Fast Evaluation of Fluctuations in Biochemical Networks With the Linear Noise Approximation

Johan Elf and Måns Ehrenberg

Genome Res. 2003 13: 2475-2484
Access the most recent version at doi:10.1101/gr.1196503

Fast Evaluation of Fluctuations in Biochemical Networks With the Linear Noise Approximation

Johan Elf[1] and Måns Ehrenberg[1]

Department of Cell & Molecular Biology, Uppsala University, BMC, 751 24 Uppsala, Sweden

Biochemical networks in single cells can display large fluctuations in molecule numbers, making mesoscopic approaches necessary for correct system descriptions. We present a general method that allows rapid characterization of the stochastic properties of intracellular networks. The starting point is a macroscopic description that identifies the system's elementary reactions in terms of rate laws and stoichiometries. From this formulation follows directly the stationary solution of the linear noise approximation (LNA) of the Master equation for all the components in the network. The method complements bifurcation studies of the system's parameter dependence by providing estimates of sizes, correlations, and time scales of stochastic fluctuations. We describe how the LNA can give precise system descriptions also near macroscopic instabilities by suitable variable changes and elimination of fast variables.

[Supplemental material is available online at www.genome.org.]

A key element in systems biology is the design of mathematical models that faithfully describe the dynamics of intracellular chemical networks. In general, chemical reactions in single cells occur far from thermodynamic equilibrium (Keizer 1987), and the molecule copy numbers can sometimes be very small (Guptasarama 1995). Both these properties make it mandatory to analyze the stochastic behavior of chemical reactions in living systems. Most people agree with the statement that if the average number of molecules of a certain type in a cell is very small, then the relative size of the copy number fluctuations can be quite significant. This intuition is often backed up with the assertion, made famous by Erwin Schrödinger (1944) in his book, *What Is Life?*, that if the average copy number is n, then the size of the fluctuations around it must be $\sqrt{n}$ (Poisson statistics). This would mean that if the average number (n) of molecules of a certain kind in an *Escherichia coli* cell is 100 or more (corresponding to a concentration >0.2 μM; Donachie and Robinson 1987), then the relative fluctuations ($\sqrt{n}/n = 1/\sqrt{n}$) would be 10% or less. Accordingly, when there are more than, say, 1000 molecules in the cell, then the relative fluctuations would be only 3% and therefore insignificant in most instances. However, intracellular chemical reactions take place far from equilibrium. Therefore, Poisson statistics may not apply, and the stochastic fluctuations around an average of n molecules can be much larger as well as much smaller than $\sqrt{n}$ (Paulsson and Ehrenberg 2001; Elf et al. 2003b). Because Poisson statistics cannot be assumed to be valid for living systems, characterization of the stochastic properties of intracellular networks is nontrivial. The purpose of this work is to outline a method that allows for such a characterization in a straightforward and simple way.

Chemical reactions in the test tube are normally described macroscopically by (1) applying the law of mass action (Cornish-Bowden 1995; Fersht 1998) and (2) neglecting random deviations from the average number of molecules during the time evolution of the concentrations involved in the studied reactions. Such macroscopic descriptions have in the past also been used to describe intracellular chemical reactions with some notable exceptions in which stochastic reasoning has been applied (e.g., Berg 1978). More recently, however, the urgent need to take stochastic aspects of chemical reactions in single cells into account has been realized, and this insight has resulted in a burst of publications dealing experimentally as well as theoretically with the intrinsically random nature of gene expression (Becskei and Serrano 2000; Kepler and Elston 2001; Paulsson and Ehrenberg 2001; Thattai and van Oudenaarden 2001; Elowitz et al. 2002; Ozbudak et al. 2002). The focus on gene expression is explained by the fact that the copy number of individual messenger RNAs can often be very small, in particular in bacterial cells like *E. coli*, and this frequently leads to highly significant relative fluctuations in messenger RNA copy numbers and, sometimes, also to large fluctuations in protein levels.

To deal with the stochastic nature of these reactions, one has to abandon the macroscopic perspective of chemical reactions, where fluctuations are neglected, in favor of mesoscopic approaches (van Kampen 1997). One technique is to solve the chemical Master equation, which describes the time evolution of the probability distribution for all the different molecular components in the system (van Kampen 1997), either analytically or numerically. Alternatively, one can simulate the molecule number trajectories for such systems, using Monte Carlo methods, implemented by, for example, the Gillespie algorithm (Gillespie 1976, 1977). These well-established approaches are based on an assumption of homogeneous probability distributions in space, corresponding to much faster diffusion of all components throughout the system than the rates of all chemical reactions (van Kampen 1997). It is likely that this assumption will turn out to be untenable for many intracellular reactions, and attempts have been made to develop Monte Carlo methods to simulate space–time trajectories for stochastic reaction-diffusion-coupled systems (Stundzia and Lumsden 1996). Recently, we developed an efficient algorithm (Elf et al. 2003a) for such simulations, built on the Gillespie and the next reaction algorithms (Gibson and Bruck 2000), but in the present work we assume spatial homogeneity throughout.

Large relative fluctuations of copy numbers characterizes many mRNA levels in bacterial cells, but there are also other cases in which noise becomes highly significant, also when the average molecule copy numbers are very large. One example is systems with a continuum of macroscopic stationary states, like a limit

[1]Corresponding authors.
E-MAIL Johan.elf@icm.uu.se; FAX 46184714262.
E-MAIL ehrenberg@xray.bmc.uu.se; FAX 46184714262.
Article and publication are at http://www.genome.org/cgi/doi/10.1101/gr.1196503.

cycle attractor (Barkai and Leibler 2000; Qian et al. 2002). Another example is systems that operate near critical points (Dogterom and Leibler 1993; Elf et al. 2003b). In all these cases, relaxation to a steady-state solution is slow or a steady state may not exist at all, meaning that very large and slowly changing fluctuations in component copy numbers can emerge.

No simple methods for the mesoscopic analysis of such systems exist today, and numerical approaches are often difficult to carry out because of the high computational demands normally associated with this type of problem. Monte Carlo methods, such as the Gillespie algorithm (Gillespie 1977), are useful only for very limited sets of parameters, because a complete simulation must be carried out for every chosen set of parameter values.

Here, we present a general method, based on the linear noise approximation (LNA; van Kampen 1997), that allows rapid characterization of the stochastic properties of intracellular networks over large parameter regions. Our analysis is useful for initial characterization of fluctuations in all kinds of intracellular reactions. We describe the method qualitatively and also give four simple, and yet realistic, examples that clarify its principles and demonstrate its applicability. The examples highlight: (1) the importance of taking stoichiometric relations into account in mesoscopic descriptions, although they enter macroscopic models in a trivial way; (2) the fundamental connection between, on one hand, the sensitivity of a system response to a change in a flow and, on the other, the relative size of random fluctuations; (3) coupled fluctuations in metabolic networks; and (4) how the LNA can be made more accurate close to near-critical points by elimination of fast variables (Gardiner 1985).

All examples are chosen so that analytical approximations of solutions of the Master equation can be obtained, and these approximations are then used to clarify the origin of fluctuations in each case and to trace the effects of different parameters on the stochastic properties of the system.

How the present theory can be applied to arbitrarily large biochemical networks is described in the Discussion.

RESULTS

The Linear Noise Approximation

The chemical Master equation describes how the joint probability distribution of the copy numbers of different chemical species evolve in time in a spatially homogenous chemical system (van Kampen 1997). As shown in the Methods section, the Master equation can be written directly from the rate constants and stoichiometries of all the elementary reactions of a chemical system, but neither analytical nor numerical solutions are in general available. Fortunately, the Master equation can often be simplified in a linear noise approximation (LNA). This is obtained through van Kampen's Ω-expansion of the Master equation (van Kampen 1997). In this formulation, the Master equation allows for analytical solutions that are locally valid close to macroscopic trajectories or stationary points of the system. These solutions often give very accurate descriptions of the size of molecule number fluctuations and how they are correlated. They can therefore be used for rapid, initial characterizations of the stochastic properties of chemical systems as guidelines for decisions about whether the fluctuations are so large that mesoscopic considerations are necessary or whether a macroscopic approach will suffice. The Ω-expansion means that the Master equation is Taylor-expanded near macroscopic system trajectories or stationary solutions in powers of $1/\sqrt{\Omega}$, where Ω is the system volume. When the Master equation is approximated near a macroscopically stable stationary solution, terms of first order in $1/\sqrt{\Omega}$ give the macroscopic rate equations, and terms of second order in $1/\sqrt{\Omega}$ give the Linear Noise Approximation (LNA). The rationale behind this approach is that for constant average concentrations, relative fluctuations will tend to decrease with the inverse of the square root of the volume. The LNA is therefore very accurate for systems for which the local linearization of the chemical rate laws is valid for component copy numbers that are frequently reached by fluctuations away from the stationary state. This is true when fluctuations are sufficiently small in relation to the component means. However, we have found that the LNA can give very good estimates of fluctuations in molecule numbers, also when they are larger than the corresponding means, and, furthermore, that the technique of separation of fast variables (see below) can be used to improve further the performance of the LNA.

Here, the LNA is derived for the stationary state of a general system of chemical reactions (Supplemental material available online at www.genome.org; Methods). The correlation matrix, **C**, containing the variances and covariances for the fluctuations of all the components in the system, is calculated from the linearized macroscopic rate equations near the stationary point as described by the Jacobian matrix, **A**, and the stoichiometries of the reactions and their rates as described by the diffusion matrix, **D**. The relation between the matrix **C**, containing information about the fluctuations, and the macroscopic matrices **A** and **D** is given by a Lyapunov matrix equation, equation 32.

$$\mathbf{AC} + \mathbf{CA}^T + \mathbf{D} = \mathbf{0} \tag{1}$$

How to calculate the **A** and **D** matrices from the chemical reactions and how to solve the equation for **C** is described in Methods. Here we focus on a few special cases in which the LNA is used to understand the origin of noise in biochemical processes.

The first example is a system that contains only one type of molecule, X. These molecules are synthesized and consumed by two different reactions with rate laws $f_1(\phi)$ and $f_2(\phi)$, respectively. The reaction rates are given in molar concentration per second and ϕ is the molar concentration of X. The first reaction adds n molecules to X in a single event, and the second reaction removes m molecules from X in a single event:

$$\varnothing \xrightarrow{f_1(\phi)} nX$$
$$mX \xrightarrow{f_2(\phi)} \varnothing \tag{2}$$

The macroscopic rate equation becomes

$$\frac{d\phi}{dt} = nf_1(\phi) - mf_2(\phi) \tag{3}$$

When equation 3 is linearized around the steady state, $\bar{\phi}$, the equation for the deviation $\delta\phi$ from steady state is

$$\frac{d\delta\phi}{dt} = \underbrace{(nf_1'(\bar{\phi}) - mf_2'(\bar{\phi}))}_{\lambda}\delta\phi = \lambda\delta\phi \tag{4}$$

λ is the compounded rate constant (in units of $\sec^{-1}$) by which the system relaxes back to steady state after a perturbation. From equation 37 in Methods, the LNA approximation C of the variance σ_X^2 for the number of X molecules at the steady state is

$$C = \Omega\frac{(n^2 + m^2)f_1(\bar{\phi})}{2\lambda} \tag{5}$$

The variance σ_X^2 normalized to the average number ($\langle X\rangle$) of molecules is often called the Fano-factor (Fano 1947). Open systems at thermodynamic equilibrium obey Poisson statistics, and for those the Fano-factor is always 1. For the creation–elimination scheme in equation 2, the Fano-factor is approximated by

$$\frac{\sigma_X^2}{\langle X\rangle} \approx \frac{C}{\Omega\bar{\phi}} = \frac{f_1(\bar{\phi})}{\bar{\phi}} \frac{1}{\lambda} \frac{(n^2+m^2)}{2} \tag{6}$$

To obtain this expression, we have also approximated $\langle X\rangle$ with the macroscopic steady-state value $\Omega\bar{\phi}$ of the number of X molecules in the steady state. The numerical magnitude of the Fano-factor is primarily determined by the ratio between the rate $f_1(\bar{\phi})/\bar{\phi}$ of turnover of all the molecules in the X-pool and the rate of relaxation λ. The Fano-factor is also shaped by the creation (n) and elimination (m) stoichiometries. Below we use equation 6 to inspect the role of the stoichiometries (Example 1) and the role of the relaxation rate λ (Example 2) for the stochastic properties of the scheme in equation 2.

Example 1: The Mesoscopic Importance of Stoichiometries in Chemical Reactions

At the macroscopic description level, one can often compensate for a change in the stoichiometric order of a reaction by a change in a rate constant, so that the dynamics of the reaction remains unchanged. At the mesoscopic level, this cannot be done.

To illustrate the role of the stoichiometries, we compare two different reaction systems (A and B) with different reaction stoichiometries:

$$\text{A}\begin{cases} \varnothing \xrightarrow{k_1} nX \\ 2X \xrightarrow{k_2\Phi^2} \varnothing \end{cases} \qquad \text{B}\begin{cases} \varnothing \xrightarrow{nk_1} X \\ 2X \xrightarrow{k_2\Phi^2} \varnothing \end{cases} \tag{7}$$

Both systems have identical macroscopic differential equations and steady-state values:

$$\frac{d\phi}{dt} = nk_1 - 2k_2\phi^2 \qquad \Rightarrow \qquad \bar{\phi} = \sqrt{nk_1/2k_2} \tag{8}$$

From equation 5, the following approximations for the variances in the two cases are obtained:

$$C_A = \frac{(n+2)}{4}\Omega\bar{\phi} \qquad C_B = \frac{3}{4}\Omega\bar{\phi}. \tag{9}$$

The Fano-factors (equation 6) for systems A and B are $(n+2)/4$ and 3/4, respectively. For system A, the Fano-factor increases linearly with the stoichiometry of the creation reaction, whereas the corresponding increase in the rate constant (but not the stoichiometry) for system B leaves the Fano-factor unchanged. This example is described in more detail and with general stoichiometries in the Supplemental material.

Example 2: Zero-Order Sensitivity and the Size of Mesoscopic Fluctuations

Goldbeter and Koshland Jr. (1981) pointed out that the extent of modification in certain enzymatically catalyzed modification–demodification reactions can be ultrasensitive to the amount of either the modifying or the demodifying enzyme. It was also shown (Berg et al. 2000) that such systems can display large fluctuations, which will tend to attenuate the macroscopically predicted ultrasensitivity.

We exemplify this by treating the case with a protein that is converted from its unmodified state X_1 to its modified state X_2 by one enzyme and back to X_1 from X_2 by another. With a constant total concentration ϕ_0 of modified and unmodified target concentrations, the state of the system is completely described by the concentration ϕ_1 of X_1, because the concentration ϕ_2 of X_2 is given by the conservation relation $\phi_0 = \phi_1 + \phi_2$. The reaction scheme is

$$X_1 \underset{f_2(\phi_1)}{\overset{f_1(\phi_1)}{\rightleftharpoons}} X_2 \tag{10}$$

The macroscopic rates of modification and demodification are $f_1(\phi_1)$ and $f_2(\phi_1)$, respectively. The modifying enzymes obey Michaelis-Menten kinetics, and when steady-state relations for their internal states are rapidly established (see discussion about elementary complex reactions in Methods), the rate equations are given by

$$f_1(\phi_1) = \frac{V_{\max 1}\phi_1}{K_m + \phi_1} \qquad f_2(\phi_1) = \frac{V_{\max 2}(\phi_0 - \phi_1)}{K_m + (\phi_0 - \phi_1)} \tag{11}$$

The same K_m value is assumed for both enzymes, and the fractions of enzyme-bound target molecules are assumed to be negligible. The macroscopic dynamics is thus given by

$$\frac{d\phi_1}{dt} = f_2(\phi_1) - f_1(\phi_1) \tag{12}$$

with the steady-state condition

$$f_1(\bar{\phi}_1) = f_2(\bar{\phi}_1). \tag{13}$$

The macroscopic stationary value $\Omega\bar{\phi}_1$ is the LNA estimate of the average number of X_1 molecules. In Figure 1A, this estimate is compared with the correct value as calculated directly from the Master equation, for different values of $V_{\max 2}$. The sensitivity amplification (logarithmic gain; Savageau 1976) of the system is

$$a_{\phi_1 V_{\max 2}} = \left.\frac{d\phi_1/\phi_1}{dV_{\max 2}/V_{\max 2}}\right|_{\bar{\Phi}} \tag{14}$$

The sensitivity amplification is the percentage change in the concentration of ϕ_1 in response to the percentage change in $V_{\max 2}$. The system can be said to be ultrasensitive if $|a_{\phi_1 V_{\max 1}}| > 1$ (Goldbeter and Koshland Jr. 1982). If $\phi_0 \gg K_m$, both enzymes are close to saturation when $V_{\max 1} = V_{\max 2}$ and the reaction displays ultrasensitivity (Goldbeter and Koshland Jr. 1981), because of zero-order kinetics in both the modification and demodification reactions. Zero-order kinetics means that the flow in either direction is insensitive to the concentrations of modified and unmodified substrates, so that the "force" that restores the system to its steady state is very weak.

The Master equation for the probability of having X_1 unmodified molecules can be solved analytically in special cases (Berg et al. 2000). The LNA approximation of the Fano-factor is (equation 6)

$$\frac{C}{\Omega\bar{\phi}_1} = \frac{1}{\bar{\phi}_1}\frac{f_1(\bar{\phi}_1)}{f_1'(\bar{\phi}_1) - f_2'(\bar{\phi}_1)} = a_{\phi V_{\max 2}} \tag{15}$$

The equivalence between the sensitivity in equation 14 and the Fano-factor in equation 15 is seen by implicit derivation of the steady-state equation (equation 13; Elf et al. 2003b):

$$\left.\frac{d\phi_1/\phi_1}{dV_{\max 2}/V_{\max 2}}\right|_{\bar{\Phi}} = \frac{V_{\max 2}}{\phi_1}\left.\frac{\partial f_2/\partial V_{\max 2}}{\partial f_2/\partial\phi_1 - \partial f_1/\partial\phi_1}\right|_{\bar{\Phi}} = \frac{f_2(\bar{\phi}_1)}{\bar{\phi}_1}\frac{1}{f_1'(\bar{\phi}_1) - f_2'(\bar{\phi}_1)} \tag{16}$$

The maximal value of the Fano-factor, $(2K_m + \phi_0)/(2K_m)$, is obtained when $V_{\max 1} = V_{\max 2}$. The LNA of the Fano-factor (equa-

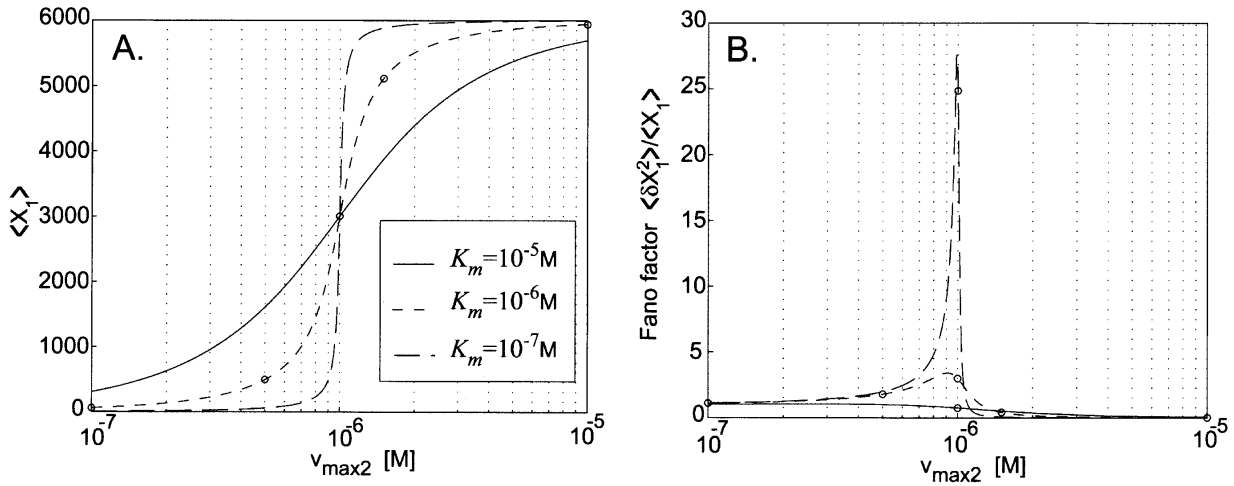

Figure 1 Properties of the modification–demodification system (one-component system). (*A*) Average number of molecules in X_1 state, $\langle X_1 \rangle$. (*B*) The Fano-factor ($C/\langle X_1 \rangle$). The lines correspond to the LNA equation 15, the circles to values from the numerical solution of the corresponding Master equation. The total concentrations of X_1 and X_2 are 10^{-5} M and $v_{max1} = 10^{-6}$ M/sec.

tion 15) is plotted in Figure 1B for increasing V_{max2} and is compared with values calculated from the exact Master equation for some parameter sets. The comparison reveals that the LNA is very accurate in these cases.

The Fano-factor are much larger than 1, as seen in Figure 1B. The reason for these large fluctuations and the strong deviation from Poisson statistics (Fano-factor 1), is high turnover rates of the molecules at steady state in combination with slow relaxation back to steady state. This is exactly what is described in equation 15, where f_1/ϕ_1 is the turnover rate of the pool and $f_1'(\phi_1) - f_2'(\phi_1)$ is the rate of relaxation. The high turnover rate creates a lot of chemical "noise" (large diffusion) that decays very slowly (weak restoring force), and this leads to very large fluctuations.

Example 3: Coupled Fluctuations in Multiple Component Systems

Similar principles apply to systems with more than one molecular species; that is, slow relaxations, high turnover rates, and large jumps in molecule numbers (high reaction stoichiometries) cause large fluctuations. The degrees of freedom are, however, larger in multicomponent systems, and intriguing stochastic phenomena of a novel type arise because of correlations between the fluctuations of different components. For instance, fluctuations in one species (e.g., mRNA) can induce fluctuations in another species (e.g., protein). Furthermore, fluctuations can bring the molecule number of one species to a very high value while compensating fluctuations can bring the molecule numbers of another species to a very low value in such a way that the flow properties of the system remain almost unchanged. The latter case is relevant for many biosynthetic reactions, and we will therefore discuss a simple example illustrating this type of degenerate behavior.

Although the Master equation is usually impossible to solve for multicomponent systems, the LNA can be used to estimate the covariance matrix directly from the macroscopic rate equations and the stoichiometry matrix as described in Methods.

To illustrate, we consider an anabolic system, where, respectively, molecules X and Y are synthesized from reservoirs A and B at rates k_1 and k_2. X and Y are assumed to associate irreversibly with association rate constant k_5 in the formation of a heterodimer C. The molecules X and Y also decay with first-order rate constants k_3 and k_4, respectively, as described by the scheme:

$$A \underset{k_3\phi_1}{\overset{k_1}{\rightleftarrows}} X \qquad B \underset{k_4\phi_2}{\overset{k_2}{\rightleftarrows}} Y \qquad X + Y \xrightarrow{k_5\phi_1\phi_2} C \tag{17}$$

The concentrations of the A, B, and C molecules are assumed to be constant, and the concentrations of X and Y are ϕ_1 and ϕ_2, respectively. This general system is analyzed in the Supplemental material. Here, we focus on the most interesting parameter regime, where X and Y are synthesized at equal rates ($k_1 = k_2 = \kappa$), and decay with equal and low rates ($k_3 = k_4 = \mu$), that is,

$$A \underset{\mu\phi_1}{\overset{\kappa}{\rightleftarrows}} X \qquad B \underset{\mu\phi_2}{\overset{\kappa}{\rightleftarrows}} Y \qquad X + Y \xrightarrow{k_5\phi_1\phi_2} C \tag{18}$$

Low decay rates imply that the degradation events are rare compared to the rate of formation of C molecules, that is, $\mu \ll \sqrt{\kappa k_5}$. This, as we shall see, is the condition for anomalously large component number fluctuations.

The macroscopic differential equations for the concentrations and their steady-state values are

$$\left.\begin{aligned} \frac{d\phi_1}{dt} &= \kappa - k_5\phi_1\phi_2 - \mu\phi_1 \\ \frac{d\phi_2}{dt} &= \kappa - k_5\phi_1\phi_2 - \mu\phi_2 \end{aligned}\right\}$$

$$\Rightarrow \begin{aligned} \mu > 0: &\quad \bar{\phi}_1 = \bar{\phi}_2 = \sqrt{\frac{\kappa}{k_5} + \frac{\mu^2}{k_5^2}} - \frac{\mu}{2k_5} \approx \sqrt{\frac{\kappa}{k_5}} \\ \mu = 0: &\quad \bar{\phi}_1 = \frac{\kappa}{\bar{\phi}_2 k_5} \end{aligned} \tag{19}$$

For $\mu = 0$, the system has infinitely many (neutrally stable) fixed points on the curve $\bar{\phi}_1 = \kappa/\bar{\phi}_2 k_5$. The stochastic reaction events will make the concentrations fluctuate freely (free diffusion) on this curve, such that the variance is infinite.

When $0 < \mu \ll \sqrt{\kappa k_5}$, the fluctuations may still be very large, as can be seen in the Monte Carlo simulation (Fig. 2C). This is again a close-to-zero-order phenomenon, but now in two dimensions. The fluctuations in Figure 2C are anticorrelated such that the major consumption term, $k_5\phi_1\phi_2$, is close to constant, that is, both synthesis and consumption rates are near zero order.

The primary source for the anomalously large fluctuations and the highly correlated fluctuations is therefore, again, high turnover of the molecule pools compared with the (decay) rate of

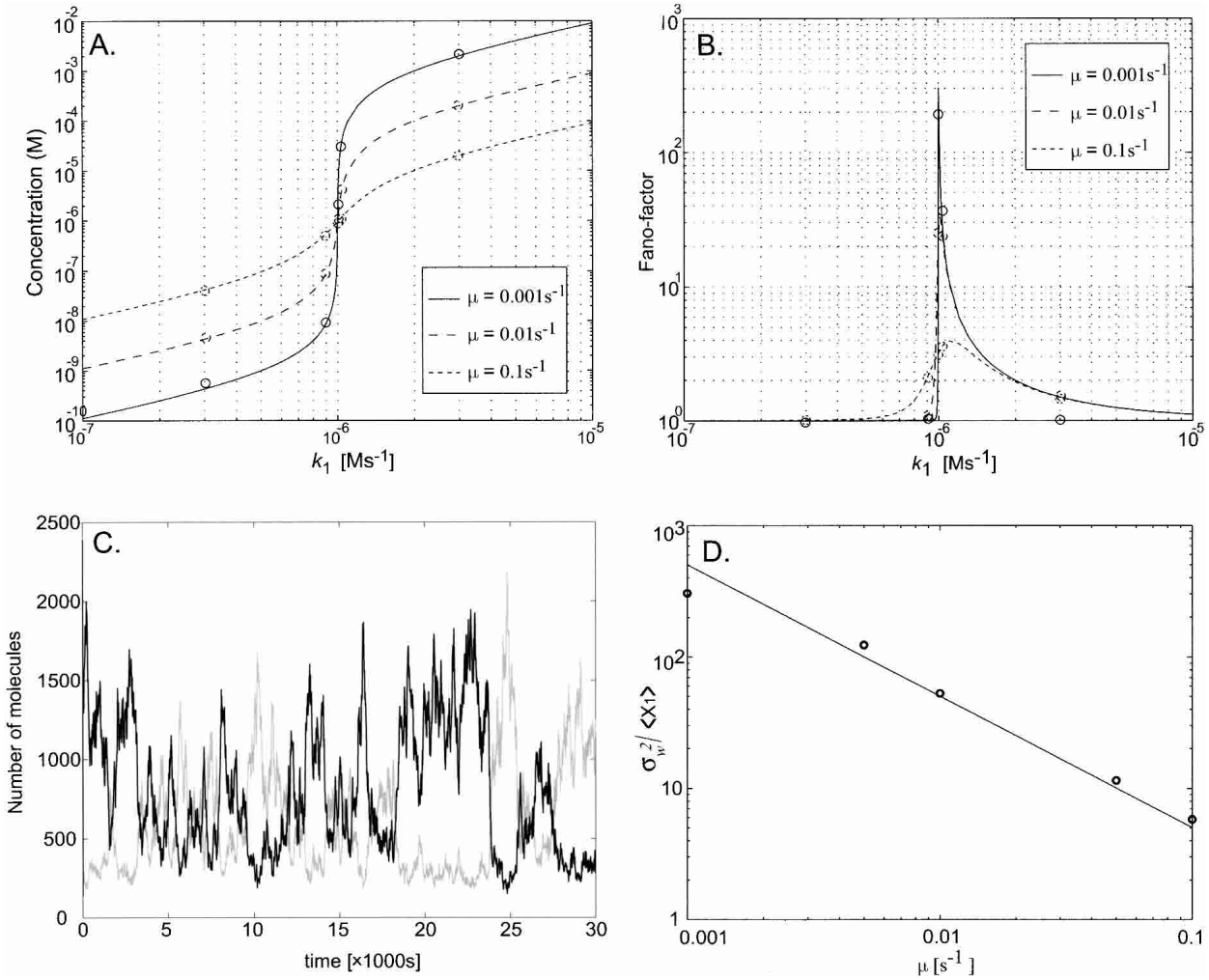

Figure 2 Properties of a system with coupled exit flow (two-component system, equation 17). (*A*) The average concentrations (*Y*-axis) of X molecules for an increasing rate of X synthesis (k_1, *X*-axis). Lines correspond to LNA and circles to estimates from Gillespie simulations. The synthesis rate of Y molecules is $k_2 = 10^{-6}$ M/sec, and the association rate constant is $k_5 = 10^6$ M/sec. These parameters are also used in *B–D*. (*B*) The Fano-factor for X molecules, $\langle \delta X^2 \rangle / \langle X \rangle$, as given by LNA (lines) and Gillespie simulation (circles). (*C*) Gillespie simulation of the variation of the number of X and Y molecules (*Y*-axis) over time (*X*-axis) for balanced synthesis (i.e., $k_1 = k_2 = 10^{-6}$ M/sec). (*D*) The modified Fano-factor, $\sigma_w^2 / \langle \Omega X_1 \rangle$, at the point of balanced synthesis (*Y*-axis) for different values of the decay rate (*X*-axis) as given by LNA (lines) and Gillespie simulation (circles).

relaxation to steady state. This is quantified in the LNA of the system as described in Methods.

In the LNA of the two-component system, given by equation 18, the fluctuations are analyzed in the variables $w = \phi_1 - \phi_2$ and $u = \phi_1 + \phi_2$. This change of variables follows automatically with our method (see Supplemental material). The advantage of using these variables is seen by subtracting the differential equations in equation 19 to get the differential equation for $w = \phi_1 - \phi_2$:

$$\frac{d\phi_1}{dt} - \frac{d\phi_2}{dt} = \frac{dw}{dt} = -\mu(\phi_1 - \phi_2) = -\mu w \quad \Rightarrow \quad w(t) = ce^{-\mu t} \tag{20}$$

The difference between X and Y decays at a rate set by μ, which is small compared with the turnover rate of the X and Y pools ($\kappa/\bar{\phi}$). The variance in $w(\sigma_w^2)$ as calculated by the LNA, normalized to the number of molecules in one of the pools at steady state, is

$$\frac{\sigma_w^2}{\Omega\bar{\phi}} = \left(\frac{\kappa}{\bar{\phi}}\frac{1}{\mu} + 1\right). \tag{21}$$

This expression is exact, except for the small difference between a true concentration average and a macroscopic stationary state value. (The exact expression can be derived with moment-generating functions.) It can now be seen that it is the large separation between the turnover rate of the X and Y pools ($\kappa/\bar{\phi}$) and the relaxation rate μ that causes the very large fluctuations in *w*.

Results from the LNA of the full system (equation 17) over large parameter regions are shown in Figure 2. The LNA is compared with solutions of the Master equation at some points. A major advantage of the LNA is that the stochastic properties of a system can be rapidly scanned over a large parameter space at a very modest computational cost.

Example 4: Elimination of Fast Variables

In Example 3, large fluctuations arise because of one very slow relaxation mode in the coupled system. This behavior is typical of systems near critical points, like phase transitions (Haken 1982). In Methods, we describe a method to make the LNA more accurate in these cases. The method requires that the dynamics of

the system is mainly determined by a few slow variables, which are linear combinations of the original concentrations, and that the fast variables will be at steady-state values, conditional on the present values of the slow variables. This approach gives precise estimates of the stochastic system behavior when the rate laws are approximately linear in the slow variables (but not necessarily in the fast variables), and gives a much more accurate description of the system dynamics than direct usage of the LNA.

It should be noted that this technique does not mean that one just assumes rapid equilibration among some of the molecular species. The number of relevant slow variables can, in fact, be lower than such a direct elimination would presuppose. To illustrate the principle, we exemplify with a four-component system, in which the large fluctuations are contained in just one slow variable. The example is an extended version of the system discussed in Example 3 above:

$$\mathrm{E_A} \xrightarrow{k_1 e_A} \mathrm{E_A} + \mathrm{A} \qquad \mathrm{E_B} \xrightarrow{k_4 e_B} \mathrm{E_B} + \mathrm{B}$$

$$\mathrm{A} + \mathrm{E_A} \underset{k_3 (e_A^{tot} - e_A)}{\overset{k_2 a e_A}{\rightleftharpoons}} \mathrm{AE_A} \qquad \mathrm{B} + \mathrm{E_B} \underset{k_6 (e_B^{tot} - e_B)}{\overset{k_5 b e_B}{\rightleftharpoons}} \mathrm{BE_B}$$

$$\mathrm{A} + \mathrm{B} \xrightarrow{k_7 ab} \mathrm{AB} \qquad (22)$$

A and B molecules are synthesized by enzymes E_A and E_B, respectively. A and B can bind E_A and E_B, and thereby inhibit their own synthesis, and can also bind each other irreversibly. Lower-case letters correspond to concentrations. The stationary, marginal distribution $P(B)$ (see Methods for a definition of marginal distribution) for the B molecules is estimated by three different methods and the results are shown in Figure 3. The distribution estimated by the Gillespie simulation (gray line) of the full system (equation 22) has similar mean and variance as the normal distribution resulting from the LNA, but the shape of the latter (black line) is very different from the more accurate distribution obtained from the Gillespie simulations. However, this failure of directly applying the LNA can be overcome by elimination of fast variables (see Supplemental material). Figure 3 shows that the elimination of all but one slow variable and application of the LNA gives a much better representation of $P(B)$ than direct application of the LNA to the four-variable system (dashed line).

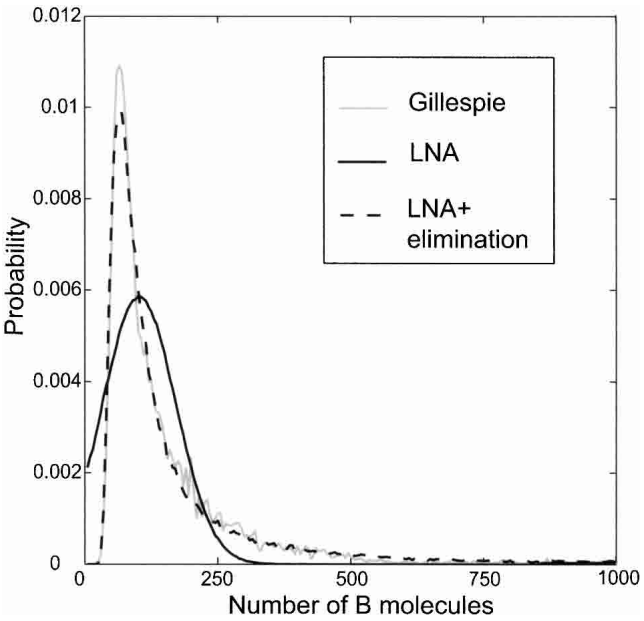

Figure 3 LNA combined with elimination of fast variables (four-component system, equation 22). The stationary marginal distribution for the number of B molecules, $P(B)$, has been estimated by three different methods: (1) by Gillespie simulation (gray line), which converges toward the correct solution; (2) direct application of LNA (solid black line), which approximates the distribution with a normal distribution; and (3) LNA combined with elimination of fast variables (dashed black line). Parameters: $\Omega = 10^{-15}l$; $k_1 = 200$ sec^{-1}; $k_2 = 1.67 \times 10^5$ M^{-1} sec^{-1}; $k_3 = 10{,}000$ sec^{-1}; $k_4 = 100$ sec^{-1}; $k_5 = 1.67 \times 10^5$ M^{-1} sec^{-1}; $k_6 = 1000$ sec^{-1}; $k_7 = 1.67 \times 10^6$ M^{-1} sec^{-1}; $e_A^{tot} = 83$ nM; $e_B^{tot} = 167$ nM.

DISCUSSION

The results in this work follow in a straightforward way from the mesoscopic treatises by van Kampen (1997), Gardiner (1985), and Keizer (1987). However, these seminal contributions to stochastic kinetics primarily deal with simple chemical reactions in large volumes so that relative fluctuations are small. The focus of the present study is, in contrast, on biochemical networks with elementary complex reactions (Keizer 1987) in the small volumes of single cells. Here, relative fluctuations can be very large, and general analytical approaches have been absent.

We have outlined a direct path (Methods) from macroscopic descriptions of intracellular chemical networks to the linear noise approximations (van Kampen 1997) of the Master equations associated with these systems. The starting point was to formulate the macroscopic dynamics in terms of elementary or elementary complex (Keizer 1987) reactions where the stoichiometries of the different reactions are made explicit. Special attention has been paid to the stochastic properties of chemical networks that are in macroscopically stable stationary states, where the LNA corresponds to a linear Fokker-Planck equation for the movement of a Brownian particle in a multidimensional harmonic potential (Wax 1954). Its solution (Risken 1984) evolves to a multidimensional normal distribution for deviations in molecule numbers from their macroscopic values. The stationary solution of the LNA is completely characterized by its covariance matrix and the Jacobian matrix associated with the macroscopic stationary state. We have used well-known results about the Lyapunov matrix equation (Horn and Johnson 1991) to derive an explicit expression for the covariance matrix in terms of the eigenvalues of the Jacobian matrix, from the macroscopic system at steady state, and the diffusion matrix. The analytical expressions for the cross- and autocorrelation functions of the stationary solution to the linear Fokker-Planck equation of the Master equation have also been described.

This type of analysis facilitates screening of any intracellular network with respect to its stochastic properties. The characteristics of the covariance and Jacobian matrices will decide when fluctuations are so small and fast that a system can be analyzed macroscopically. They will reveal the existence of anomalously large fluctuations that require mesoscopic approaches for precise system descriptions.

Our results show how giant fluctuations in multidimensional systems can arise owing to large terms in the diffusion matrix (large turnover rates of component pools that generate internal noise) and small eigenvalues of the Jacobian matrix (weak restoring forces). When there is a single stable, but weakly attracting, stationary state so that diffusion is almost free in its neighborhood, very large relative fluctuations can occur for systems of finite size. A striking example is the behavior of anabolic reactions in which two or more substrates are independently synthesized and then irreversibly joined by an unsaturated enzyme. When the rates of synthesis of all substrates are balanced, all combinations of substrate concentrations that are compatible with the joint exit flow are macroscopically allowed, and there is

free diffusion throughout this continuum of states. In growing systems or in systems in which the rates of synthesis are attenuated by feedback or product inhibition there will be a single stationary state, but it is often so weakly constrained that highly significant fluctuations prevail (Elf et al. 2003b).

Our discussion of the LNA complements existing numerical tools (Sauro 1993; Mendes 1997) for the analysis of stability properties of stationary states and bifurcations (Strogatz 1994) in kinetic reaction networks. These tools normally make use of the Jacobian matrix and its eigenvalues. The explicit formulation of the diffusion matrix from the Liapunov equation now makes it easy to obtain also mesoscopic system properties near stationary points.

Our general analysis, exemplified by concrete examples of how anomalous fluctuations arise in intracellular networks, has also been extended beyond direct applications of the LNA to intracellular systems. When fluctuations are large compared with the range of validity of linear rate laws, then the LNA is expected to fail in many cases. We have analyzed systems in which the macroscopic stationary states are almost degenerate. In those cases, the eigenvalues of the Jacobian are very small, although the flows through the system are significant. Here, large fluctuations are associated with slow relaxation modes, and we have used this system property to systematically eliminate fast variables (Gardiner 1985) with small fluctuations. This approach improves the performance of the LNA in cases in which the slow variables can be described by linear rate laws, although the fast variables display nonlinear dynamics.

We have described several cases in which the LNA or the LNA improved by elimination of fast variables gives accurate mesoscopic system descriptions. However, in cases in which fluctuations bring a system far from the region of validity of the local LNA, a more detailed mesoscopic analysis is required. This is, for instance, the case in systems with noise-induced transitions (Horsthemke and Lefever 1984), which is exemplified (Fig. 4) by how the LNA performs for a system with noise-induced oscillations (Vilar et al. 2002). In the parameter region where an asymptotically stable macroscopic steady state exists, the LNA works, but where periodic oscillations occur because of random fluctuations, the LNA fails to reproduce the oscillations. However, the LNA correctly predicts a marked increase in the size of the fluctuations at the point where the stochastically induced oscillations begin to emerge (Fig. 4).

The linear Fokker-Planck equation used here to discuss the stochastic properties of intracellular chemical networks can be reformulated in terms of stochastic differential equations (Methods). This formulation is mathematically equivalent to the Fokker-Planck equations and allows direct comparisons of the effects of intrinsic noise in intracellular chemical networks, on the one hand, with the consequences of external noise on technical feedback systems, on the other (Glad and Ljung 2000).

METHODS

Macroscopic and Mesoscopic Dynamics

Consider an intracellular biochemical system with volume Ω and N different chemical components. The concentrations of the components are summarized in the vector $\mathbf{x} = (x_1, \ldots, x_N)^T$ and the numbers of molecules in $\mathbf{X} = \Omega\mathbf{x} = (X_1, X_2, \ldots, X_N)^T$. The state of the system is defined by $\mathbf{X}$ (or $\mathbf{x}$), and a state change takes place by one of R elementary reactions. The probability that the elementary reaction number j will occur in a small time interval δt is given by $\Omega\tilde{f}_j(\mathbf{x}, \Omega)\delta t$, where $\tilde{f}_j(\mathbf{x}, \Omega)$ is a transition rate. By such an event the chemical component number i changes from X_i to $X_i + S_{ij}$ molecules. The integers S_{ij}, $i = 1, 2, \ldots, N$; $j = 1, 2, \ldots, R$ are the elements of the $N \times R$ stoichiometric matrix $\mathbf{S}$ of the reaction network (Heinrich and Schuster 1996).

When the system volume Ω increases further and further, the relative fluctuations in the network will eventually become insignificant. In the limit of infinitely large Ω, the stochastic concentration vector $\mathbf{x}$ will normally become deterministic, and the transition rates $\tilde{f}_j(\mathbf{x}, \Omega)$ will simplify to their macroscopic rate law counterparts $f_j(\boldsymbol{\varphi})$. We define

$$\boldsymbol{\varphi} = \lim_{\substack{\Omega\to\infty \\ \mathbf{x}\to\infty}} \Omega^{-1}\mathbf{X} = \lim_{\Omega\to\infty} \mathbf{x},$$

$$\lim_{\Omega\to\infty} \tilde{f}_j(\mathbf{x},\Omega) = f_j(\boldsymbol{\varphi}) \qquad (23)$$

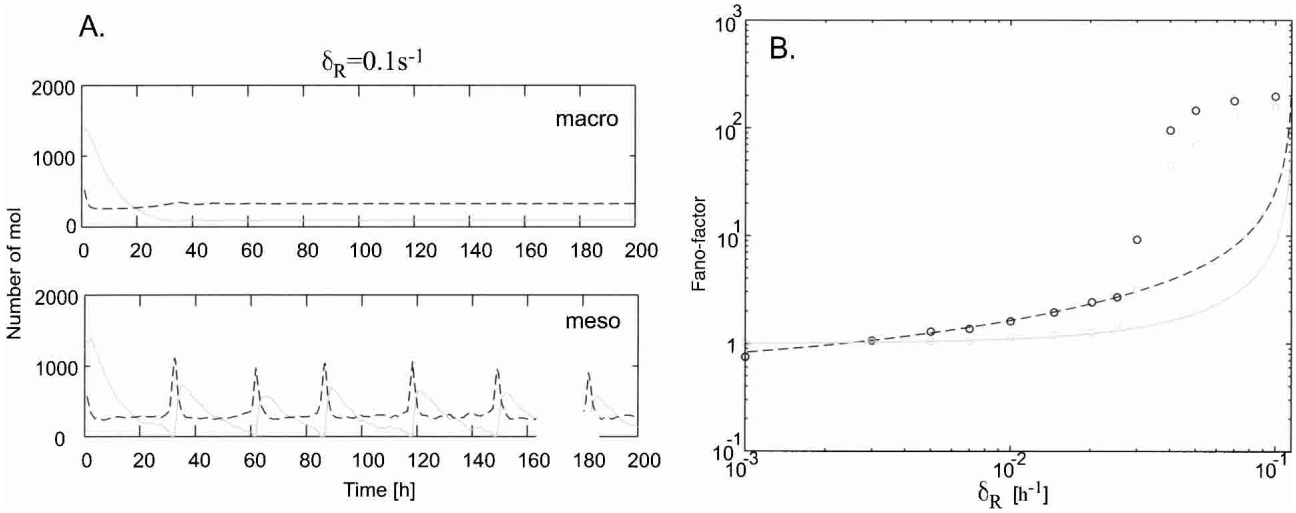

Figure 4 Example of failure of local linearization and LNA. This figure demonstrates the application of LNA to the system described by Vilar et al. (2002). We use the same equations and parameters as in the original article. In a certain parameter regime, the mesoscopic system displays periodic oscillation, whereas the macroscopic model describes an asymptotically stable steady state (*A*). As can be seen in *B*, the LNA well describes the Fano-factor for the repressor and repressor–inhibitor complex up to a certain value for the decay rate for the repressor ($\delta_R \approx 0.0275h^{-1}$). Above this point, the system displays large stochastic fluctuations and for even higher values ($\delta_R > 0.05h^{-1}$) also periodic oscillations that cannot be described by the LNA. However, the large fluctuations predicted by LNA indicate that a more thorough stochastic analysis should be conducted. At the bifurcation, at $\delta_R \approx 0.101h^{-1}$, the macroscopic steady state loses stability and the oscillations are described also by a macroscopic model.

Here, $\boldsymbol{\varphi} = (\phi_1, \ldots, \phi_N)^T$ is the vector of macroscopic concentrations for which the dynamics is determined by N ordinary rate equations (Cornish-Bowden 1995), according to

$$\frac{d\phi_i}{dt} = \sum_{j=1}^{R} S_{ij} f_j(\boldsymbol{\varphi}) \qquad i = \{1, 2, \ldots, N\}$$

$$\Leftrightarrow \dot{\boldsymbol{\varphi}} = \mathbf{S}\mathbf{f}(\boldsymbol{\varphi}) \qquad \mathbf{f}(\boldsymbol{\varphi}) = (f_1(\boldsymbol{\varphi}), \ldots, f_R(\boldsymbol{\varphi}))^T \tag{24}$$

The Master equation (van Kampen 1997) for the system determines the time evolution of the probability distribution $P(\mathbf{X}, t)$ for the molecule number vector $\mathbf{X}$ and can be written in compact form as

$$\frac{dP(\mathbf{X},t)}{dt} = \Omega \sum_{j=1}^{R} \left(\prod_{i=1}^{N} E^{-S_{ij}} - 1 \right) \tilde{f}_j(\mathbf{x},\Omega) P(\mathbf{X},t), \tag{25}$$

where $E^{-S_{ij}}$ is a step operator. When it acts on a function $g(X_1, \ldots, X_i, \ldots, X_N)$, it removes S_{ij} molecules from X_i, that is, $E^{-S_{ij}} g(X_1, \ldots, X_i, \ldots, X_N) = g(X_1, \ldots, X_i - S_{ij}, \ldots, X_N)$ and $E^{-S_{ij}} E^{-S_{kj}} g(\ldots, X_i, \ldots, X_k, \ldots) = g(\ldots, X_i - S_{ij}, \ldots, X_k - S_{kj}, \ldots)$. Equation 24 is the complete macroscopic system description and equation 25 is the full Master equation for the same reaction network. Both equations 24 and 25 assume spatial homogeneity by fast diffusion of all components. When this is not the case, the Master equation has to be extended (Elf et al. 2003a). The stoichiometric matrix $\mathbf{S}$ appears in both equations 24 and 25, whereas the mesoscopic rate law $\tilde{f}_j$ in equation 25 is replaced by its macroscopic counterparts f_j in equation 24. How the macroscopic equation 24 emerges from the mesoscopic equation 25 when the system volume goes to infinity is shown in the derivation of the linear noise approximation in the Supplemental material (equation A8).

A crucial point in the mesoscopic description is correct identification of the elementary reactions in equations 24 and 25. Elementary means that the transition rates $\tilde{f}_j(\mathbf{x}, \Omega)$ should be constant as long as $\mathbf{x}$ does not change, that is, the rates do not depend on variables not included in the state description. A reaction can be "elementary complex" (Keizer 1987) as when it is catalyzed by an enzyme with internal states that equilibrate on a much shorter time scale than the time scale for product formation. Conversion of a substrate X_1 to a product X_2, carried out by an enzyme with Michaelis-Menten kinetics, is under such conditions elementary complex and can be characterized by the compounded transition rate $\tilde{f}(x_1) = v_{\max} x_1/(K_m + x_1)$.

Linearization of Macroscopic Dynamics and the Linear Noise Approximation

In this section we first recapitulate selected parts of the standard treatment (Strogatz 1994) of linearized dynamics around a stationary state and then derive an analytical solution to the linear noise approximation of the Master equation, equation 25. If a stationary state exists for the macroscopic dynamics in equation 24, it follows from the algebraic system of equations:

$$\mathbf{0} = \mathbf{S}\mathbf{f}(\boldsymbol{\varphi}) \tag{26}$$

Linearization of equation 24 around the stationary solution vector $\bar{\boldsymbol{\varphi}} = (\bar{\phi}_1, \bar{\phi}_2, \ldots, \bar{\phi}_N)$ associated with equation 26 leads to a matrix equation for the deviations $\boldsymbol{\delta\varphi} = (\delta\phi_1, \ldots, \delta\phi_N)$ from $\bar{\boldsymbol{\varphi}}$ according to

$$\dot{\boldsymbol{\delta\varphi}} = \mathbf{A}\boldsymbol{\delta\varphi}. \tag{27}$$

$\mathbf{A}$ is the Jacobian matrix, with the elements

$$A_{ik} = \left. \frac{\partial (S_{i\cdot} \mathbf{f}(\boldsymbol{\varphi}))}{\partial \phi_k} \right|_{\varphi = \bar{\varphi}} \tag{28}$$

When $\mathbf{A}$ can be written as $\mathbf{A} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^{-1}$, where $\boldsymbol{\Lambda}$ is a diagonal matrix with the eigenvalues $(\lambda_1, \lambda_2, \ldots, \lambda_N)$ of $\mathbf{A}$ as elements and $\mathbf{Q}$ is a matrix with the eigenvectors of $\mathbf{A}$ in the columns, the solution to equation 27 is

$$\boldsymbol{\delta\varphi} = \mathbf{Q} e^{\Lambda t} \mathbf{Q}^{-1} \boldsymbol{\delta\varphi}(0), \tag{29}$$

where $\boldsymbol{\delta\varphi}(0)$ is the initial condition (Strang 1988). The linear transformation $\tilde{\boldsymbol{\delta\varphi}} = \mathbf{Q}^{-1}\boldsymbol{\delta\varphi}$, where $\tilde{\boldsymbol{\delta\varphi}} = (\delta\tilde{\phi}_1, \ldots, \delta\tilde{\phi}_N)$, leads to a new set of variables (normal modes) with the property that

$$\delta\tilde{\phi}_i(t) = e^{\lambda_i t}\, \delta\tilde{\phi}_i(0) \qquad i = 1, 2..N \tag{30}$$

The Linear Noise Approximation

The mesoscopic counterpart to the macroscopic solutions in equations 29 or 30, the linear noise approximation (van Kampen 1997), follows from the Master equation, equation 25, by linearizing the rate laws as in equation 27 and by approximating its difference relations by differentials. To our knowledge, the general multidimensional LNA has not been derived in the literature and is therefore given in the Supplemental material. In the LNA, the probability distribution for the vector $\boldsymbol{\delta X}$, describing random deviations in molecule numbers from their macroscopic values $\Omega\bar{\varphi} = \Omega(\bar{\phi}_1, \ldots, \bar{\phi}_N)^T$, is obtained from a linear Fokker-Planck equation (Supplemental equation A9). This equation (Risken 1984) depends on the Jacobian $\mathbf{A}$ defined in equation 28 and on the noise-generating diffusion matrix $\mathbf{B}\mathbf{B}^T$, which is given by

$$\mathbf{B}\mathbf{B}^T = (\mathbf{S}\sqrt{diag\,(\mathbf{f}(\boldsymbol{\varphi}))})\,(\mathbf{S}\sqrt{diag\,(\mathbf{f}(\boldsymbol{\varphi}))})^T = \mathbf{S}\,diag\,(\mathbf{f}(\boldsymbol{\varphi}))\,\mathbf{S}^T. \tag{31}$$

The matrix diag($\mathbf{f}(\boldsymbol{\varphi})$) is diagonal with the elements $f_j(\boldsymbol{\varphi})$; $j = 1, 2, \ldots, R$. It depends on the macroscopic rate laws $\mathbf{f}(\boldsymbol{\varphi})$, rather than on the mesoscopic transition rates $\tilde{\mathbf{f}}(\boldsymbol{\varphi})$, because the difference between the two disappears in the LNA (see Supplemental material). The stationary distribution of the Fokker-Planck equation is a multivariate Gaussian with zero mean vector and covariance matrix $\mathbf{C} = \langle \boldsymbol{\delta X} \boldsymbol{\delta X}^T \rangle$, which is determined by the (Lyapunov) matrix equation

$$\mathbf{A}\mathbf{C} + \mathbf{C}\mathbf{A}^T + \Omega\mathbf{B}\mathbf{B}^T = \mathbf{0}, \tag{32}$$

where $\mathbf{A}$ and $\mathbf{B}\mathbf{B}^T$ are evaluated at the stationary state $\boldsymbol{\varphi} = \bar{\boldsymbol{\varphi}}$. $\mathbf{C}$ contains the second-order moments of the stationary Gaussian distribution for the stochastic vector $\boldsymbol{\delta X}$, according to

$$\langle \delta X_i \delta X_k \rangle = C_{ij} \qquad i,k = \{1, 2, \ldots N\} \tag{33}$$

Knowledge of $\mathbf{A}$ and $\mathbf{C}$ also gives the time correlation matrix for $\boldsymbol{\delta X}$ according to

$$\langle \boldsymbol{\delta X}(t+\tau)(\boldsymbol{\delta X}(t))^T \rangle = \mathbf{C} \exp(\mathbf{A}|\tau|) \tag{34}$$

Equation 32 can be solved for $\mathbf{C}$ numerically with standard methods, but analytical expressions can also be obtained. The properties of $\mathbf{C}$ are best explained from the analytical solution that exists for the covariance matrix $\tilde{\mathbf{C}}$ for the new variables $\boldsymbol{\delta\tilde{X}} = \mathbf{Q}^{-1}\boldsymbol{\delta X}$, that is, fluctuations as normal modes (compare with equation 30). $\boldsymbol{\delta\tilde{X}}$ is associated with the linearly transformed stoichiometric matrix $\tilde{\mathbf{S}} = \mathbf{Q}^{-1}\mathbf{S}$, and it follows from equation 31 that $\mathbf{B}_{\tilde{\mathbf{B}}} = \mathbf{Q}\tilde{\mathbf{B}}$. The covariance matrix $\tilde{\mathbf{C}}$ for $\boldsymbol{\delta\tilde{X}}$ is defined from $\mathbf{C} = \mathbf{Q}\tilde{\mathbf{C}}\mathbf{Q}^*$ (Horn and Johnson 1985), where $\mathbf{Q}^*$ means $\mathbf{Q}$ complex conjugated and transposed, and $\mathbf{A} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^{-1}$. Insertion in equation 32 gives

$$\mathbf{A}\mathbf{C} + \mathbf{C}\mathbf{A}^T + \Omega\mathbf{B}\mathbf{B}^T = \mathbf{0} \Leftrightarrow$$

$$\boldsymbol{\Lambda}\tilde{\mathbf{C}} + \tilde{\mathbf{C}}\boldsymbol{\Lambda}^* + \Omega\tilde{\mathbf{B}}\tilde{\mathbf{B}}^* = \mathbf{0} \Leftrightarrow$$

$$\tilde{C}_{jk} = \Omega \frac{(\tilde{B}\tilde{B}^*)_{jk}}{\lambda_j + \bar{\lambda}_k} \tag{35}$$

The diagonal elements of $\tilde{\mathbf{C}}$, that is, the variances $\langle \delta\tilde{X}_i^2 \rangle$, are the summed squares of the transformed stoichiometric coefficients $\tilde{S}_{ij}$ (which can be complex valued) multiplied with their corresponding rate laws (flows), divided by the real part of the rate of relaxation for the normal mode $\delta\tilde{\phi}_i$ in equation 30:

$$\tilde{C}_{ii} = \frac{\Omega}{2|\mathrm{Re}\ \lambda_i|} \sum_j f_j(\boldsymbol{\varphi}) |\tilde{S}_{ij}|^2. \tag{36}$$

It is seen that internal noise is boosted by high rates f_j and large stoichiometry coefficients $|\tilde{S}_{ij}|$ and that fluctuations are attenuated by a large (negative) real part of the eigenvalue of the normal mode $\delta\tilde{\phi}_i$ in equation 30. When **A** is symmetric, **Q** can be made unitary and then $\tilde{\mathbf{S}} = \mathbf{Q}^T\mathbf{S}$, which simplifies analysis and interpretation (see Example 3 in the Supplemental material).

Single–Component Systems

When there is only one chemical species X_1, then equation 32 is scalar and the LNA of the stationary variance becomes

$$2\lambda_1 C_{11} + \Omega\sum_j S_{1j}^2 f_i(\bar{\phi}_1) = 0 \Rightarrow C_{11} = -\frac{\Omega\sum_j S_{1j}^2 f_j(\bar{\phi}_1)}{2\lambda_1} \qquad (37)$$

The Fano-Factor and the Correlation Coefficient

The Fano-factor (Fano 1947) is a measure of the relative size of the internal fluctuations. It is defined as the variance divided by the average value, that is, $F_{Xi} = \langle\delta X_i^2\rangle/\langle X_i\rangle$, which in LNA is approximated by $F_{Xi} \approx C_{ii}/(\Omega\bar{\phi}_i)$. $F_X = 1$ when X has Poisson statistics. As both the average value and variance of internal fluctuations are proportional to the volume of the system, the Fano-factor is independent of volume. Another volume-independent stochastic measure is the correlation coefficient $r_{ki} = \sigma_{ki}{}^2/\sqrt{\sigma_k{}^2\sigma_i{}^2}$, which in the LNA becomes $r_{ki} \approx C_{ki}/\sqrt{C_{kk}C_{ii}}$. It is zero for uncorrelated fluctuations and +1 or −1 if fluctuations are perfectly linearly correlated or anticorrelated, respectively.

Elimination of Fast Variables Can Make the LNA More Accurate Near Critical Points

Starting from the full system description in terms of **S** and $\mathbf{f}(\boldsymbol{\varphi})$, we first identify the correlation matrix $\tilde{\mathbf{C}}$ from the LNA in equation 35, using equations 28 and 31. We identify the variables $(\tilde{\phi}_1\ldots\tilde{\phi}_j)$, which correspond to the slow relaxation modes where $|\mathrm{Re}\lambda_{1\ldots j}| \ll |\mathrm{Re}\lambda_{j+1\ldots N}|$. The macroscopic dynamics of the system (equation 24) are then formulated in terms of transformed variables $\tilde{\boldsymbol{\varphi}} = \mathbf{Q}^{-1}\boldsymbol{\varphi}$ similar to equation 30:

$$\dot{\boldsymbol{\varphi}} = \mathbf{S}\mathbf{f}(\boldsymbol{\varphi}) \Leftrightarrow \dot{\tilde{\boldsymbol{\varphi}}} = \mathbf{Q}^{-1}\mathbf{S}\mathbf{f}(\mathbf{Q}\tilde{\boldsymbol{\varphi}}) = \mathbf{g}(\tilde{\boldsymbol{\varphi}}) \qquad (38)$$

Next, the fast variables $\tilde{\phi}_{j+1}\ldots\tilde{\phi}_N$ are locked to their macroscopic steady-state values, $\bar{\tilde{\phi}}_{j+1}\ldots\bar{\tilde{\phi}}_N$ conditional on $\tilde{\phi}_1\ldots\tilde{\phi}_j$, by setting the differential equations $g_{j+1}\ldots g_N$ equal to zero and solving the following system of equations:

$$\begin{pmatrix} g_{j+1}(\tilde{\phi}_1,..,\tilde{\phi}_N) \\ \vdots \\ g_N(\tilde{\phi}_1,..,\tilde{\phi}_N) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} \bar{\tilde{\phi}}_{j+1} \\ \vdots \\ \bar{\tilde{\phi}}_N \end{pmatrix} = \begin{pmatrix} h_{j+1}(\tilde{\phi}_1,\ldots,\tilde{\phi}_j) \\ \vdots \\ h_N(\tilde{\phi}_1,\ldots,\tilde{\phi}_j) \end{pmatrix} \qquad (39)$$

Subsequently, the LNA is applied only to the slow variables to get their stationary probability density function $P_1(\tilde{X}_1, \ldots, \tilde{X}_j)$. This is a normal distribution centered at the macroscopic steady state with a covariance matrix $\tilde{\mathbf{C}}_{1\ldots j,\,1\ldots j}$. For each value of $(\tilde{X}_1, \ldots, \tilde{X}_j)$, the fast variables are locked to their conditional stationary values $\bar{\tilde{\phi}}_{j+1}\ldots\bar{\tilde{\phi}}_N$, and the overall stationary distribution function is therefore:

$$\tilde{P}(\tilde{\mathbf{X}}) = \int_{\tilde{X}_1,\ldots,\tilde{X}_j} \delta(\tilde{X}_{j+1} - \Omega h_{j+1}(\tilde{x}_1,\ldots,\tilde{x}_j)) \cdots\delta(\tilde{X}_N - \Omega h_N(\tilde{x}_1,..,\tilde{x}_j))P_1(\tilde{X},..,\tilde{X}_j)d\tilde{X},..,d\tilde{X}_j, \qquad (40)$$

where $\delta(\tilde{X}_k - \Omega_{h_k}(\tilde{x}_1, \ldots, \tilde{x}_j))$ is the Dirac function and $\tilde{x}_i = \Omega^{-1}\tilde{X}_i$. $\tilde{P}(\tilde{\mathbf{X}})$ is not a normal distribution any more because the $\mathbf{h} = (h_{j+1}, \ldots, h_N)$ functions generally are nonlinear. Finally, we get the probability distribution function in the original variables as

$$P(\mathbf{X}) = \mathcal{N}\cdot\tilde{P}(\mathbf{Q}^{-1}\mathbf{X}), \qquad (41)$$

where $\mathcal{N}$ is a normalization constant. For a multidimensional $P(\mathbf{X})$, it is convenient to consider only the marginal distributions, defined as

$$P(X_j) = \int_{X_1,\ldots,X_{j-1},X_{j+1},\ldots,X_N} P(\mathbf{X})dX_1..dX_{j-1}dX_{j+1}dX_N. \qquad (42)$$

ACKNOWLEDGMENTS

This work was supported by the National Graduate School of Scientific Computing (NGSSC). We thank Martin Lovmar for comments on the manuscript and Lars Elden for pointing out the Horn and Johnson reference.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

REFERENCES

Barkai, N. and Leibler, S. 2000. Circadian clocks limited by noise. *Nature* **403:** 267–268.

Becskei, A. and Serrano, L. 2000. Engineering stability in gene networks by autoregulation. *Nature* **405:** 590–593.

Berg, O.G. 1978. A model for the statistical fluctuations of protein numbers in a microbial population. *J. Theor. Biol.* **71:** 587–603.

Berg, O.G., Paulsson, J., and Ehrenberg, M. 2000. Fluctuations and quality of control in biological cells: Zero-order ultrasensitivity reinvestigated. *Biophys. J.* **79:** 1228–1236.

Cornish-Bowden, A. 1995. *Fundamentals of enzyme kinetics*. Portland Press, London.

Dogterom, M. and Leibler, S. 1993. Physical aspects of the growth and regulation of microtubule structures. *Phys. Rev. Lett.* **70:** 1347–1350.

Donachie, W.D. and Robinson, A. 1987. Cell division: Parameter values and the process. In *Escherichia coli and Salmonella typhimurium: Cellular and molecular biology* (eds. F. Neidhardt et al.), pp. 1578–1593. ASM Press, Washington, DC.

Elf, J., Doncic, A., and Ehrenberg, M. 2003a. Mesoscopic reaction-diffusion in intracellular signaling. In *Fluctuations and noise in biological, biophysical and biomedical systems* (eds. S. Bezrukov et al.), pp. 114–124. SPIE, Bellingham, WA.

Elf, J., Paulsson, J., Berg, O.G., and Ehrenberg, M. 2003b. Near-critical phenomena in intracellular metabolite pools. *Biophys. J.* **84:** 154–170.

Elowitz, M.B., Levine, A.J., Siggia, E.D., and Swain, P.S. 2002. Stochastic gene expression in a single cell. *Science* **297:** 1183–1186.

Fano, U. 1947. Ionization yield of ratios. II. The fluctuations of the number of ions. *Phys. Rev.* **72:** 26–29.

Fersht, A.R. 1998. *Structure and mechanism in protein science*. Freeman, New York.

Gardiner, C. 1985. *Handbook of stochastic methods*. Springer Verlag, Berlin.

Gibson, M. and Bruck, J. 2000. Efficient exact stochastic simulation of chemical systems with many species and channels. *J. Phys. Chem. A* **104:** 1876–1889.

Gillespie, D. 1976. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.* **22:** 403–434.

———. 1977. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **81:** 2340–2361.

Glad, T. and Ljung, L. 2000. *Control theory: Multivariable and nonlinear methods*. Taylor & Francis, London.

Goldbeter, A. and Koshland Jr., D.E. 1981. An amplified sensitivity arising from covalent modification in biological systems. *Proc. Natl. Acad. Sci.* **78:** 6840–6844.

———. 1982. Sensitivity amplification in biochemical systems. *Q. Rev. Biophys.* **15:** 555–591.

Guptasarama, P. 1995. Does replication-induced transcription regulate the myriad of low copy number proteins of *Escherichia coli*? *Bioessays* **17:** 987–997.

Haken, H. 1982. Synergetics. Nonequilibrium phase transitions and self-organization in biological systems. In *Thermodynamics and kinetics in biological processes* (eds. I. Lamprecht and A.I. Zotin). Walter de Gruyter, Berlin.

Heinrich, R. and Schuster, S. 1996. *The regulation of cellular systems*. Chapman and Hall, New York.

Horn, R.A. and Johnson, C.R. 1985. *Matrix analysis*. Cambridge University Press, Cambridge, UK.

———. 1991. *Topics in matrix analysis*. Cambridge University Press, Cambridge, UK.

Horsthemke, W. and Lefever, R. 1984. *Noise-induced transitions. Theory and applications in physics, chemistry, and biology*. Springer Verlag, Berlin.
Keizer, J. 1987. *Statistical thermodynamics of nonequilibrium processes*. Springer-Verlag, Berlin.
Kepler, T.B. and Elston, T.C. 2001. Stochasticity in transcriptional regulation: Origins, consequences, and mathematical representations. *Biophys. J.* **81:** 3116–3136.
Mendes, P. 1997. Biochemistry by numbers: Simulation of biochemical pathways with Gepasi 3. *Trends Biochem. Sci.* **22:** 361–363.
Ozbudak, E.M., Thattai, M., Kurtser, I., Grossman, A.D., and van Oudenaarden, A. 2002. Regulation of noise in the expression of a single gene. *Nat. Genet.* **31:** 69–73.
Paulsson, J. and Ehrenberg, M. 2001. Noise in a minimal regulatory network: Plasmid copy number control. *Q. Rev. Biophys.* **34:** 1–59.
Qian, H., Saffarian, S., and Elson, E.L. 2002. Concentration fluctuations in a mesoscopic oscillating chemical reaction system. *Proc. Natl. Acad. Sci.* **99:** 10376–10381.
Risken, H. 1984. *The Fokker-Planck equation*. Spinger Verlag, Berlin.
Sauro, H.M. 1993. SCAMP: A general-purpose simulator and metabolic control analysis program. *CABIOS* **9:** 441–450.
Savageau, M.A. 1976. *Biochemical systems analysis: A study of function and design in molecular biology*. Addison-Wesley, Reading, MA.
Schrödinger, E. 1944. *What is life?* Cambridge University Press, Cambridge, UK.
Strang, G. 1988. *Linear algebra and its applications*. Harcourt Brace Jovanovich Collage Publishers, Fort Worth, TX.
Strogatz, S.H. 1994. *Nonlinear dynamics and chaos: With applications to physics, biology, chemistry, and engineering*. Addison-Wesley, Cambridge, MA.
Stundzia, A.B. and Lumsden, C.J. 1996. Stochastic simulation of coupled reaction-diffusion processes. *J. Comput. Phys.* **127:** 196–207.
Thattai, M. and van Oudenaarden, A. 2001. Intrinsic noise in gene regulatory networks. *Proc. Natl. Acad. Sci.* **98:** 8614–8619.
van Kampen, N.G. 1997. *Stochastic processes in physics and chemistry*, 2nd ed. Elsevier, Amsterdam.
Vilar, J.M., Kueh, H.Y., Barkai, N., and Leibler, S. 2002. Mechanisms of noise-resistance in genetic oscillators. *Proc. Natl. Acad. Sci.* **99:** 5988–5992.
Wax, N. 1954. *Selected papers on noise and stochastic processes*. Dover, New York.

Received January 20, 2003; accepted in revised form July 22, 2003.