrated effects that are consistent with slow promoter kinetics in Eq. (16).<sup>8,56</sup> While the topic will not be discussed in further detail, it is important to recognize that gene expres-

sion systems involving slow promoter kinetics may display significant deviations from the fundamental statistical-mechanical system-size scaling predicted for chemical reaction systems (see also Ref. 55).

### III. LINEAR NOISE APPROXIMATION

Unfortunately, it is in most cases not possible to solve the Master equation analytically when the reaction system of interest involves nonlinear steps and feedback loops. In such cases, the impact of intrinsic noise on system function require approximate methods or brute force numerical simulations. Two frequently employed approximate methods are the Langevin approach and the LNA. Both methods rely upon asymptotic expansion in a small parameter—small excursions from a steady state in the Langevin approach,<sup>57</sup> and small deviations from the macroscopic trajectory in the LNA.<sup>46</sup>

The LNA provides a convenient method for calculating the intrinsic fluctuations in systems that display nonsteady behavior, such as oscillations,<sup>58</sup> and is the most suitable approach in this study. The approximation involves the smoothing of the discrete state space (the number of molecules) into a continuum (the macroscopic concentrations) and the calculation of additional terms describing the fluctuations about the macroscopic trajectory. In other words, the number of molecules  $n_i(t)$  of a species  $i$  is approximated as a continuous variable in terms of the concentration  $c_i(t)$  and a term  $\alpha_i(t)$  describing the deviation from  $c_i(t)$ . Because the impact of this fluctuation is expected to scale with the square root of the system size  $\Omega$ , the number of molecules is approximated by setting

$$n_i(t) = \Omega c_i(t) + \sqrt{\Omega} \alpha_i(t). \quad (17)$$

In addition to the suitability of the ansatz in Eq. (17), the LNA assumes that the step operator [Eq. (11)] associated with the transition probabilities in the Master equation is well described by a Taylor series for large  $\Omega$ . This is reasonable since an integer change in molecule numbers  $n_i$  has infinitesimal effect on the concentration  $c_i$  in the macroscopic limit. In terms of Eq. (17), adding  $k$  molecules increases the value of  $\alpha_i$  by  $k/\sqrt{\Omega}$ . Therefore, for large system size, the completion of a single reaction brings about a small concentration change and the operator  $E_i^k$  is well approximated by a truncated Taylor series

$$E_i^k \approx \left[ 1 + \frac{k}{\sqrt{\Omega}} \partial_i + \frac{k^2}{2\Omega} \partial_i^2 \right], \quad (18)$$

where  $\partial_i$  and  $\partial_i^2$  denote  $\partial/(\partial\alpha_i)$  and  $\partial^2/(\partial\alpha_i^2)$ , respectively.

To calculate the average fluctuations  $\langle\alpha_i\rangle$  and the variances  $\langle\alpha_i\alpha_j\rangle$ , it is necessary to derive an equation that describes the fluctuations  $\alpha(t) = \alpha_1, \dots, \alpha_d$ , where  $d$  is the number of distinct species in the system. This is done by introducing the probability density  $\Pi = \Pi(\alpha, t) = \Omega^{d/2} P(\mathbf{n}, t)$  where the scaling  $\Omega^{d/2}$  ensures that  $\Pi$  is appropriately normalized. The next step is to obtain an approximation of the equation that governs the evolution of  $\Pi$ . This is done by inserting Eqs. (17) and (18) into the Master equation, followed by collecting terms in which  $\Omega$  appears with the same

order (see Ref. 46 for details). Terms that are independent of  $\Omega$  constitute the zeroth order approximation (since  $\Omega^0=1$ ), corresponding to the macroscopic rate equations

$$\frac{d\mathbf{c}}{dt} = \mathbf{f}(\mathbf{c}). \quad (19)$$

The equation that governs the evolution of  $\Pi(\alpha, t)$  contains terms multiplied by  $1/\sqrt{\Omega}$ . Collecting these terms yields the general form of a linear Fokker-Planck equation

$$\frac{\partial\Pi}{\partial t} = - \sum_{i,j} A_{ij} \partial_i (\alpha_j \Pi) + \frac{1}{2} \sum_{i,j} B_{ij} \partial_i \partial_j \Pi, \quad (20)$$

where  $A_{ij}$  are called *drift* or *dissipation* terms and  $B_{ij}$  are called the *diffusion* terms. The matrix  $A$  is given by the Jacobian evaluated pointwise along the trajectory

$$A_{ij}(t) = \left. \frac{\partial f_i}{\partial c_j} \right|_{\mathbf{c}(t)}, \quad (21)$$

and reflects the *local stability* of the macroscopic system to small perturbations at the point  $\mathbf{c}(t)$ . The matrix  $B$  describes the strength of the local fluctuations and is a bit more challenging to calculate (see Appendix A). It is noted that the time-dependent coefficient matrices  $A$  and  $B$  are *independent* of the fluctuations  $\alpha$ , which appear only linearly in Eq. (20). This ensures that the solution  $P(\mathbf{n}, t)$  conditioned by the initial state  $P(\mathbf{n}, 0) = \delta(\mathbf{n} - \mathbf{n}_0)$  is normal and completely characterized by the first two moments.

Equations describing the evolution equations of the means  $\langle\alpha_i\rangle$  and the variances  $\langle(\alpha_i - \langle\alpha_i\rangle)(\alpha_j - \langle\alpha_j\rangle)\rangle$  are obtained by multiplying Eq. (20) by  $\alpha_i$  and integrating over all  $\alpha = \alpha_1, \dots, \alpha_d$ . For the means, one obtains the equation

$$\frac{d}{dt} \langle\alpha\rangle = A \langle\alpha\rangle. \quad (22)$$

Notice that if the initial state is known precisely so that  $\langle\alpha(0)\rangle = 0$ , then the mean will remain zero for all time. For simplicity, we shall always assume that  $\langle\alpha\rangle = 0$ , so that for the variances  $C_{ij} = \langle\alpha_i\alpha_j\rangle$ , one obtains a set of equations that in compact matrix-vector form are given by

$$\frac{dC}{dt} = AC + CA^T + B. \quad (23)$$

Because the probability distribution  $P(\mathbf{n}, t)$  is completely characterized by its first two moments, we may further write the solution  $\Pi(\alpha, t)$  as a  $d$ -dimensional Gaussian probability distribution spanned by the fluctuation variables  $\alpha$

$$\Pi(\alpha, t) = \frac{1}{\sqrt{(2\pi)^d \det[C]}} \exp \left[ -\frac{1}{2} \alpha^T C^{-1} \alpha \right]. \quad (24)$$

The fluctuations away from the macroscopic trajectory are given by the contours of  $\Pi$ , which are referred to as the error curves. Choosing the number  $K$  as the probability contained within the contour, the equation for the corresponding error curve is given by

$$\frac{1}{2}\boldsymbol{\alpha}^T\mathbf{C}^{-1}\boldsymbol{\alpha}=K. \quad (25)$$

This equation provides a useful tool for visualizing variances along the macroscopic trajectory in higher dimensions.

Since the Fokker-Planck equation in Eq. (20) and the error curves in Eq. (25) employ a continuous approximation of the number of molecules, the last step is to convert to a description in terms of fluctuating concentrations  $\tilde{\alpha}_i = \alpha_i/\sqrt{\Omega}$  and find an equation for corresponding cross-correlation matrix  $\tilde{\mathbf{C}} = \langle \tilde{\boldsymbol{\alpha}}\tilde{\boldsymbol{\alpha}}^T \rangle$ . This is done by substituting Eq. (17) into Eq. (23)

$$\frac{d\langle \tilde{\boldsymbol{\alpha}}\tilde{\boldsymbol{\alpha}}^T \rangle}{dt} = \mathbf{A}\langle \tilde{\boldsymbol{\alpha}}\tilde{\boldsymbol{\alpha}}^T \rangle + \langle \tilde{\boldsymbol{\alpha}}\tilde{\boldsymbol{\alpha}}^T \rangle\mathbf{A}^T + \frac{\mathbf{B}}{\Omega}. \quad (26)$$

It is interesting to note that some basic features manifest directly in this equation. First, the internal perturbations to the system *variables* are brought about by a diffusion term ( $\mathbf{B}/\Omega$ ), that diminishes as the system size increases. It would thus be expected that a system with a design and parameter values that yields high values of  $\mathbf{B}$  is more sensitive to a decrease in system size than a system characterized by low values of  $\mathbf{B}$ .<sup>47</sup> Secondly, since  $\mathbf{A}$  gives the sensitivity of the system to perturbations in the variables, a system that is more *dissipative* (i.e., large, all negative eigenvalues of  $\mathbf{A}$ ), is more efficient in dampening small excursions away from the macroscopic trajectory yielding lower intrinsic noise.<sup>59</sup>

### A. Intrinsic noise in the expression of a single gene

To illustrate the use of the LNA, consider the simple model of gene expression in Fig. 1. When mRNA synthesis is associated with a constant rate  $\kappa_M$ , the Master equation is given by [cf. Eq. (14)],

$$\begin{aligned} \frac{dP}{dt} = & \kappa_M\Omega(\mathbf{E}_m^{-1} - 1)P + \kappa_P(\mathbf{E}_p^{-1} - 1)\bar{m}P + \delta_M(\mathbf{E}_m^1 - 1)\bar{m}P \\ & + \delta_P(\mathbf{E}_p^1 - 1)\bar{p}P. \end{aligned} \quad (27)$$

The Master equation is then expressed in terms of  $\Pi = \Pi(\alpha_1, \alpha_2, t)$  by applying the chain rule (see Ref. 46 for details). The result is the equation given by

$$\frac{dP}{dt} = \frac{1}{\sqrt{\Omega}} \left[ \frac{1}{\sqrt{\Omega}} \frac{\partial \Pi}{\partial t} - \frac{dm}{dt} \frac{\partial \Pi}{\partial \alpha_m} - \frac{dp}{dt} \frac{\partial \Pi}{\partial \alpha_p} \right]. \quad (28)$$

Inserting  $\bar{m} = \Omega(m + \alpha_m/\sqrt{\Omega})$  and  $\bar{p} = \Omega(p + \alpha_p/\sqrt{\Omega})$  [Eq. (17)] and the approximate step-operator [Eq. (18)] into the Master equation [Eq. (27)], yields an intermittent equation containing terms multiplied by  $1/\sqrt{\Omega}$  raised to different powers

$$\begin{aligned} & \frac{1}{\sqrt{\Omega}} \frac{\partial \Pi}{\partial t} - \frac{dm}{dt} \frac{\partial \Pi}{\partial \alpha_m} - \frac{dp}{dt} \frac{\partial \Pi}{\partial \alpha_p} \\ & = \kappa_M \left[ \frac{\partial_m^2}{2\sqrt{\Omega}} - \partial_m \right] \Pi + \left( \kappa_P \left[ \frac{\partial_p^2}{2\sqrt{\Omega}} - \partial_p \right] \right. \\ & \quad \left. + \delta_M \left[ \partial_m + \frac{\partial_m^2}{2\sqrt{\Omega}} \right] \right) \left( m + \frac{\alpha_m}{\sqrt{\Omega}} \right) \Pi \end{aligned}$$

$$+ \delta_P \left[ \partial_p + \frac{\partial_p^2}{2\sqrt{\Omega}} \right] \left( p + \frac{\alpha_p}{\sqrt{\Omega}} \right) \Pi. \quad (29)$$

Isolating terms entering Eq. (29) independent of  $\Omega$ , corresponding to zeroth order in  $1/\sqrt{\Omega}$ , yields

$$\frac{dm}{dt} \frac{\partial \Pi}{\partial \alpha_m} + \frac{dp}{dt} \frac{\partial \Pi}{\partial \alpha_p} = [\kappa_M - \delta_M m] \partial_m \Pi + [\kappa_P m - \delta_P p] \partial_p \Pi. \quad (30)$$

Equating the coefficients of  $\partial_i \Pi$  then yields the macroscopic rate equations in Eq. (1).

An approximate evolution equation for  $\Pi(\alpha_m, \alpha_p)$  is obtained by isolating terms entering Eq. (29) with first order in  $1/\sqrt{\Omega}$ . This first-order approximation is given by

$$\begin{aligned} \frac{d\Pi}{dt} = & \frac{\kappa_M}{2} \partial_m^2 \Pi + \frac{\kappa_P m}{2} \partial_p^2 \Pi - \kappa_P \partial_p (\alpha_m \Pi) + \delta_M \partial_m (\alpha_m \Pi) \\ & + \frac{\delta_M m}{2} \partial_m^2 \Pi + \frac{\delta_P p}{2} \partial_p^2 \Pi + \delta_P \partial_p (\alpha_p \Pi), \end{aligned} \quad (31)$$

which is a Fokker-Planck equation with the coefficient matrices

$$\mathbf{A} = \begin{bmatrix} -\delta_M & 0 \\ \kappa_P & -\delta_P \end{bmatrix}, \quad (32)$$

$$\mathbf{B} = \begin{bmatrix} \kappa_M + \delta_M m & 0 \\ 0 & \kappa_P m + \delta_P p \end{bmatrix}.$$

The intrinsic noise in the steady state is found by setting the left-hand side of Eq. (23) to zero and solving for  $\langle \alpha_i \alpha_j \rangle$ . This yields the following set of equations:

$$\begin{aligned} -2\delta_M \langle \alpha_m \alpha_m \rangle + \kappa_M + \delta_M m^s &= 0, \\ -(\delta_M + \delta_P) \langle \alpha_m \alpha_p \rangle + \kappa_P \langle \alpha_m \alpha_m \rangle &= 0, \\ -2\delta_P \langle \alpha_p \alpha_p \rangle + \kappa_P \langle \alpha_m \alpha_p \rangle + \kappa_P m^s + \delta_P p^s &= 0, \end{aligned} \quad (33)$$

the solution of which, with  $\phi = \delta_M/\delta_P$ , is written compactly

$$\mathbf{C}^s = \begin{bmatrix} m^s & \frac{p^s}{1+\phi} \\ \frac{p^s}{1+\phi} & \left( p^s + \frac{(p^s)^2}{m^s(1+\phi)} \right) \end{bmatrix}. \quad (34)$$

Using the definition  $\langle \alpha_p \alpha_p \rangle^s = \mathbf{C}_{22}^s$ , it is a simple matter to calculate the intrinsic noise in the protein concentration

$$\eta_{\text{int}}^2 = \frac{\langle p^2 \rangle}{\langle p \rangle^2} = \frac{1}{\Omega} \frac{\langle \alpha_p \alpha_p \rangle}{(p^s)^2} = \frac{1}{\Omega} \left[ \frac{1}{p^s} + \frac{1}{m^s(1+\phi)} \right], \quad (35)$$

in agreement with Eq. (16) in the limit of fast promoter kinetics. The real power of the linear noise approximation, however, lies in the possibility of calculating the intrinsic noise in systems with nonlinear reaction rates where moment generating functions are of no use. The following section illustrates the point using a simple two-dimensional chemical oscillator.

### IV. EFFECTS OF INTRINSIC AND EXTRINSIC NOISE ON THE BRUSSELATOR

The Brusselator is a two-species autocatalytic reaction network described by the macroscopic rate equations

$$\frac{dx}{dt} = 1 + ax^2y - (b + 1)x, \quad \frac{dy}{dt} = -ax^2y + bx, \quad (36)$$

where  $x$  and  $y$  are the concentrations of the two species,  $X$  and  $Y$ , respectively.<sup>60</sup> The macroscopic description admits a unique steady state given by  $(x^s, y^s) = (1, b/a)$ . This steady state is stable when  $b < 1 + a$  and unstable when  $b > 1 + a$ . The critical point,  $b = 1 + a$ , is a Hopf bifurcation where a limit cycle attractor is created. The system is simple enough that it provides a clear forum to present the tools we shall use in later sections since unfortunately the higher-order systems do not yield such simple analytical results.

#### A. Evaluating extrinsic noise

##### 1. Steady state regime

The derivation of sensitivities and noise susceptibilities can be carried out explicitly for  $b < 1 + a$  since an analytic description of the dependence of the steady state on the parameters is available. Applying Eq. (7) at the steady state  $(x^s, y^s) = (1, b/a)$  yields steady state noise susceptibilities that are zero for  $x$  (since  $x^s$  is independent of the parameters) and unity for  $y$  (since  $y^s$  is linear in  $a^{-1}$  and in  $b$ ). Hence, in the linear SNA, the system neither amplifies nor attenuates external perturbations. It is noted that the linear dependence of  $y^s$  on  $b$  and  $a^{-1}$  is analogous to the linear dependence of the steady state protein abundance  $p^s$  on  $\kappa_M$ ,  $\kappa_P$ ,  $\delta_M^{-1}$ , and  $\delta_P^{-1}$  in Sec. II A.

The result  $\chi_{y,a}^s = 1$  is only valid for infinitesimal perturbations in  $a$ . Because of the inverse dependence of  $y^s$  on  $a$  (or of  $p^s$  on  $\delta_M$  and  $\delta_P$ ), perturbations of finite size change the steady state nonlinearly and lead to a breakdown of the linear approximation in Eq. (4). To determine noise susceptibility in this more general case, it is necessary to employ an approach that takes nonlinear effects into consideration. For  $\chi_{y,a}^s$ , this can be accomplished directly by sampling  $N$  different values of the parameter, and then calculating the resulting external noise as described in Sec. II A,

$$\eta_{\text{ext}}^2 = \frac{1}{[\langle y^s(a) \rangle]^2 N} \sum_i [y^s(a_i) - \langle y^s(a) \rangle]^2. \quad (37)$$

Alternatively, when the static parametric noise is associated with the probability distribution  $G(a)$ , the variance in  $y^s$  can be calculated directly via

$$\begin{aligned} \sigma_{y^s}^2(a) &= \langle y^s(a)^2 \rangle - \langle y^s(a) \rangle^2 \\ &\approx \int_{a^-}^{a^+} [y^s(a)]^2 G(a) da - \left\{ \int_{a^-}^{a^+} y^s(a) G(a) da \right\}^2, \end{aligned} \quad (38)$$

where  $(a^-, a^+)$  is a range containing  $a_0$  that is sufficiently large to make the tails of the improper integrals negligible but sufficiently small so that  $b < 1 + a$  for any value of  $a$ . Once the integrals on the right-hand side of Eq. (38) have

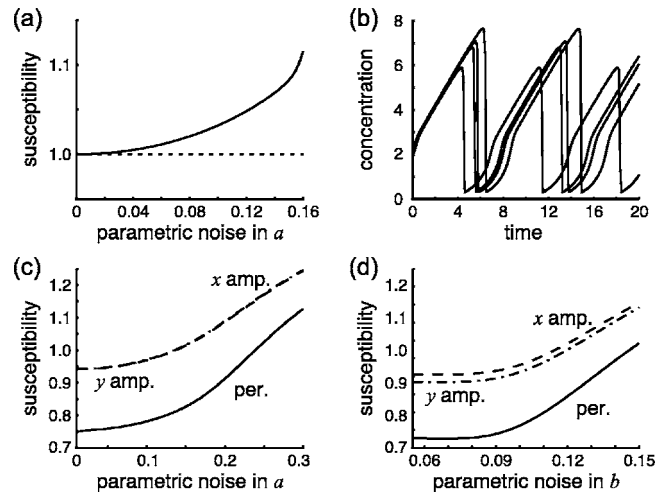


FIG. 2. Parametric noise in the Brusselator system. (a) Steady state susceptibility for  $y$  calculated using the linear (broken line) and nonlinear SNA (full line). The calculation used  $a^- = 0.0001$ ,  $a^+ = 10$ . The truncated tails of the integration are negligible over the range of  $\sigma_a \in (0.0005, 0.08)$ . (b) Desynchronization of initially identical oscillators due to variability in  $a$ . (c), (d) Susceptibilities in oscillation amplitudes and period with parametric noise in (c)  $a$  or (d)  $b$ . In panels (a), (c), and (d), the susceptibilities coincide with the sensitivity at low parametric noise and increases as the parametric noise is increased.

been calculated, which in general must be done numerically, the noise susceptibility  $\chi_{y,a}^s$  can then be calculated as in Eq. (3). This nonlinear SNA method can be used to evaluate the effect of noise in a parameter on any chosen measure of system behavior, independent of the form of the trajectory.

The results of such calculations are shown in Fig. 2(a), for the nominal parameter value  $(a_0, b_0) = (0.5, 1)$ . The figure shows how the steady state noise susceptibility  $\chi_{y,a}^s$  changes as a function of Gaussian parametric noise defined by  $\eta_a = \sigma_a/a_0$ . Notice that the noise susceptibility depends on the magnitude of the parametric noise. For small  $\eta_a$ , we recover the limiting case of  $\chi_{y,a}^s = 1$  corresponding to infinitesimal parametric perturbations. As  $\eta_a$  increases, the nonlinear effects come into play causing an increase in the noise susceptibility as the external noise intensifies. In other words, parametric noise is generally amplified by the system and the degree of the amplification increases with the magnitude of the external perturbations.

##### 2. Oscillatory regime

The Brusselator displays stable oscillations when  $b > 1 + a$ . In this regime, there are several ways to evaluate the impact of external fluctuations. Figure 2(b) shows how parametric noise in the parameter  $a$  affects an ensemble of deterministic oscillators with identical initial conditions. The different values of  $a$  yield oscillators with varying amplitudes and periods. As a result, they progress at different rates along limit cycles, resulting in the desynchronization of individual oscillators. The rate of desynchronization, which is an important indicator of robustness and noise resilience, depends on the variance in the oscillation period, and can be obtained by calculating the autocorrelation function associated with the ensemble-averaged trajectory. However, this approach requires quite intensive computations.

Alternatively, we can apply the SNA as outlined in Sec. II by evaluating the effect of parametric perturbations on characteristics of the trajectory, such as the oscillation amplitude and its period. However, since no analytic description is available for these characteristic, it is necessary to determine the parametric noise susceptibility of the period  $\chi_\rho^{\text{per}}$  and amplitude  $\chi_\rho^{\text{amp}}$  with respect to  $\rho$  using numerical simulations. In the linear regime these measures can be obtained by evaluating parameter sensitivity<sup>34,61–63</sup> (i.e., the linear SNA). For nonlinear dependencies, the measures can be obtained by brute force Monte Carlo sampling where the variance (e.g.,  $\sigma_{\text{per}}^2$  or  $\sigma_{\text{amp}}^2$ ) is obtained from a probability distribution obtained by applying a sufficiently large number of perturbations  $\delta_i$  and running a simulation with the altered parameter value  $\rho_i = \rho_0 + \delta_i$  for each sampled value.

A more efficient method is to calculate the variances by integration as in Eq. (38). Because analytical forms are not available, this implementation of the nonlinear SNA involves tabulating dependencies of the desired characteristics on the parameter  $\rho$  over the entire range ( $\rho^-, \rho^+$ ) of relevant values. While this requires some initial computations, the calculation for each nominal parameter value  $\rho_0$  and for each value of the parametric noise  $\eta_\rho$  can be done very rapidly.

Additionally, numerical continuation packages are available to efficiently map how nontrivial characteristics depend on system parameters and the extension to parametric dependencies, e.g., when  $\rho_0$  is varied and  $\eta_\rho$  is fixed, other oscillation characteristic, such as absolute amplitudes, and more complex characteristics, for example in bursting and period doubling phenomena, can also be addressed. The extension to parametric dependencies, e.g., when  $\rho_0$  is varied and  $\eta_\rho$  is fixed is straightforward.

Figures 2(c) and 2(d) show the noise susceptibilities for the oscillation amplitudes ( $\chi_{x,\rho}^{\text{amp}}$  and  $\chi_{y,\rho}^{\text{amp}}$ ) and the period ( $\chi_\rho^{\text{per}}$ ) when the parametric noise  $\eta_\rho$  is varied when  $a$  and  $b$  are noisy parameters, respectively. As in the analysis for the steady state, the noise susceptibility obtained with the nonlinear SNA coincides with the sensitivity for low noise intensities, and increases as the magnitude of the perturbations becomes larger and nonlinearities becomes important. It is noted that  $\chi_a^{\text{per}}$  can be linked directly to a measure of the rate at which an initially synchronized population goes out of phase. The phase of an oscillator grows linearly in time in proportion to the frequency of the motion and so the rate at which two oscillators are desynchronized is equal to the difference in the reciprocals of the periods. This measure is proportional to the difference in the periods, and so is essentially captured by the susceptibility.

## B. Evaluation of intrinsic noise

With the notations introduced in Sec. III, the Master equation describing the Brusselator system is

$$\begin{aligned} \frac{\partial P}{\partial t} = & \Omega(\mathbf{E}_1^{-1} - 1)P + \frac{a}{\Omega^2}(\mathbf{E}_1^{-1}\mathbf{E}_2^1 - 1)\bar{x}(\bar{x} - 1)\bar{y}P \\ & + (\mathbf{E}_1^1 - 1)\bar{x}P + b(\mathbf{E}_1^1\mathbf{E}_2^{-1} - 1)\bar{x}P, \end{aligned} \quad (39)$$

where  $\bar{x}$  and  $\bar{y}$  are the number of  $X$  and  $Y$  molecules, respectively. This nonlinear equation is approximated using the

LNA by inserting the ansatz in Eq. (17) and the approximate step operator in Eq. (18). In a manner that is completely analogous to the approach taken in Sec. III, the macroscopic rate equations are obtained as the zeroth order approximation by collecting terms that are independent of  $\Omega$  [i.e.,  $(1/\sqrt{\Omega})^0$ ] while the Fokker-Planck equation describing the fluctuations about the macroscopic trajectory is obtained by collecting terms containing  $1/\sqrt{\Omega}$ . Without providing additional details (see Ref. 46), the coefficient matrices associated with the Fokker-Planck equation are given by

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} 2axy - (b+1) & ax^2 \\ b - 2axy & -ax^2 \end{bmatrix}, \\ \mathbf{B} &= \begin{bmatrix} 1 + (b+1)x + ax^2y & -bx - ax^2y \\ -bx - ax^2y & bx + ax^2y \end{bmatrix}. \end{aligned} \quad (40)$$

These matrices are used to evaluate how increased internal fluctuations (i.e., decreased system size) impact both the steady state ( $b < 1+a$ ) and the oscillations ( $b > 1+a$ ).

### 1. The stable regime $b < 1+a$

The equations governing the evolution of the mean and variance of  $\alpha_x$  and  $\alpha_y$  are obtained from Eqs. (22) and (23). In the parameter regime where the macroscopic solution  $(x^s, y^s) = (1, b/a)$  is stable, the mean of the fluctuations vanish as the system approaches the steady state while the variances remain finite. In the steady state the cross-correlation matrix can be shown to be given by

$$\begin{aligned} \mathbf{C}^s &= \begin{bmatrix} \langle \alpha_1^2 \rangle^s & \langle \alpha_1 \alpha_2 \rangle^s \\ \langle \alpha_1 \alpha_2 \rangle^s & \langle \alpha_2^2 \rangle^s \end{bmatrix} \\ &= \frac{1}{1+a-b} \begin{bmatrix} b+1+a & -2b \\ -2b & b^2/a + b/a + b \end{bmatrix}. \end{aligned} \quad (41)$$

The steady state intrinsic noise in the concentrations  $x$  and  $y$  are then obtained using  $\langle \alpha_x^2 \rangle^s = C_{11}^s$  and  $\langle \alpha_y^2 \rangle^s = C_{22}^s$ ,

$$\eta_{x,\text{int}}^2 = \frac{1+a+b}{\Omega(1+a-b)}, \quad \eta_{y,\text{int}}^2 = \frac{a}{b} \eta_{x,\text{int}}^2. \quad (42)$$

As discussed in Sec. III, the approximate steady state probability distribution  $\Pi^s(\alpha_1, \alpha_2)$  is a bivariate Gaussian centered around the steady state  $(1, b/a)$  and the fluctuations are constrained to lie within an error curve [Eq. (25)] obtained from the steady state cross-correlation matrix in Eq. (41). The error curve obtained in the  $x, y$  phase plane with  $K=0.6826$ , corresponding to one standard deviation, is shown in Fig. 3(a) together with points of trajectories generated using stochastic simulations of the Master equation.

Recent investigations of stochasticity in genetic circuits have focused upon the fluctuations around a stable steady state. With a few notable exceptions,<sup>58,59</sup> fluctuations around macroscopic limit cycles have received less attention. In the next section, we review the work of Tomita and co-workers<sup>58</sup> on the two-dimensional Brusselator, extending their work by providing an explicit scheme for carrying out the analysis on oscillators with arbitrary state dimension. Along the way, we develop noise measures that are well suited to quantify fluctuations around a limit cycle.

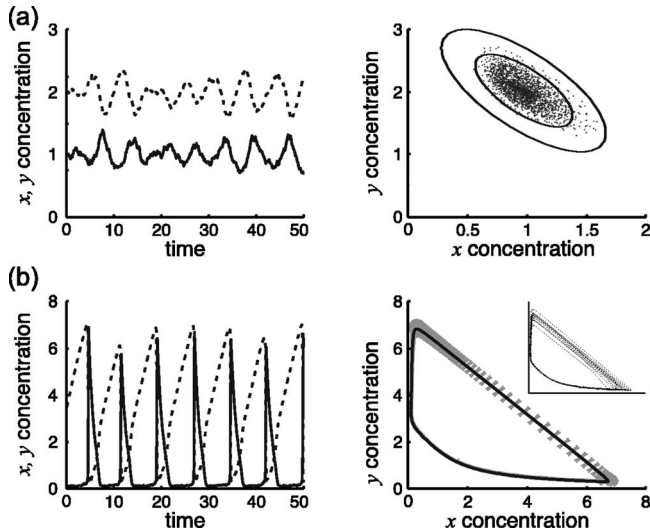


FIG. 3. (a) and (b) Time series and phase plane portraits showing the error curves (one standard deviation) for the Brusselator system in (a) the steady state regime ( $a=0.95$ ,  $b=1.90$ ). The curves represent the first and second standard deviations predicted by the linear noise approximation. (b) The oscillatory regime ( $a=5$ ,  $b=10$ ) for  $\Omega=1000$ . The predictions of the linear noise approximation are shown as gray error curves.

## 2. Limit cycle regime $b > 1 + a$

To compute the intrinsic noise in the parameter regime beyond the Hopf bifurcation, it is necessary to solve the Fokker-Planck equation along the limit cycle. The analysis proceeds almost identically to that outlined in the previous section, except that we must first introduce a new reference frame that moves with the macroscopic trajectory. This is necessary to allow the separation of fluctuations that are perpendicular to the limit cycle from those that are tangent to it.<sup>58,59</sup> The fluctuations along these two directions are both important. The fluctuations along the tangent, with variance that grows unbounded, affects the oscillation phase (and its period) and tends to cause initially identical systems to desynchronize. On the other hand, trajectories perturbed away from the limit cycle are drawn back to it, and the fluctuations perpendicular to the limit cycle contains the information that is required to calculate the variability along the phase-plane trajectory as well as the fluctuations in amplitudes.

To resolve motion into tangent and normal directions pointwise along the limit cycle, we choose the so-called *Frénet frame* (see Appendix B) because it is readily scaled to higher dimensional systems. This change of reference frame transforms the matrices  $\mathbf{A}$ ,  $\mathbf{B}$ , and  $\mathbf{C}$  of the Fokker-Planck equation into the new matrices  $\mathbf{A}'$ ,  $\mathbf{B}'$ , and  $\mathbf{C}'$ . The new matrices are then inserted into Eq. (23) to obtain an equation that governs the evolution of the cross-correlation matrix  $\mathbf{C}'$ . Analogous to the steady state analysis, the elements of  $\mathbf{A}'$  and  $\mathbf{B}'$  give the local stability and the diffusion away from the trajectory, respectively. Specifically, the transformed matrices are both in block-diagonal form (Appendix B) with the first elements  $A'_{11}$  and  $B'_{11}$  giving the *phase drift* and the *phase diffusion along* the macroscopic trajectory. The  $d-1$  sub-blocks of  $\mathbf{A}'$  and  $\mathbf{B}'$  determine these properties along the orthogonal directions. In other words, the variances perpendicular to the macroscopic trajectory are obtained from an

equation that governs the evolution of the  $d-1$  sub-block of  $\mathbf{C}'$  [cf. Eq. (B3) in Appendix B].

The evolution equation for the  $d-1$  sub-block of  $\mathbf{C}'$  associated with the two variable Brusselator system is particularly simple as it contains only scalars. It is given by (see Appendix B)

$$\frac{dC'_{22}}{dt} = 2A'_{22}C'_{22} + B'_{22}. \quad (43)$$

After the transients have passed, the trajectory in the original  $x, y$  coordinates can be anywhere along the limit cycle, but transverse fluctuations are largely confined to a narrow region of width  $\sqrt{C'_{22}}/\sqrt{\Omega}$ , where the  $1/\sqrt{\Omega}$  dependency is introduced by the variable transformation from numbers of molecules into concentrations. That is to say the fluctuations away from the macroscopic trajectory scale as  $1/\sqrt{\Omega}$  in both the oscillatory and the steady state regime. However, because  $C'_{22}$  is evaluated pointwise along the limit cycle, the magnitude of the fluctuations is different in different regions of the phase plane. This is illustrated in Fig. 3(b), which shows the error curve [Eq. (25)] evaluated pointwise along the macroscopic trajectory in the  $x, y$  phase plane together with numerical simulations of the Master equations.

As Ali and Menzinger<sup>59</sup> demonstrate in great detail, a globally stable limit cycle is more usefully thought of as a patchwork of attractive and repulsive segments along the orbit. This phase dependent local stability has observable consequences in stochastic systems; fluctuations during an unstable phase will be amplified, while fluctuations along a locally stable phase are strongly damped. Formally, from Eq. (23), since  $\mathbf{A}$  is a function of the macroscopic trajectory  $\mathbf{x}(t)$ , the eigenvalues of  $\mathbf{A}$  will likely vary along the orbit, in some cases perhaps even change sign. In regions of local instability, fluctuations through  $\mathbf{B}$  will be amplified leading to an increase in variance  $\mathbf{C}$  as seen in the plots drawn from a stochastic simulation,<sup>67</sup> shown in Fig 3(b).

Because the variances change during the oscillation, there is no direct analog to the steady state intrinsic noise measure in the oscillatory regime. We will employ two different measures—the intrinsic amplitude noise  $\eta_{\text{int},i}^{\text{amp}}$  and the total intrinsic noise  $\eta_{\text{int},i}^{\text{tot}}$ . The intrinsic amplitude noise in species  $i$  is defined as the standard deviation in the absolute concentration amplitude divided by the average absolute amplitude

$$\eta_{\text{int},i}^{\text{amp}} = \sqrt{\frac{\sigma_{\text{max},i}^2 + \sigma_{\text{min},i}^2}{(c_{\text{max},i} - c_{\text{min},i})^2}}, \quad (44)$$

where  $c_{\text{max},i}$  and  $c_{\text{min},i}$  are the highest and lowest macroscopic concentrations attained during the oscillation, respectively, and  $\sigma_{\text{max},i}$  and  $\sigma_{\text{min},i}$  are obtained by evaluating the relevant entries in the matrix  $\mathbf{C}'$  at these points. As the second measure, we define the total intrinsic noise in species  $i$  by evaluating the average intrinsic noise over the limit cycle trajectory  $c_i(t)$ ,

$$\eta_{\text{int},i}^{\text{tot}} = \sqrt{\frac{1}{T} \int_{t'}^{t'+T} \frac{C_{ii}'(t)}{\Omega c_i(t)^2} dt}. \quad (45)$$

Based on these definitions, it is clear that the LNA yields intrinsic amplitude noise [Eq. (44)] and intrinsic total noise [Eq. (45)] that scales as  $1/\sqrt{\Omega}$ . However, because they yield a single scalar for each species, they are convenient when evaluating the effect of changing system parameters.

As in the previous section, we would like to evaluate how perturbations affect the variance in the oscillation period. Unfortunately, the period is a nonlocal characteristic that is not immediately available from the local measures of phase drift and diffusion determined by the elements  $A'_{11}$  and  $B'_{11}$ , respectively. The quantity  $C'_{11}$  is of no use as the variance in oscillation phase grows unbounded.

The correlation time  $\tau_c$ , which is determined by the inverse damping rate of the autocorrelation function, is typically calculated numerically from stochastic simulations of the Master equation, although some work has been done recently to quantify  $\tau_c$  analytically.<sup>64</sup> The analytic development is complicated, however, and the phase diffusion does provide somewhat similar information. Recall from Sec. III that the magnitude of  $\mathbf{B}$  in part determines how sensitive a given point in the phase plane is to a decrease in system size. In a similar fashion, the magnitude of term  $B'_{11}$  is related to the sensitivity of the oscillation phase at a given point along the limit cycle and the rate at which initially identical oscillators will desynchronize. As a scalar measure of this sensitivity, we define the averaged intrinsic phase diffusion  $\Delta_{\text{int}}^{\text{diff}}$  as

$$\Delta_{\text{int}}^{\text{diff}} = \frac{1}{T} \int_{t'}^{t'+T} \frac{B'_{11}(t)}{\Omega \|c(t)\|}, \quad (46)$$

where the norm  $\|c(t)\|$  of the concentration vector is introduced to allow direct comparison when a parameter value is changed.

For the Brusselator, the transformation of  $\mathbf{B}$  into the Frénet frame (Appendix B) yields a phase diffusion coefficient that at time  $t$  [and position  $x(t), y(t)$  in the phase-plane] is given by

$$B'_{11} = f_x^2 B_{11} + 2f_x f_y B_{12} + f_y^2 B_{22}, \quad (47)$$

where  $f_x$  and  $f_y$  are the macroscopic rate equations for the concentrations of  $x$  and  $y$  defined in Eq. (36) and  $B_{ij}$  are the elements of the diffusion matrix defined in Eq. (40). Hence, to calculate the average intrinsic diffusion, we need to evaluate Eq. (47) pointwise along the limit cycle and insert the result into Eq. (46). The result for fixed system size and varying  $a$  is shown in Fig. 4(a). It is seen that for the total noise, the concentration of  $x$  is more sensitive to changes in  $a$  than the  $y$  species which can be argued heuristically by noticing that an increase in  $a$  will make the top-left entry of the Jacobian matrix [Eq. (40)] more positive resulting in a greater noise susceptibility in  $x$ . The amplitude noise is roughly constant over  $a$  for both species as seen in Fig. 4(b). As a control, the dependency of  $\Delta_{\text{int}}^{\text{diff}}$  on  $a$ , which is shown in Fig. 4(c), should be compared with the correlation time  $\tau_c$  obtained from numerical simulations of the Master equation<sup>64</sup> in Fig. 4(d). The two measures follow the same

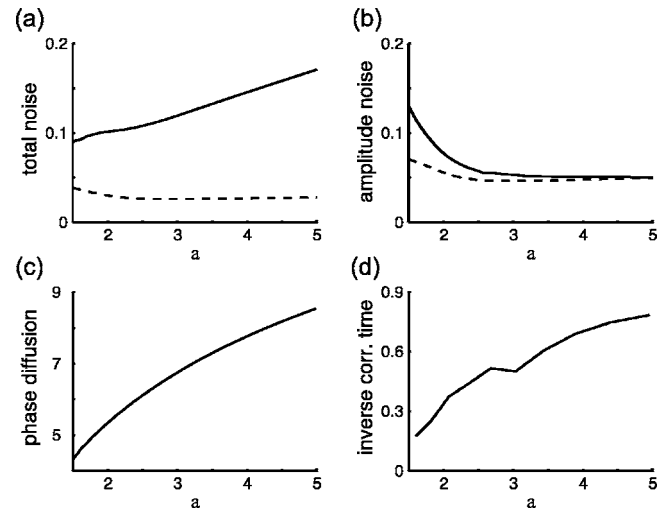


FIG. 4. Intrinsic noise the oscillations of  $x$  (full curves) and  $y$  (broken curves) generated by the Brusselator system for  $\Omega=1000$  when  $a$  is varied, holding the ratio  $b/a=2$  fixed. (a) Total noise. (b) Amplitude noise. (c) Average phase diffusion. (d) Inverse correlation time computed from 500 trials of a stochastic simulation.

general trend, suggesting that the averaged intrinsic phase diffusion provides an estimate of the effect of system size on the oscillation period that can be calculated very rapidly from the macroscopic rate equations and the macroscopic phase-plane trajectory.

## V. GENETIC OSCILLATORS

In this section, we apply the principles discussed in the previous sections in a comparative study of two genetic networks, the Repressilator and the Atkinson oscillator. These networks yield oscillatory gene expression when implemented in *E. coli*, but employ different designs to achieve oscillatory dynamics (Fig. 5). The Repressilator network consists of three transcriptional repressors whose genes and promoters are connected serially to form a negative feedback loop circuit. On the other hand, the Atkinson oscillator employs two transcriptional regulators, an activator and a repressor, organized in a circuit that combines positive and negative feedback loop regulation. While both networks yield oscillations, they appear to have markedly different sensitivity to noise. Only a fraction of the cells harboring the Repressilator oscillate and these oscillation desynchronize at a very short time scale.<sup>25</sup> On the other hand, cells harboring the Atkinson network maintain damped oscillations at the population level over a long time scale<sup>31</sup> indicative of high synchrony at the single-cell level. One possible explanation for the observed differences in apparent robustness is that the two network designs have inherently different susceptibilities to internal and external noise. This is the question that we seek to explore using the tools demonstrated in Sec. IV.

In order to make a comparative analysis, we adopt a consistent framework to model the two networks. In this framework, the three-gene Repressilator network in Fig. 5(a) is modeled as a six-variable system (three mRNA species and three proteins) governed by the normalized macroscopic rate equations

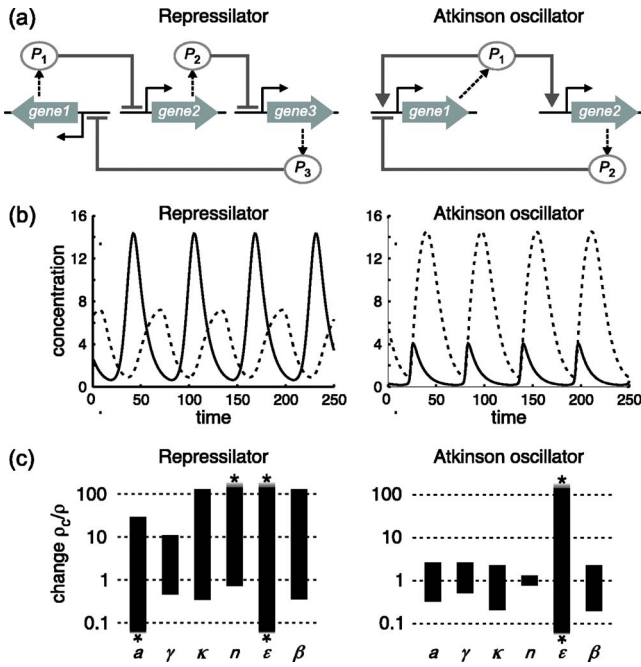


FIG. 5. Macroscopic oscillations in models of the Repressilator and the Atkinson networks. (a) Network design. (b) Oscillations in  $p_1$  (full curves) and  $p_2$  (broken curves) when all parameters except  $\kappa$  are identical (see text). The range in (individual) parameter values over which the system exhibits oscillations. A fold change of 100 corresponds to a 100-fold increase while a fold change of 1/10 corresponds to a tenfold (90%) decrease. Asterisks indicate that oscillations cannot be suppressed by increasing or decreasing the parameter.

$$\frac{dm_i}{d\tau} = a_i \kappa_i + \frac{\kappa_i}{1 + p_{i+1}^{n_i}} - m_i, \tag{48}$$

$$\frac{dp_i}{d\tau} = -\varepsilon(p_i - \beta_i m_i), \quad i = 1, 2, 3,$$

where the effect of the repressor on the rate mRNA synthesis is modeled using Hill-type kinetics and the decay rates are assumed to be the same for each mRNA and protein species. In these equations, which are obtained by introducing the dimensionless time  $\tau = \delta_M \cdot t$  and expressing mRNA and protein concentrations relative to the Hill constant, the term  $a \cdot \kappa_i$  captures repressor-independent (i.e., basal) mRNA synthesis,  $\beta_i$  is proportional to the rate of protein synthesis, and  $\varepsilon$  is the ratio of mRNA and protein decay rates.

The Atkinson network in Fig. 5(b) involves transcription regulated by two proteins and is modeled in a similar fashion as a four-dimensional system governed by the rate equations

$$\begin{aligned} \frac{dm_1}{dt} &= a_1 \kappa_1 + \frac{\kappa_1}{1 + p_2^{n_2}} \cdot \frac{p_1^{n_1}}{1 + p_1^{n_1}} - m_1, \\ \frac{dm_2}{dt} &= a_2 \kappa_2 + \frac{\kappa_2 p_1^{n_1}}{1 + p_1^{n_1}} - m_2, \end{aligned} \tag{49}$$

$$\frac{dp_i}{dt} = -\varepsilon(p_i - \beta_i m_i), \quad i = 1, 2.$$

In these equations, it is assumed that the synthesis rate of activator-encoding mRNA ( $m_1$ ) can be expressed as the product of Hill-type functions describing the activating effects of  $p_1$  and the repressing effect of  $p_2$ . For simplicity, we will additionally assume that the all relative basal transcription rates, Hill coefficients and relative rates of protein synthesis are identical (i.e.,  $a_i = a$ ,  $n_i = n$ ,  $\beta_i = \beta$ ). Moreover, we set  $\kappa_3 = \kappa_2$  for the Repressilator and define  $\kappa_2 = \kappa$  for both models. To have  $\kappa$  as a single parameter governing overall transcription, we further introduce  $\gamma$  as the ratio of the maximal transcription rates and express  $\kappa_1$  in terms of  $\kappa$  through  $\kappa_1 = \gamma \kappa$ . With these restrictions, the two models involves a total of six free parameters ( $a, \gamma, \kappa, n, \varepsilon, \beta$ ).

For our comparative analysis, we chose parameter values that, with exception of  $\kappa$ , are identical ( $a = 10^{-3}$ ,  $\gamma = 5$ ,  $n = 3$ ,  $\varepsilon = 0.1$ ,  $\beta = 20$ ) and yield oscillations in the deterministic description of both systems. The oscillations observed with the chosen parameters are shown in Fig. 5(b). To obtain oscillations with comparable maximal amplitudes, a higher value of  $\kappa$  was used in the simulations of the Atkinson network ( $\kappa = 1$ ) than in the simulations of the Repressilator ( $\kappa = 0.4$ ). In Fig. 5, we show the change in these parameter values that are required to suppress the oscillation. This comparison indicates that the likelihood of observing oscillations is greater in the Repressilator than in the Atkinson network. With the exception of the ratio of protein and mRNA decay rates  $\varepsilon$ , whose value does not affect the ability of either system to oscillate, the Repressilator shows oscillations in a much greater range of parameter values that the Atkinson oscillator.

In the following sections, we compare in a similar fashion the effects of intrinsic and extrinsic variability on the oscillations in the two systems by applying the nonlinear SNA and the LNA as described in the previous sections. The results of these analyses are by no means guaranteed to agree with the (deterministic) bifurcation analysis shown in Fig. 5(c).

### A. Effects of extrinsic variability

To evaluate the robustness of the two oscillating gene networks to external perturbations, we apply the nonlinear SNA method as demonstrated in Sec. IV A 2 using the Brusselator system. In Fig. 6 we show the change in external noise susceptibility of the oscillation period in the Repressilator and the Atkinson network with increased variability in one or more of the system parameters.

The noisy parameters are chosen as the ratio of protein and mRNA decay rates ( $\varepsilon$ ), the maximal transcription rate of mRNA ( $\kappa$ ), and inhomogeneity in maximal transcription rates ( $\gamma$ ), though analysis involving other parameters could also be of interest. Simultaneous noise in all the maximal transcription rates is biologically relevant as factors that are global to each cell, such as enzymatic activities and metabolic state, are expected to affect the expression of each gene similarly.

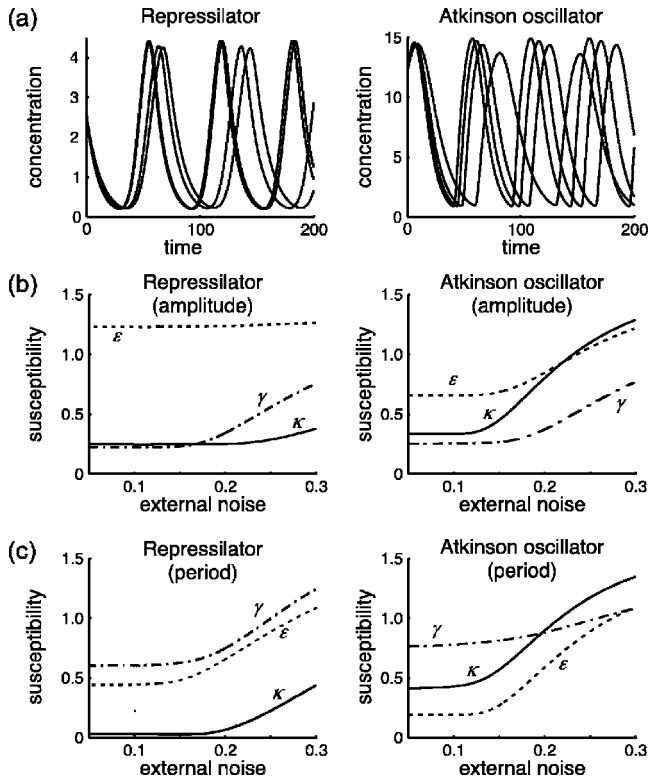


FIG. 6. Parametric noise in the Repressilator and the Atkinson network. (a) Desynchronization of initially identical oscillators due to the variability in  $\epsilon$ . (b) Parametric noise susceptibility in the amplitude of oscillations in protein  $p_2$  when parametric noise is applied to the parameters  $\epsilon$ ,  $\kappa$ , and  $\gamma$ . (c) Parametric noise susceptibility in the period. The nominal parameter values are  $a=10^{-3}$ ,  $\gamma=5$ ,  $n=3$ ,  $\epsilon=0.1$ ,  $\beta=20$ ,  $\kappa=1$  (Atkinson) or  $\kappa=0.4$  (Repressilator). The susceptibilities coincide with the sensitivity at low parametric noise and increases as the parametric noise is increased.

As for the case of the Brusselator system, these graphs show the susceptibility for both networks coincide with the sensitivity in the linear regime (small noise intensity), and the nonlinearity having an effect as the distribution widens. It is noted that with the exception of  $\epsilon$ , the noise susceptibilities obtained for the Repressilator are consistently lower than those obtained for the Atkinson network.

An interesting feature of the Atkinson oscillator is the difference between the susceptibility in the period with respect to  $\kappa$  and  $\gamma$ . For small perturbations, a shift in  $\gamma$ , which affects only a single rate, has a larger effect on the period than a corresponding shift in the coordinated rate  $\kappa$ . However, as evidenced by the crossover between the two curves, the situation is reversed for larger perturbations, for which changes in the coordinated rate are felt more strongly.

While the analysis is not comprehensive, it suggests that the two networks have comparable robustness, though the Repressilator appears more sensitive to shifts in the ratio of mRNA and protein decay rates.

**B. Effects of intrinsic fluctuations**

In this section, we perform a limited comparative LNA analysis of the impact on the oscillations generated by the two genetic networks for decreased system size. The analysis

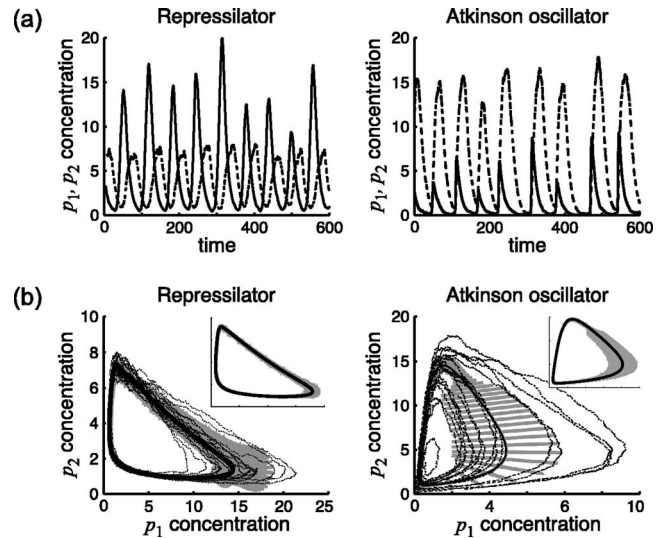


FIG. 7. Comparison of the effect of reduced system size ( $\Omega=100$ ) on the oscillations generated in the Repressilator and the Atkinson networks. (a) Oscillatory time series. The concentration of  $p_1$  and  $p_2$  are drawn with full and broken lines, respectively. (b) Phase planes. The deterministic and stochastic trajectory are shown as solid and broken curves, respectively. The error curves are shown in gray. Insets show the deterministic trajectories and the error-curves obtained for  $\Omega=1000$ .

is carried out using the nominal parameter set and for variation in ratio of mRNA and protein decay rates.

The oscillations generated by the two networks for a system size of  $\Omega=100$  are shown in Fig. 7(a). This system size corresponds to a maximal amplitude of about 1500 molecules in  $p_1$  for the Repressilator and in  $p_2$  for the Atkinson oscillator. To maintain consistency, the propensities used in the stochastic simulation<sup>67</sup> are identical to those used in the macroscopic description, including the time averaging implied by the Hill-type kinetics. It is noted that the oscillation amplitude of  $p_1$  is fairly irregular in the Atkinson network and seems to occasionally reach abundance levels that are significantly higher than in the macroscopic limit. Apparently, when the concentration of the repressor  $p_2$  is low, stochastic effects may facilitate rapid synthesis and accumulation of the autocatalytic protein  $p_1$ . Direct inspection of the stochastic trajectories thus suggests that the Atkinson network is more sensitive to intrinsic noise than the Repressilator, at least in terms of the maximal oscillation amplitude of  $p_1$ .

The phase plane portraits associated with the stochastic time series are shown in Fig. 7(b) together with the macroscopic trajectories and error curves calculated along this trajectory. In both cases, the stochastic phase-plane trajectories are largely confined to the region predicted by the LNA. Additionally, while being only an approximate measure, the LNA is able to capture the relative high variability in the amplitude of  $p_1$  in the Atkinson network.

Analogously to the prediction of a  $1/\sqrt{\Omega}$  scaling of intrinsic noise in steady state gene expression, the LNA predicts that the variance perpendicular to the direction of the limit cycle scales as  $1/\sqrt{\Omega}$  [see Fig. 7(b) insets]. This is essentially because the calculation of the cross-correlation matrix is based on macroscopic rate equations that are inde-

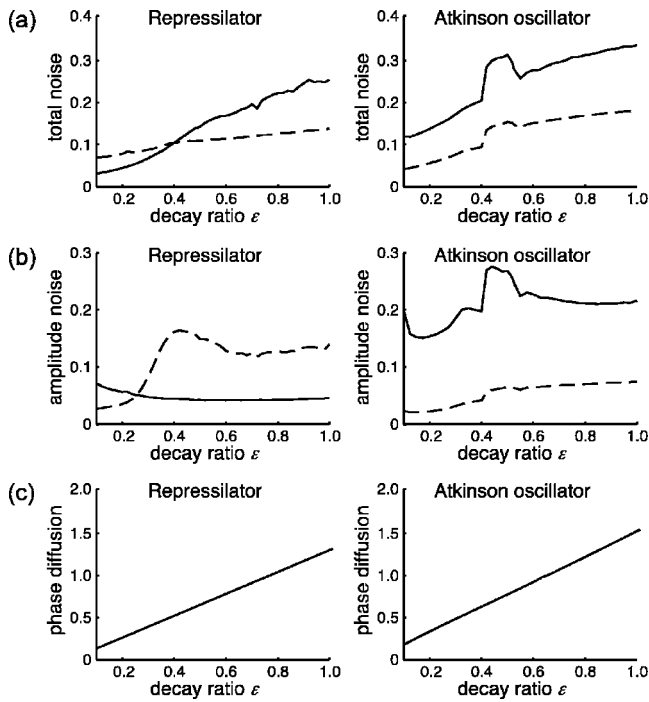


FIG. 8. Comparison of (a) total noise, (b) amplitude noise, and (c) phase diffusion for the two networks when the parameter  $\epsilon$  is varied for  $\Omega=100$ . The noise associated with  $p_1$  and  $p_2$  are drawn with full and broken lines, respectively, in (a) and (b).

pendent of system size. It is expected that LNA will not provide sufficiently adequate estimates when the system size becomes very low, or near critical points where the system changes its qualitative behavior.<sup>47</sup> The advantage of the LNA is that it is much more efficient when the number of molecules is relatively high and brute-force simulations of the Master equation become prohibitively long.

The LNA is very efficient in estimating the effects of varying a system parameter for fixed system size. An example of such an analysis is shown in Fig. 8, which shows how the total intrinsic noise, the intrinsic amplitude noise and the average phase diffusion changes when the value of  $\epsilon$  is increased from its nominal value. In the two networks, the total noise, Fig. 8(a), is consistently higher for  $p_1$  regardless of the value of  $\epsilon$ . The largest difference in total  $p_1$  noise arising at low values of  $\epsilon$  where, for the Repressilator, the total noise in  $p_1$  is less than that of  $p_2$ .

It is interesting to note that the total noise in Fig. 8(a) are not necessarily correlated with the variability in oscillation amplitude noise. Figure 8(b) illustrates how the estimated noise in the absolute amplitudes changes with  $\epsilon$ . For the Repressilator, the amplitude noise in  $p_1$  decreases with an increase in  $\epsilon$  in contrast to an increase of the total  $p_1$  noise seen in Fig. 8(a). This effect presumably arises from an increase in variance in regions along the trajectory with intermediate abundance of  $p_1$ . For the Atkinson oscillator, the total  $p_1$  noise correlates with the amplitude noise, consistent with the total noise being dominated by large variability in the rate of autocatalytic  $p_1$  synthesis.

Calculations of the average phase diffusion (Fig. 8) indicates that increasing  $\epsilon$  makes the oscillation period more

sensitive to intrinsic noise in both the Repressilator and the Atkinson network. This, in turn, suggests that employing stable proteins may decrease the rate of desynchronization between initially identical oscillators. This is interesting considering this result in the light that decreased mRNA half-life, i.e., increased  $\epsilon$  through time averaging decreases intrinsic gene expression noise in the steady state. Additionally, even though the oscillation period of the Repressilator appears slightly more resilient to intrinsic noise, it appears that the two oscillator designs yield oscillations that are about equally robust to intrinsic noise under comparable conditions. However, given the limited scope of our analysis, the generality of this result remains to be determined.

## VI. DISCUSSION

The comparative analysis of the oscillations generated by the Repressilator and the Atkinson networks suggests that the former is more robust. This increased robustness is manifested on several levels. It is much more likely to observe deterministic oscillations in the Repressilator given a random set of parameter values and this system has an overall increased resilience to both extrinsic and intrinsic noise in terms of both oscillation period and amplitudes. This is in direct contrast to experimental observations and suggests that factors other than network design may be more important.

It is possible that the use of destabilized proteins in the Repressilator experiment may account for the observed lack of robustness in this system. Protein half-life is in the present models inversely proportional to the parameter  $\epsilon$  and increasing the  $\epsilon$  leads to increased total noise, amplitude noise, and phase diffusion in both systems. Interestingly, if we compare the total noise in Fig. 8(a) obtained with  $\epsilon=0.1$  for the Atkinson network to that obtained with  $\epsilon=1.0$  for the Repressilator (i.e., a tenfold decrease in protein half-life for the Repressilator), we notice that the total noise is significantly higher for the Repressilator system. Similarly, in Fig. 8(c), the phase diffusion is significantly higher for the Repressilator with  $\epsilon=1.0$  than for the Atkinson network with  $\epsilon=0.1$ . Hence, differences in protein stability could contribute to the experimentally observed differences in apparent robustness.

Another potential explanation for the experimentally observed differences in apparent robustness is the use of self-replicating plasmids to carry the Repressilator network instead of chromosomal genes in the case of the Atkinson network. The number of plasmids per cell are known to fluctuate greatly<sup>65</sup> and this could in principle introduce a high level of cell-cell variability in the system parameters thereby compromising network function.

It is important to emphasize that the analysis presented here presumes the systems are in a parameter regime conducive to *macroscopic* oscillations, but for stochastic systems, that is only half the story as noise tends to extend the range of parameters that sustain oscillations. These so-called noise-induced oscillations are most easily realized in excitable systems, such as the Atkinson oscillator, and have been suggested as a design feature that confers noise resistance to chemical oscillators<sup>26</sup> by harnessing noise to serve a constructive end.<sup>28,66</sup> The methods described here are not di-

rectly applicable to estimate the effect of fluctuations on local macroscopic stability of a fixed steady state, although the LNA does provide a quantitative measure of the magnitude of the fluctuations. The discrepancy between the present analysis and the observed behavior of the two genetic circuits may then be explained by hypothesizing that it is actually the noise that *causes* the oscillations and not the case that the systems are macroscopically oscillatory with noise superimposed upon that motion.

The comparative analysis could be extended in a number of directions to give more conclusive results. With respect to the effect of external noise in parameter values, a number of generalizations present themselves. This preliminary analysis was restricted to perturbations in single parameters while others remained fixed. To extend this analysis to a complete local investigation of parameter space, it would be necessary to address the covariance in the metric arising from the simultaneous perturbation of multiple parameters. Such computations could be time consuming if the dimension of the parameter space is large, but should be feasible on simple models such as those considered here. An additional generalization would be to consider the effect of time-varying perturbations in the parameter values. This complicates matters considerably, requiring the description of the system dynamics as a stochastic process, e.g., as a set of stochastic differential equations. Finally, the choice of probability distribution could be reconsidered. Here we have made a natural choice of Gaussian distributions, but one could imagine situations in which an alternative distribution is deemed more relevant. In particular, for parameters near a bifurcation point, a log normal distribution may serve to explore a larger region of the appropriate parameter regime.

The intrinsic noise estimate using the LNA can be improved by implementing an elimination of the fast-decaying eigenmodes from the drift and diffusion matrices  $\mathbf{A}$  and  $\mathbf{B}$ ,<sup>47</sup> though for ease of presentation the details of this additional complication were not discussed here. Additionally, the LNA relies upon a *stable* asymptotic macroscopic state, either a fixed-point or a limit cycle. More esoteric dynamical behavior, such as bifurcations, multistabilities, and so on, require more sophisticated treatment (see Ref. 46 for details). For multistable systems, for example, either the attractors must be considered individually<sup>49</sup> or the full chemical Langevin equation must be used.<sup>68</sup>

We have assumed throughout that the reactants are distributed *homogeneously* in space. The cell interior is a crowded environment and several strategies are used to increase local concentrations of reactants to suppress noise (e.g., DNA looping, scaffolding, etc.), as well as exploitation of spatial inhomogeneity to achieve important physiological functions, such as precise location of cell division.<sup>69</sup> In circumstances where spatial inhomogeneity is important, often numerical simulation is the only resort,<sup>70</sup> although the linear noise approximation can sometimes be generalized to such systems.<sup>71</sup>

## APPENDIX A: THE DIFFUSION MATRIX

To arrive at a formal definition of  $\mathbf{B}$  in terms of known quantities without having to expand the Master equation di-

rectly, we briefly outline the approach of Elf and Ehrenberg.<sup>47</sup> Consider the capture reaction in the Brusselator example,



which proceeds at a rate  $\nu = a x^2 y$  resulting in the change  $n_1 \rightarrow n_1 + 1$  and  $n_2 \rightarrow n_2 - 1$ . Each individual reaction in the network proceeds at a certain rate  $\nu_j$  and changes species  $n_i$  by an integer  $S_{ij}$ . The collection of elements  $S_{ij}$  is called the *stoichiometry matrix* and the collection  $\nu_j$  is often called the *velocity vector*. For the scaled Brusselator,

$$\mathbf{S} = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 0 & -1 & 1 & 0 \end{bmatrix} \quad \text{and} \quad \boldsymbol{\nu} = [1 \ ax^2y \ bx \ x]^T.$$

By definition,  $\mathbf{f}(\mathbf{c}) = \mathbf{S} \cdot \boldsymbol{\nu}$ . The matrix  $\mathbf{A}$  in the linear noise approximation is then,

$$A_{ij} = \left. \frac{\partial f_i}{\partial c_j} \right|_{\mathbf{c}(t)} = \left. \frac{\partial (\mathbf{S} \cdot \boldsymbol{\nu})_i}{\partial c_j} \right|_{\mathbf{c}(t)}. \quad (\text{A2})$$

One can also show that in terms of the stoichiometry matrix,<sup>47</sup>

$$\mathbf{B} = \mathbf{S} \cdot \text{diag}[\boldsymbol{\nu}] \cdot \mathbf{S}^T. \quad (\text{A3})$$

Immediately, we see that if a particular reaction initiates a large stoichiometric change ( $S_{ij} > 1$ ), the fluctuations through that channel will be large.

## APPENDIX B: MOVING REFERENCE FRAME

From Eq. (19), the rate vector  $\mathbf{f}(\mathbf{c})$  is always tangent to the limit cycle. The first  $d-1$  time derivatives of  $\mathbf{f}(\mathbf{c})$  are never parallel with one another, and therefore provide a convenient algorithm for constructing an orthonormal basis moving along the macroscopic trajectory,

$$\hat{\mathbf{e}}_1(t) = \frac{\mathbf{f}(\mathbf{c}(t))}{\|\mathbf{f}(\mathbf{c}(t))\|}, \quad \hat{\mathbf{e}}_i(t) = \frac{\mathbf{e}_i(t)}{\|\mathbf{e}_i(t)\|}; \quad (\text{B1})$$

$$\mathbf{e}_i(t) = \frac{d^{i-1} \mathbf{f}(\mathbf{c}(t))}{dt^{i-1}} - \sum_{j=1}^{i-1} \left\langle \frac{d^{i-1} \mathbf{f}(\mathbf{c}(t))}{dt^{i-1}}, \hat{\mathbf{e}}_j(t) \right\rangle \hat{\mathbf{e}}_j(t). \quad (\text{B2})$$

This new coordinate frame is called the *Frénet frame*. The orthonormal basis vectors  $\{\hat{\mathbf{e}}_i\}$  are used to construct the unitary rotation matrix  $\mathbf{U} = [\hat{\mathbf{e}}_1(t), \hat{\mathbf{e}}_2(t), \dots, \hat{\mathbf{e}}_d(t)]^T$ , that resolves motion into tangent and normal directions pointwise along the limit cycle. The evolution of the variance in the moving coordinate system is

$$\frac{d\mathbf{C}'}{dt} = \mathbf{A}' \mathbf{C}' + \mathbf{C}' \mathbf{A}'^T + \mathbf{B}', \quad (\text{B3})$$

where  $\mathbf{A}'$  and  $\mathbf{B}'$  are the transformed matrices,

$$\mathbf{A}' = \mathbf{U} \mathbf{A} \mathbf{U}^T + \frac{d\mathbf{U}}{dt} \mathbf{U}^T, \quad \mathbf{B}' = \mathbf{U} \mathbf{B} \mathbf{U}^T. \quad (\text{B4})$$

The transformed Jacobian matrix  $\mathbf{A}'$  has the structure

$$\mathbf{A}' = \begin{bmatrix} \lambda_{tt} & \cdots \\ \vdots & \tilde{\mathbf{A}} \end{bmatrix}.$$

The reduced evolution equation for the stable variance is

$$\frac{d\tilde{C}}{dt} = \tilde{A}\tilde{C} + \tilde{C}\tilde{A}^T + \tilde{B}, \quad (\text{B5})$$

where  $\tilde{B}$  and  $\tilde{C}$  are defined analogously to  $\tilde{A}$  and represent the lower-right block of  $B'$  and  $C'$ , respectively. The variance in the original coordinate system is calculated by inverting the transformation,

$$C = U^T \begin{bmatrix} 0 & 0 \\ 0 & \tilde{C} \end{bmatrix} U.$$

<sup>1</sup>C. V. Rao, D. M. Wolf, and A. P. Arkin, *Nature (London)* **420**, 231 (2002).

<sup>2</sup>J. Paulsson, *Nature (London)* **427**, 415 (2004).

<sup>3</sup>M. Kærn, T. C. Elston, W. J. Blake, and J. J. Collins, *Nat. Rev. Genet.* **6**, 451 (2005).

<sup>4</sup>J. M. Raser and E. K. O'Shea, *Science* **309**, 2010 (2005).

<sup>5</sup>M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain, *Science* **297**, 1183 (2002).

<sup>6</sup>E. M. Ozbudak, M. Thattai, I. Kurtser, A. D. Grossman, and A. van Oudenaarden, *Nat. Genet.* **31**, 69 (2002).

<sup>7</sup>W. J. Blake, M. Kærn, C. R. Cantor, and J. J. Collins, *Nature (London)* **422**, 633 (2003).

<sup>8</sup>J. M. Raser and E. K. O'Shea, *Science* **304**, 1811 (2004).

<sup>9</sup>A. Becskei and L. Serrano, *Nature (London)* **405**, 590 (2000).

<sup>10</sup>A. Becskei, B. Seraphin, and L. Serrano, *EMBO J.* **20**, 2528 (2001).

<sup>11</sup>F. J. Isaacs, J. Hasty, C. R. Cantor, and J. J. Collins, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 7714 (2003).

<sup>12</sup>E. M. Ozbudak, M. Thattai, H. N. Lim, B. I. Shraiman, and A. Van Oudenaarden, *Nature (London)* **427**, 737 (2004).

<sup>13</sup>J. M. Pedraza and A. van Oudenaarden, *Science* **307**, 1965 (2005).

<sup>14</sup>S. Hooshangi, S. Thiberge, and R. Weiss, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 3581 (2005).

<sup>15</sup>M. Acar, A. Becskei, and A. van Oudenaarden, *Nature (London)* **435**, 228 (2005).

<sup>16</sup>N. Rosenfeld, J. W. Young, U. Alon, P. S. Swain, and M. B. Elowitz, *Science* **307**, 1962 (2005).

<sup>17</sup>L. A. Colman-Lerner, A. Gordon, E. Serra, T. Chin, O. Resnekov, D. Endy, C. G. Pesce, and R. Brent, *Nature (London)* **437**, 699 (2005).

<sup>18</sup>L. S. Weinberger, J. C. Burnett, J. E. Toettcher, A. P. Arkin, and D. V. Schaffer, *Cell* **122**, 169 (2005).

<sup>19</sup>J. Stelling, U. Sauer, Z. Szallasi, F. Doyle Jr., and J. Doyle, *Cell* **118**, 675 (2004).

<sup>20</sup>H. Kitano, *Nat. Rev. Genet.* **5**, 826 (2004).

<sup>21</sup>N. Barkai and S. Leibler, *Nature* **387**, 913 (1997).

<sup>22</sup>P. A. Spiro, J. S. Parkinson, and H. G. Othmer, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 7263 (1997).

<sup>23</sup>M. D. Levin, C. J. Morton-Firth, W. N. Abouhamad, R. B. Bourret, and D. Bray, *Biophys. J.* **74**, 175 (1998).

<sup>24</sup>U. Alon, M. G. Surette, N. Barkai, and S. Leibler, *Nature (London)* **397**, 168 (1999).

<sup>25</sup>M. B. Elowitz and S. Leibler, *Nature (London)* **403**, 335 (2000).

<sup>26</sup>J. M. G. Vilar, H. Y. Kueh, N. Barkai, and S. Leibler, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 5988 (2002).

<sup>27</sup>D. Gonze, J. Halley and A. Goldbeter, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 673 (2002).

<sup>28</sup>R. Steuer, C. Zhou, and J. Kurths, *BioSystems* **72**, 241 (2003).

<sup>29</sup>B. A. Mello and Y. Tu, *Biophys. J.* **84**, 2943 (2003).

<sup>30</sup>Z. H. Hou and H. W. Xin, *J. Chem. Phys.* **119**, 11508 (2003).

<sup>31</sup>M. R. Atkinson, M. A. Savageau, J. T. Myers, and A. J. Ninfa, *Cell* **113**, 597 (2003).

<sup>32</sup>I. Mihalcescu, W. H. Hsing, and S. Leibler, *Nature (London)* **430**, 81 (2004).

<sup>33</sup>E. Nagoshi, C. Saini, C. Bauer, T. Laroche, F. Naef, and U. Schibler, *Cell* **119**, 693 (2004).

<sup>34</sup>J. Stelling, E. D. Gilles, and F. J. Doyle, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 13210 (2004).

<sup>35</sup>C. V. Rao, J. R. Kirby, and A. P. Arkin, *PLoS Biol.* **2**, E49 (2004).

<sup>36</sup>E. Korobkova, T. Emonet, J. M. G. Vilar, T. S. Shimizu, and P. Cluzel, *Nature (London)* **428**, 574 (2004).

<sup>37</sup>A.-J.F. Carr and D. Whitmore, *Nat. Cell Biol.* **7**, 319 (2005).

<sup>38</sup>W. Bialek and S. Setayeshgar, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 10040 (2005).

<sup>39</sup>J. Wolf, S. Becker-Weimann, and R. Heinrich, *IEE Systems Biology* **2**, 35 (2005).

<sup>40</sup>D. Bell-Pedersen, V. M. Cassone, D. J. Earnest, S. S. Golden, P. E. Hardin, T. L. Thomas, and M. J. Zoran, *Nat. Rev. Genet.* **6**, 544 (2005).

<sup>41</sup>A. Wagner, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 11775 (2005).

<sup>42</sup>N. Friedman, S. Vardi, M. Ronen, U. Alon, and J. Stavans, *PLoS Biol.* **3**, 1261 (2005).

<sup>43</sup>R. Gunawan, Y. Cao, L. Petzold, and F. Doyle, *Biophys. J.* **88**, 2530 (2005).

<sup>44</sup>N. Barkai and S. Leibler, *Nature (London)* **403**, 267 (2000).

<sup>45</sup>N. G. van Kampen, *Can. J. Phys.* **39**, 551 (1961).

<sup>46</sup>N. G. van Kampen, *Adv. Chem. Phys.* **34**, 245 (1976).

<sup>47</sup>J. Elf and M. Ehrenberg, *Genome Res.* **13**, 2475 (2003).

<sup>48</sup>J. Elf, J. Paulsson, O. G. Berg and M. Ehrenberg, *Biophys. J.* **84**, 154 (2003).

<sup>49</sup>R. Tomioka, H. Kimura, T. J. Kobayashi and K. Aihara, *J. Theor. Biol.* **229**, 501 (2004).

<sup>50</sup>P. S. Swain, M. B. Elowitz, and E. D. Siggia, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 12795 (2002).

<sup>51</sup>M. A. Savageau, *Biochemical Systems Analysis* (Addison-Wesley, Reading, 1976).

<sup>52</sup>R. Heinrich and S. Schuster, *Regulation of Cellular Systems* (Springer, New York, 1996).

<sup>53</sup>D. McQuarrie, *J. Appl. Probab.* **4**, 413 (1967).

<sup>54</sup>N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1992).

<sup>55</sup>T. B. Kepler and T. C. Elston, *Biophys. J.* **81**, 3116 (2001).

<sup>56</sup>A. Becskei, B. B. Kaufmann, and A. van Oudenaarden, *Nat. Genet.* **37**, 937 (2005).

<sup>57</sup>P. Swain, *J. Mol. Biol.* **344**, 965 (2004).

<sup>58</sup>K. Tomita, T. Ohta and H. Tomita, *Prog. Theor. Phys.* **52**, 1744 (1974).

<sup>59</sup>F. Ali and M. Menzinger, *Chaos* **9**, 348 (1999).

<sup>60</sup>S. H. Strogatz, *Nonlinear Dynamics and Chaos* (Addison-Wesley, Reading, 1994).

<sup>61</sup>E.G. Buré and E. N. Rozenvasser, *Autom. Remote Control (Engl. Transl.)* **7**, 1045 (1974).

<sup>62</sup>M. A. Kramer, H. Rabitz, and J. M. Calo, *Appl. Math. Model.* **8**, 328 (1984).

<sup>63</sup>B. P. Ingalls, *IEE Systems Biology* **1**, 62 (2004).

<sup>64</sup>P. Gaspar, *J. Chem. Phys.* **117**, 8905 (2002).

<sup>65</sup>J. Paulsson and M. Ehrenberg, *Q. Rev. Biophys.* **34**, 1 (2001).

<sup>66</sup>G. M. Suel, J. Garcia-Ojalvo, L. M. Liberman, and M. B. Elowitz, *Nature* **440**, 545 (2006).

<sup>67</sup>D. T. Gillespie, *J. Chem. Phys.* **81**, 2340 (1977).

<sup>68</sup>D. T. Gillespie, *J. Chem. Phys.* **113**, 297 (2000).

<sup>69</sup>R. A. Kerr, H. Levine, T. J. Sejnowski and W.-J. Rappel, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 347 (2006).

<sup>70</sup>J. Hattne, D. Fange and J. Elf, *Bioinformatics* **21**, 2923 (2005).

<sup>71</sup>M. A. Burschka, *J. Stat. Phys.* **45**, 715 (1986).