Random Signal Fluctuations Can Reduce Random Fluctuations in Regulated Components of Chemical Regulatory Networks

Johan Paulsson* and Måns Ehrenberg†

Department of Cell and Molecular Biology, Uppsala University, BMC, Box 596, Uppsala, 75124 Sweden

(Received 9 February 2000)

Many intracellular components are present in low copy numbers per cell and subject to feedback control. We use chemical master equations to analyze a negative feedback system where species X and S regulate each other's synthesis with standard intracellular kinetics. For a given number of X-molecules, S-variation can be significant. We show that this signal noise does not necessarily increase X-variation as previously thought but, surprisingly, can be necessary to reduce it below a Poissonian limit. The principle resembles Stochastic Resonance in that signal noise improves signal detection.

PACS numbers: 87.16.Yc, 87.15.Ya, 87.10.+e, 87.16.Xa

Intracellular processes are regulated by signal molecules that often are present in a few to a few hundred copies and display significant internal noise [1]. Since a noisy signal only represents the underlying state of the cell in a probabilistic sense, this has generally been assumed to randomize control [2–4]. In this Letter we use chemical master equations [5,6] to demonstrate that signal noise instead can attenuate the concentration noise in a regulated component. Our minimal regulatory network consists of two components, X and S, that regulate each other's synthesis, and performance is defined by the capacity to reduce random X-variation. We chose this particular copy number control (CNC) system because its kinetic mechanisms constitute simple intracellular standards, but also because it is virtually identical to models of CNC of bacterial plasmids [7]. Plasmid CNC systems are comparatively lucid and have evolved primarily to attenuate copy number fluctuations [4].

The macroscopic equations describing the feedback system are

$$\begin{cases} [\dot{x}] = \frac{k[x]}{1+[s]/K} - [x], \\ [\dot{s}] = k_s[x] - k_d[s], \end{cases} \tag{1}$$

where $[x]$ and $[s]$ are continuous concentration variables. For plasmids, the autocatalysis of X-molecules comes from the constant frequency with which each plasmid molecule tries to replicate itself [8]. S-molecules inhibit X-synthesis trials through so-called hyperbolic inhibition [Eq. (1)] so that the probability that a trial is successful depends on the concentration of S-molecules at the time of the trial [8].

Normalizing the variables by their nonzero steady states $[x_r] = [x]/[\overline{x}]$ and $[s_r] = [s]/[\overline{s}]$ gives $[\dot{x}_r] = k[x_r][1 + [s_r](k - 1)]^{-1} - [x_r]$ and $[\dot{s}_r] = k_d([x_r] - [s_r])$. The parameters k_s and K in Eq. (1) thus determine two characteristic concentration scales, and k_d determines how rapidly $[s_r]$ adjusts to changes in $[x_r]$. When k_d is small, relaxation to steady state is oscillatory, and when $k_d \to \infty$, $[s_r]$ is strictly proportional to $[x_r]$ at all times. The sensitivity with which the rate of $[x_r]$ synthesis responds to changes in $[s_r]$ increases with k, but for $k \gg 1$, sensitivity approaches an asymptote, where the rate of synthesis of $[x_r]$ per unit of $[x_r]$ is inversely proportional to $[s_r]$, i.e., $[\dot{x}_r] \approx [x_r][s_r]^{-1} - [x_r]$. The basic control principle is that $[s_r]$ follows changes in $[x_r]$, thereby boosting the relative rate of $[x_r]$ synthesis at low concentrations and restraining it at high.

If X-synthesis were not inhibited by S-molecules, Eq. (1) would predict that any initial condition is perpetuated if $k = 1$. However, due to the random nature of chemical reactions, this would correspond to an accelerated unbiased random walk. X-variation would then be limited only by X-extinctions or by physical restrictions, such as a depletion of the cell's resources, when the number of X-molecules becomes too large. Both these effects are evolutionary unfavorable for the cell and negative feedback CNC has therefore evolved to attenuate the fluctuations [4]. Since the evolutionary rationale for CNC is to reduce fluctuations, it cannot be properly inspected using macroscopic equations. Here we instead use the chemical master equation [5,6].

Four types of events are included in our mesoscopic model of CNC. When there are m X-molecules and n S-molecules, S-molecules are synthesized with rate $k_s m$ and degraded with rate $k_d n$. X-molecules are synthesized with rate $g_{m,n} = km/(1 + n/K)$, where K now includes the reaction volume, and degraded with rate m. For plasmids, this straightforward mesoscopic version of the inhibition mechanism is a good approximation since the duration of an X-synthesis event is only a few seconds while the half-life of S-molecules is about a minute [8]. If the number of S-molecules instead changed significantly over the duration of an X-synthesis trial, so that the trial cannot be considered instantaneous, the behavior is more complicated.

The rates are transition probabilities per time unit in the master equation for the probability $p_{m,n}$ of being in a state with m X-molecules and n S-molecules. With the step operator [6] $E_n^j f(n) = f(n + j)$, where $j = \pm 1$, this can

be written as

$$\dot{p}_{m,n} = (E_m^{-1} - 1)g_{m,n}p_{m,n} + (E_m - 1)mp_{m,n} + k_s m(E_n^{-1} - 1)p_{m,n} + k_d(E_n - 1)np_{m,n} + p_{m,n}\sum_{n=0}^{\infty} p_{1,n} \quad (2)$$

for $\{m > 0, n \geq 0\}$. The last term comes from conditioning the distribution on $m > 0$. This is done because the state with zero X- and S-molecules is absorbing and all others are transient. When X is essential for the survival of the cell, the conditioning has its natural counterpart in the death of X-free cells.

When k_d is so high that the number of S-molecules rapidly adjusts to the current number of X-molecules, the number of S-molecules is a fast variable that can be removed from the equations. Equation (2) then simplifies to a master equation for the probability p_m of being in a state with m X-molecules

$$\dot{p}_m = (E_m^{-1} - 1)g_m p_m + (E_m - 1)mp_m + p_1 p_m . \quad (3)$$

We will analyze Eq. (3) both for noise-free and noisy signals, i.e., when conditional S-variation is insignificant and significant, respectively.

When the conditional S-variation for a given value of X is negligible, then

$$g_m = km/[1 + (mk_s/k_d)/K] . \quad (4)$$

The most efficient control is then obtained when k is so high that g_m is approximately constant. The number of X-molecules then only deviates from the Poisson distribution at very low averages, $\langle m \rangle$. For a fixed average, lower values of k inevitably broaden the distribution. In fact, when conditional S-variation is negligible, the X-distribution can never be narrower than (approximately) Poisson even if all parameters in Eq. (2) can be chosen freely. Lower X-variation requires more efficient control kinetics, for instance, $g_m \propto m^{1-i}$. High i means high sensitivity amplification, defined as the percentage change in the response, g_m, over the percentage change in the signal, m [9]. In this case it means that the X-synthesis rate is high below the average and negligible above so that fluctuations are insignificant. Throughout this Letter we use $\langle m \rangle = 10$. For $i \geq 1$, the X-variances are then $\sigma_m^2 \approx \langle m \rangle / i$.

Considering the impact of a noisy signal, i.e., where conditional S-variation cannot be ignored, an adiabatic elimination of the fast variable S instead gives

$$g_m = km \sum_{n=0}^{\infty} \overline{p}_{n|m}/(1 + n/K) . \quad (5)$$

The quasistationary conditional probabilities of n S-molecules given m X-molecules, $\overline{p}_{n|m}$, are Poissonian with conditional average $\langle n \rangle_m = mk_s/k_d$, since all synthesis and degradation events are independent. The number of S-molecules at any given time thus only represents the number of X-molecules in a probabilistic sense. For instance, when $\langle n \rangle_m = m$, a sample drawn from $\overline{p}_{n|10}$ is lower than a sample drawn from $\overline{p}_{n|9}$ with probability 0.365. Since regulation is nonlinear, this will affect X-variation. In the notation of Haken [10] and Gardiner [5], the fast variable S can be said to be a noisy slave of the slow variable X.

To efficiently eliminate X-fluctuations, a higher m should correspond to a higher n so that the X-synthesis rate decreases efficiently. However, increasing the conditional S-variance increases the probability that a high m instead corresponds to a low n. It is tempting to assume that conditional S-fluctuations reduce performance by randomizing the critical step in regulation. However, having overlapping conditional S-distributions means that the probabilities of avoiding inhibition may become *more* separated than was possible without fluctuations. The reason is that the nonlinear regulatory mechanism receives a disproportional contribution from the tail of the distribution, and the probability mass in the tail in turn responds sensitively to changes in the average. This is a general principle for sensitivity amplification (Paulsson *et al.* [11]) for which the name Stochastic Focusing (SF) has been suggested.

Calculating the stationary distribution of Eq. (3) using Eq. (5) shows (Fig. 1) that X-fluctuations can be efficiently eliminated without conventional sensitivity amplification. In fact, X-variation can be reduced *indefinitely* if the rate constants can be chosen freely. Efficiency requires a high k, but there is an upper limit in k for given $\langle m \rangle$ and $\overline{p}_{n|m}$ since $g_m \geq km\overline{p}_{0|m}$ [Eq. (5)].

We also studied the impact of k_d using the two-dimensional master equation (2) when $k = 100$, both for $\langle n \rangle_{10} = 8$ and for insignificant conditional S-variation [12]. Decreasing k_d impairs CNC much more when conditional S-variation is significant (Fig. 2). This is not an effect of the higher sensitivity amplification. Using conventional sensitivity amplification in Eq. (2), such as $g_{m,n} = km/(1 + n^2/K)$, CNC works better at all values of k_d when conditional S-variation is negligible (Fig. 2). The difference between conventional sensitivity

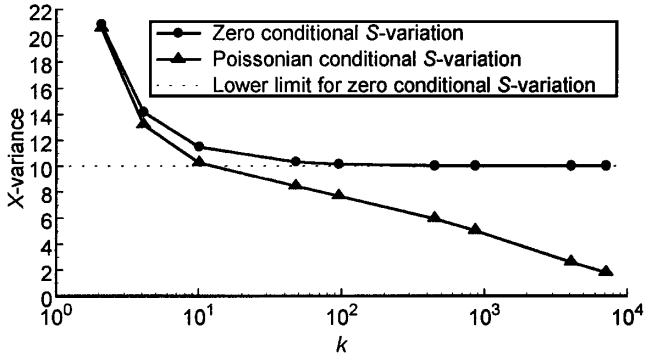

FIG. 1. X-variance as a function of k for hyperbolic inhibition when Eq. (5) is used in Eq. (3). The averages were $\langle m \rangle = 10$ and $\langle n \rangle_m = m$. The inhibition constant K changes along the curve to keep $\langle m \rangle = 10$.

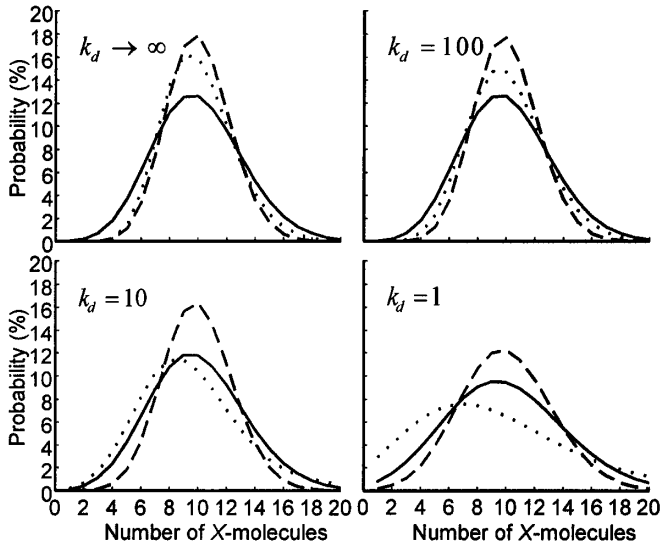

FIG. 2. Stationary X-distributions calculated by numerical integration of Eq. (2). The solid curve is hyperbolic inhibition without conditional S-variation [12], and the dotted curve is the same with Poissonian conditional S-fluctuations. The broken curve was obtained without conditional S-variation when $g_{m,n} = km/(1 + n^2/K)$. K changes to keep $\langle m \rangle = 10$ and $\langle n \rangle_{10} = 8$ in all curves.

amplification and fluctuation enhanced sensitivity (SF) is that low k_d and significant conditional S-fluctuations result in correlations between subsequent events. If n by chance is low at $t = 0$, it is likely to remain low for some time. Figure 2 shows that as time correlations become more and more significant, the beneficial effects of noise become overshadowed by stronger correlations between outcomes of subsequent X-synthesis trials.

The results presented in Figs. 1 and 2 are not restricted to Poisson fluctuations or exceptionally low conditional averages $\langle n \rangle_m$. We performed the same analysis when S-molecules were produced in instantaneous, geometrically distributed bursts. Geometric bursts are common in intracellular processes since they arise for Poisson processes in exponentially distributed time windows [3]. With k_s' as the rate constant for initiation of S-synthesis and $G_j = q^j(1 - q)$ as the probability for a burst of j molecules, adiabatic elimination of the fast variable S results in Eqs. (3) and (5) with $\overline{p}_{n|m}$ as the stationary distribution of

$$\dot{p}_{n|m} = k_s' m \left(\sum_{j=1}^{n} G_j p_{n-j|m} - q p_{n|m} \right) + k_d (E_n - 1) n p_{n|m} . \quad (6)$$

This gives the same type of macroscopic equation as before [Eq. (1)], and $\overline{p}_{n|m}$ is the negative binomial (NB)

$$\overline{p}_{n|m} = \frac{\beta^n}{(1 + \beta)^{\lambda m + n}} \frac{\Gamma(\lambda m + n)}{\Gamma(\lambda m) n!} , \quad (7)$$

where $\beta = q/(1 - q)$ is the average burst size and $\lambda = k_s'/k_d$. The corresponding conditional S-average and variance are $\langle n \rangle_m = m\lambda\beta$ and $\sigma^2_{n|m} = m\lambda\beta(\beta + 1)$, respectively. Fluctuations can thus be significant also at high averages. For plasmids, NB distributions can arise when S-synthesis depends on a transcription activator or repressor. The NB also arises as the stationary distribution for a large number of other simple birth and death processes of biochemical relevance, for instance autocatalysis [4].

To study the impact of increased variation using Eq. (7) in Eqs. (3) and (5), we numerically compared different combinations of β and λ for fixed averages $\langle m \rangle = 10$ and $\langle n \rangle_{10} = 50$. The rate constant k was fixed to 100, and K was changed to keep $\langle m \rangle = 10$. The results are shown in Fig. 3.

When $\beta \to 0$, the NB distribution converges to a Poissonian, and the conditional S-fluctuations have no significance because of the high average, resulting in an approximately Poisson distributed number of X-molecules (Fig. 3). When β increases, so does the conditional S-variation, but the X-variation instead decreases. In fact, the uncertainty in the number of X-molecules can only be decreased below the Poissonian limit by simultaneously increasing the conditional uncertainty in the number of S-molecules. This "kinetic uncertainty principle" is obviously not universal but rather depends on the kinetic mechanisms of the individual systems.

In conclusion, intrinsic noise in one component of a regulatory chemical network can be exploited to reduce intrinsic noise in another component, in direct contradiction to what has been previously assumed [2]. Signal noise arising from simplistic birth and death processes combined with simple and realistic biochemical mechanisms can, in fact, work as efficiently as a threshold mechanism combined with insignificant signal noise. This principle

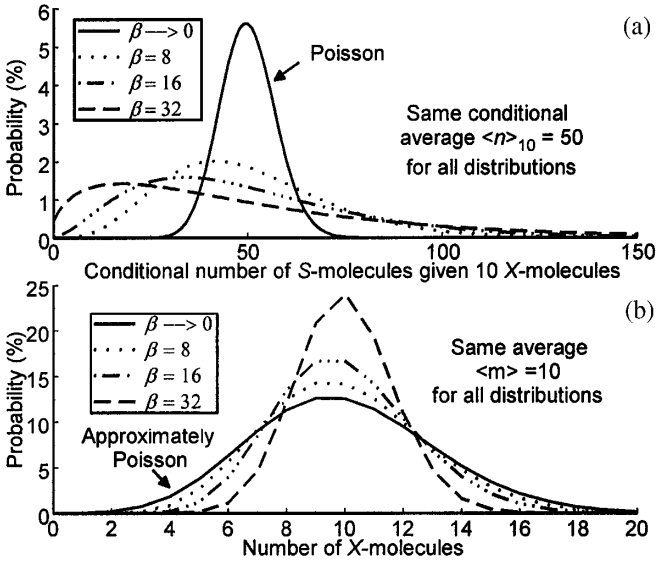

FIG. 3. Copy number distributions. (a) Conditional NB S-distribution [Eq. (7)] corresponding to ten X-molecules, i.e., $\overline{p}_{n|10}$. (b) X-distributions using Eqs. (5) and (7) in Eq. (3). See main text for parameters.

resembles Stochastic Resonance (SR) [13] in that signal noise can improve performance of a nonlinear system. However, signal, noise and performance are all different from SR standards. Instead, the results are more closely related to Renyi's classic paper [14] showing that average rates of bimolecular reactions are affected by copy number variances.

We acknowledge the support from the National Graduate School of Scientific Computing and the Swedish Natural Science Research Council, and thank O. G. Berg for valuable comments on the manuscript.

*Electronic address: Johan.Paulsson@icm.uu.se
†Electronic address: Mans.Ehrenberg@icm.uu.se

[1] P. Guptasarma, BioEssays **17**, 987 (1995).
[2] S. Benzer, Biochim. Biophys. Acta **11**, 383 (1953); J. L. Spudich and D. E. Koshland, Jr., Nature (London) **262**, 467 (1976); M. S. H. Ko, BioEssays **14**, 341 (1992); H. H. McAdams and A. Arkin, Proc. Natl. Acad. Sci. U.S.A. **94**, 814 (1997); A. Arkin, J. Ross, and H. H. McAdams, Genetics **149**, 1633 (1998); D. L. Cook, A. N. Gerber, and S. J. Tapscott, Proc. Natl. Acad. Sci. U.S.A. **26**, 15 641 (1998); H. H. McAdams and A. Arkin, Trends Genet. **15**, 65 (1999).
[3] O. G. Berg, J. Theor. Biol. **71**, 587 (1978).
[4] J. Paulsson and M. Ehrenberg, J. Mol. Biol. **279**, 73 (1998).
[5] C. W. Gardiner, *Handbook of Stochastic Methods* (Springer-Verlag, Berlin, 1985).
[6] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier, Amsterdam, 1992).
[7] Plasmids are ubiquitous self-replicating extrachromosomal gene clusters that genetically code for their own CNC.
[8] J. Paulsson, K. Nordström, and M. Ehrenberg, Plasmid **39**, 215 (1998); D. K. Summers, *The Biology of Plasmids* (Blackwell Science Ltd., Oxford, 1996).
[9] J. Higgins, Ann. N.Y. Acad. Sci. **108**, 305 (1963); M. A. Savageau, Nature (London) **229**, 542 (1971); D. E. Koshland, Jr., A. Goldbeter, and J. B. Stock, Science **217**, 220 (1982).
[10] H. Haken, *Synergetics* (Springer-Verlag, Berlin, 1978).
[11] J. Paulsson, O. G. Berg, and M. Ehrenberg, Proc. Natl. Acad. Sci. U.S.A. (to be published).
[12] The simulations were made with $\langle n \rangle_{10} = 100$ so that fluctuations were negligible. Virtually identical results were obtained both with $\langle n \rangle_{10} = 20$ and $g_m = km/(1 + \langle n \rangle_m / K)$.
[13] K. Wiesenfeld and F. Moss, Nature (London) **373**, 33 (1995); J. J. Collins, C. C. Chow, and T. T. Imhoff, Nature (London) **376**, 236 (1995); L. Gammaitoni, P. Hanggi, P. Jung, and F. Marchesoni, Rev. Mod. Phys. **70**, 223 (1998).
[14] A. Renyi, MTA Alk. Mat. Int. Közl. **2**, 83 (1953).