Noise effects in nonlinear biochemical signaling

Neda Bostani, David A. Kessler, Nadav M. Shnerb, Wouter-Jan Rappel, and Herbert Levine

Phys. Rev. E **85**, 011901 — Published 3 January 2012

DOI: 10.1103/PhysRevE.85.011901

Noise Effects in Nonlinear Biochemical Signaling

Neda Bostani,[1] David A. Kessler,[2,*] Nadav M. Shnerb,[2,†] Wouter-Jan Rappel,[3,‡] and Herbert Levine[3,§]

[1] *Key Laboratory of Particle Astrophysics, Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049, China*

[2] *Department of Physics, Bar-Ilan University, Ramat-Gan 52900 Israel*

[3] *Center for Theoretical Biological Physics, University of California San Diego, La Jolla, CA 92093-0319 USA*

It has been generally recognized that stochasticity can play an important role in the information processing accomplished by reaction networks in biological cells. Most treatments of that stochasticity employ Gaussian noise even though it is a priori obvious that this approximation can violate physical constraints, such as the positivity of chemical concentrations. Here, we show that even when such nonphysical fluctuations are rare, an exact solution of the Gaussian model shows that the model can yield unphysical results. This is done in the context of a simple incoherent-feedforward model which exhibits perfect adaptation in the deterministic limit. We show how one can use the natural separation of time scales in this model to yield an approximate model, that is analytically solvable, including its dynamical response to an environmental change. Alternatively, one can employ a cutoff procedure to regularize the Gaussian result.

PACS numbers: 02.50.Le, 05.65.+b, 87.23.Ge, 87.23.Kg

I. INTRODUCTION

The role of stochasticity in the functioning of cellular signal transduction networks is a question of great topical interest [1]. Unlike typical condensed-matter systems, biological cells must carry out chemical manipulations with small numbers of molecules, an inherently noisy situation. Noise comes in a variety of forms, including fluctuations in chemicals to be sensed [2], fluctuations in the binding-unbinding of receptor arrays [3], fluctuations during the processing of information [4], and fluctuations in the implementation of downstream actions [5].

In this context, almost all analytic studies of stochastic reaction dynamics utilize a small noise Gaussian approximation. This assumption emerges naturally, for example, in the van Kampen system-size expression [6] where the fluctuations in particle number are formally lower order and hence are treated as small and centered around the mean value set by the deterministic reaction equations. The initial purpose of this paper is to point out that this approach may give highly misleading results especially when some of the downstream reactions are nonlinear. We do this by studying a specific example, that of an incoherent feedforward module processing data from a small number of receptors [7]. Afterwards, we show how an alternate approach for the rapid activation versus slow inhibition limit can provide a complementary analytic approach.

The example we choose to study is a part of the gradient sensing module underlying the chemotactic response of Dictyostelium cells [8]. These cells appear to implement a control circuit incorporating a simple incoherent feedforward loop topology for adapting out the constant concentration background [9, 10]. This circuit is instantiated by using RAS-GEF as a positive signal and RAS-GAP as the complementary inhibitor [11]. These are both activated upon the binding of cAMP by the G-protein coupled cAMP receptor and in turn drive the signaling hub protein RAS into its active (GTP-bound), respectively inactive (GDP-bound) form. The circuit diagram is illustrated in Fig. 1. In the limit where we are far from saturation, the system can be described by the following equations [10]:

$$\begin{aligned} \frac{dA}{dt} &= \alpha S(t) - \gamma A \ , \\ \frac{dB}{dt} &= \beta S(t) - \delta B \ , \\ \frac{dE}{dt} &= A(1-E) - BE \ . \end{aligned} \tag{1}$$

*Electronic address: kessler@dave.ph.biu.ac.il
†Electronic address: shnerbn@mail.ph.biu.ac.il
‡Electronic address: rappel@physics.ucsd.edu
§Electronic address: hlevine@ucsd.edu

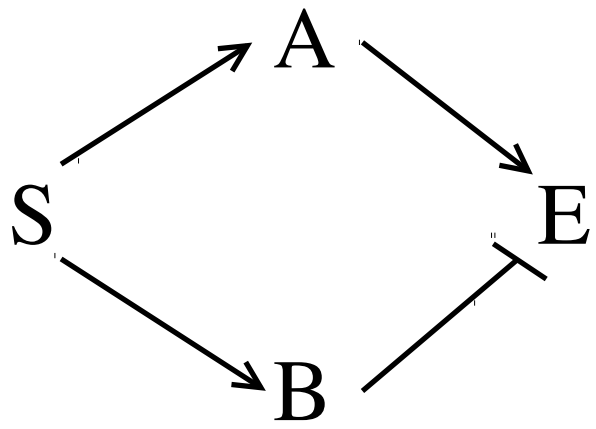

FIG. 1: Circuit diagram showing the activation of both A and B by the signal S. A in turn activates E, while B inhibits it.

Here E is the fraction of RAS molecules that have been activated and S is the external signal which drives both the activator A and inhibitor B. It is trivial to verify that if S is a constant signal, the steady-state value of E, $E_0 = \alpha\delta/(\alpha\delta + \beta\gamma)$ does not depend on its value. Hence this system in the deterministic limit is a perfectly adapting module, exhibiting only a transient response to changes in its input.

In order to study the effect of noise in the input signal $S(t)$ on this system, it is standard to assume that $S(t)$ is the sum of a deterministic signal $S_0(t)$ plus stationary Gaussian noise. For simplicity, we consider the case where the deterministic signal is a constant, S_0, and the noise, $\eta(t)$, has zero mean and correlator

$$\langle \eta(t)\eta(t')\rangle \equiv G(t-t') = \sigma^2 e^{-|t-t'|/\tau} \ . \tag{2}$$

We point out for future reference that the white noise limit is obtained, with these definitions, for $\sigma^2 \to \infty$, $\tau \to 0$, $\sigma^2\tau$ fixed, since as $\tau \to 0$ we have $\exp(-|t-t'|/\tau)/\tau \to \delta(t-t')$. The advantage of assuming Gaussian noise is that the system is then analytically tractable. The fields A and B can be expressed in term of the noise η as follows:

$$\begin{aligned} A(t) &= \tfrac{\alpha}{\gamma} S_0 + \alpha \int_{-\infty}^{t} dt' \ e^{-\gamma(t-t')}\eta(t') \ , \\ B(t) &= \tfrac{\beta}{\delta} S_0 + \beta \int_{-\infty}^{t} dt' \ e^{-\delta(t-t')}\eta(t') \ . \end{aligned} \tag{3}$$

Substituting this into the effector equation allows us to find

$$E(t) = \int_{-\infty}^{t} dt_1 \ g_1(t_1) e^{-\int_{t_1}^{t} dt_2 \ g_2(t_2)} \ , \tag{4}$$

with the definitions

$$\begin{aligned} g_1(t) &= \frac{\alpha}{\gamma} S_0 + \int_{-\infty}^{t} dt' \ \alpha e^{-\gamma(t-t')}\eta(t') \ , \\ g_2(t) &= \left(\frac{\alpha}{\gamma} + \frac{\beta}{\delta}\right) S_0 + \int_{-\infty}^{t} dt' \ \left(\alpha e^{-\gamma(t-t')} + \beta e^{-\delta(t-t')}\right)\eta(t') \ . \end{aligned} \tag{5}$$

From this expression, all moments of E can be calculated exactly. For example, let us focus on the expectation value of E. The standard expressions for Gaussian processes, for example

$$\left\langle e^{-\int \eta(t')h(t')dt'} \right\rangle = e^{\frac{1}{2}\int dt' dt'' h(t') G(t',t'') h(t'')} \ , \tag{6}$$

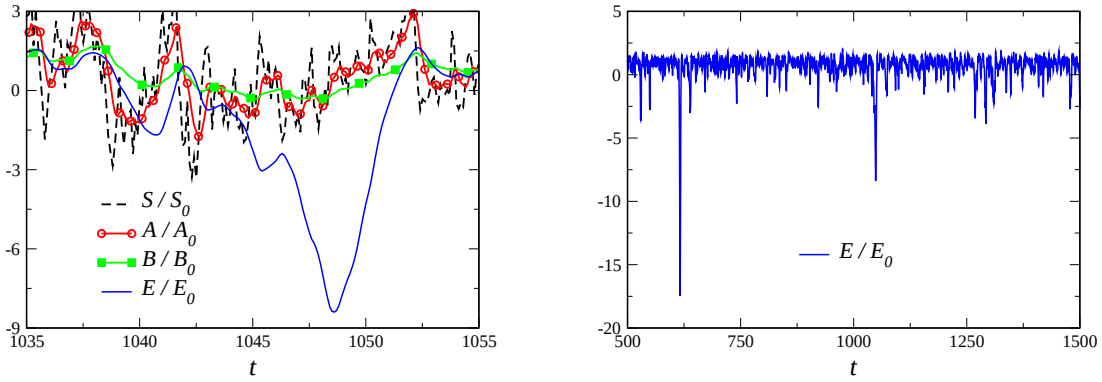

FIG. 2: (color online). Left) Excerpt of a simulation of the Gaussian model with $S_0 = 1/3$, $\sigma^2 = 2/9$, $\alpha = 1$, $\gamma = 2.5$, $\delta = 0.4$, $\beta = 1.7$, showing a large negative fluctuation of E/E_0 (solid blue trace). Note that β is chosen to be exactly the critical value at which Eq. (9) is violated and $\langle E \rangle$ diverges, Also shown are $S(t)/S_0$ (dashed black line), $A(t)/A_0$ (solid red line with open circles) and $B(t)/B_0$ (solid green line with filled squares). Right) Larger time series of $E(t)/E_0$ showing the intermittent nature of the large fluctuations.

for general functions $h(t)$, allow us after a tedious calculation to derive the following expression for $\langle E \rangle$:

$$\langle E \rangle = \int_0^\infty du\; e^{-S_0\left(\frac{\alpha}{\gamma}+\frac{\beta}{\delta}\right)u} e^{\sigma^2 \Phi(u)} \left(\frac{\alpha S_0}{\gamma} - \alpha\sigma^2 \Delta_E(u) \right) \; . \tag{7}$$

From this, one can immediately recover the aforementioned $\sigma^2 = 0$ deterministic result, $\langle E \rangle = \frac{\delta\alpha}{\delta\alpha+\beta\gamma}$. The exact expression for $\Delta_E(u)$, given in Appendix A, is not particularly informative; its only critical feature is that it decays to a constant at large u. The factor $\sigma^2\Phi(u)$ in the exponent is much more significant. Again, we leave the full form for Appendix A, and merely give the large u behavior:

$$\sigma^2 \Phi(u) \sim u\sigma^2\tau \left(\frac{\alpha}{\gamma} + \frac{\beta}{\delta} \right)^2 \; . \tag{8}$$

This term represent the diffusive growth of (the integral of) S^2 and has a well-defined form even in the white-noise limit for S where $\tau \to 0$ with $\sigma^2\tau$ fixed.

The starting point of our work is the observation that the integral defining $\langle E \rangle$ fails to converge unless

$$S_0 > \sigma^2\tau \left(\frac{\alpha}{\gamma} + \frac{\beta}{\delta} \right) \; . \tag{9}$$

It is easy to show that similar, but more stringent, bounds hold for all moments of E, which therefore are predicted to grow without bound (starting from any initial state) if the noise is too large. The problem arises from the fact that the S fluctuations are unbounded below for Gaussian noise and hence can drive A and/or B negative. This gives rise to transient periods during which $|E|$ *grows* exponentially. This behavior can be directly seen in a simulation of the Gaussian noise model, as presented in the left panel of Fig. 2. (In passing, these large excursions are rare events if σ is sufficiently small, and hence getting an accurate measure of $\langle E \rangle$ from the simulations can be difficult. We will return to this point in more detail later). As the noise gets large these growth periods lead to ever increasing values of $|E|$ and the anomalous contributions to $\langle E \rangle$ never saturate. It is important to note that even when $S_0/\sigma \ll 1$, so that the negative fluctuations of A and B are rare, nevertheless the Gaussian model can yield unphysical, infinite results, if $\alpha/\gamma + \beta/\delta$ is large enough. Thus, our finding depends essentially on the nonlinear coupling of the noisy signal to the E field, so that the noise acts multiplicatively on E. This troublesome behavior persists in the white noise limit, since the above condition, Eq. (9), involves the white noise amplitude, $\sigma^2\tau$.

II. BINOMIAL NOISE

Clearly, the Gaussian noise approach is in general unacceptable. We must treat the noise in a more realistic fashion if we are to have a well-defined model. If the source of the noise is the finite number of receptors [12, 13], we are led to consider a model wherein the incoming signal is a random variable reflecting the fraction of bound sensors. Assuming that the N receptors are independent and bind a ligand of fixed concentration, c_0, the signal can be exactly described via the master equation for the probability distribution, $P(s,t)$, for the number $s = 0, 1 \ldots, N$ of occupied receptors. where the signal $S = s/N$; i.e., the fraction of occupied receptors:

$$\begin{aligned} \frac{\partial P(s,t)}{\partial t} &= -\big[k_{\rm off}s + k_{\rm on}c_0(N-s)\big]P(s,t) + k_{\rm off}(s+1)P(s+1,t) \\ &\quad + k_{\rm on}c_0(N-s+1)P(s-1,t) \ . \end{aligned} \tag{10}$$

We call the model with this discrete noise the binomial noise model, as the equilibrium distribution of s is binomial,

$$P_{eq}(s) = \frac{N!}{s!(N-s)!}\frac{(k_{\rm on}c_0)^s(k_{\rm off})^{N-s}}{(k_{\rm on}c_0+k_{\rm off})^N} \ . \tag{11}$$

For large N, we recover a Gaussian noise process, except in the tails, as can be seen via the following argument. Our discrete stochastic process, Eq. 10, is well approximated for large N by the Ornstein-Uhlenbeck process, described by the Fokker-Planck equation

$$\tau\frac{\partial P_G(S,t)}{\partial t} = \frac{\partial}{\partial S}\left[(S-S_0)P_G\right] + \sigma_S^2\frac{\partial^2 P_G}{\partial S^2} \ , \tag{12}$$

with $S_0 = k_{\rm on}c_0/(k_{\rm on}c_0 + k_{\rm off})$, $\tau^{-1} = k_{\rm off} + k_{\rm on}c_0$ and $\sigma_S^2 = k_{\rm off}k_{\rm on}c_0/(N(k_{\rm on}c_0 + k_{\rm off})^2)$ so that in the limit the stochastic function $S(t)$ is Gaussian with variance σ_S. Here, of course, the equilibrium distribution is Gaussian,

$$P_{G,eq}(S) \simeq \exp\left[-\left(S-S_0\right)^2/2\sigma_S^2\right] \ , \tag{13}$$

and the steady-state autocorrelation function is

$$< S(t)S(t') > - \ S_0^2 = \sigma_S^2 e^{-|t-t'|/\tau} \ . \tag{14}$$

Nevertheless, for all finite N, the signal S in the binomial process is always non-negative, and no anomalous behavior can occur, with $E(t)$ strictly bounded from below by 0. The striking difference in the two models is apparent by comparing the Gaussian simulation presented in Fig. 2 to the simulation of the corresponding binomial model in Fig. 3. The well-behaved nature of the binomial model for all N is consistent with the fact that our condition for the convergence of the first moment in the Gaussian model, Eq. (9), is always satisfied in the large N limit, given that the noise amplitude σ_S^2 is small, of order $O(1/N)$. The large-N limit, however, is only an asymptotic approximation, since for any finite value of the noise amplitude σ_S^2 of the Gaussian model, sufficiently high moments of E do indeed diverge.

An additional perspective on the difference between the binomial and its parallel Gaussian model is afforded by examining the equilibrium $\langle E\rangle$ as a function of the parameter β. This is presented in Fig. 4. In the Gaussian noise model, the divergence condition condition is obviously satisfied for $\beta > \beta_c$, where

$$\beta_c \equiv \delta\left[\frac{S_0}{\sigma^2\tau} - \frac{\alpha}{\gamma}\right] \ , \tag{15}$$

since the right-hand size of Eq. (9) grows linearly with β. The incipient divergence of $\langle E\rangle$ for the Gaussian model at $\beta_c = 1.7$ is apparent from the figure. The binomial model, on the other hand, shows no special behavior at the critical β_c at which $\langle E\rangle$ diverges in the Gaussian model. The binomial model $\langle E\rangle$ exhibits a broad minimum at $\beta \approx 3.9$ and then rises toward unity as β increases. Even before the divergence, the Gaussian model deviates significantly from its binomial counterpoint, since binomial noise is far from Gaussian when N is small. The third curve in this figure results from a cutoff version of the Gaussian model, to be discussed later. For smaller noise, (equivalently, larger N in the binomial model), however, as depicted in Fig. 5, the difference between the two models is small, essentially right up to the divergence, as binomial noise is very well approximated by Gaussian noise, except in the tails, which are irrelevant (at least for the first moment, $\langle E\rangle$) as long as one is a finite distance below β_c. The onset of the deviation for small noise in the Gaussian model is extremely close to β_c. For example, the Gaussian model $\langle E\rangle$ crosses zero at a distance of the order of 10^{-18} below β_c for the parameters of Fig. 5.

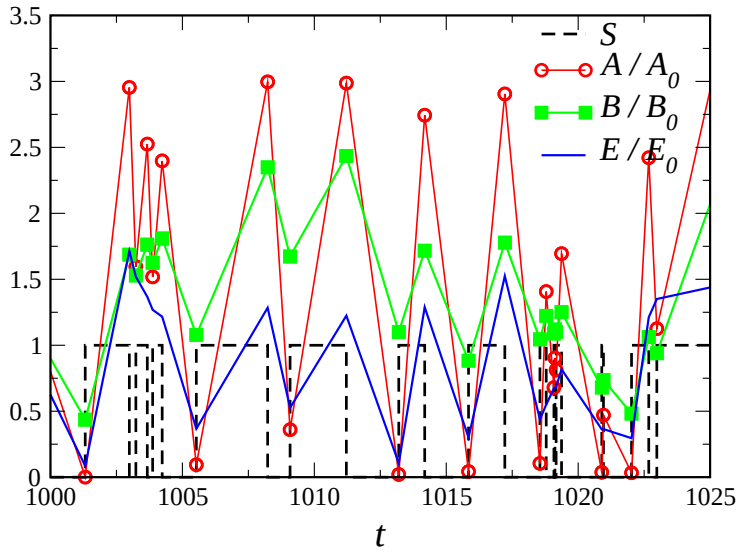

FIG. 3: (color online). Excerpt of a simulation of the $N = 1$ binomial model with parameters parallel to those of the Gaussian simulation in Fig. 2: $k_{\rm on}c_0 = 1$, $k_{\rm off} = 2$, $\alpha = 1$, $\gamma = 2.5$, $\delta = 0.4$, $\beta = 1.7$.

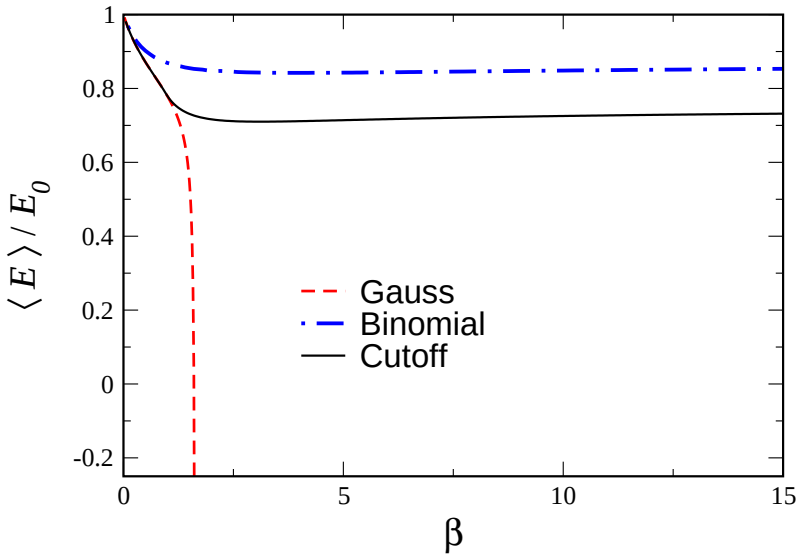

FIG. 4: (color online). Variation of $\langle E\rangle$ with β for the Gaussian model, Eq. 7, (red short dashed line) and the corresponding $N = 1$ binomial model (blue dashed-dotted line), derived from averaging 10^4 simulations. The parameters of the Gaussian model (except β) are as in Fig. 2; for the binomial model as in Fig. 3. These results are compared to that of the cutoff model, defined later in the text (solid black line).

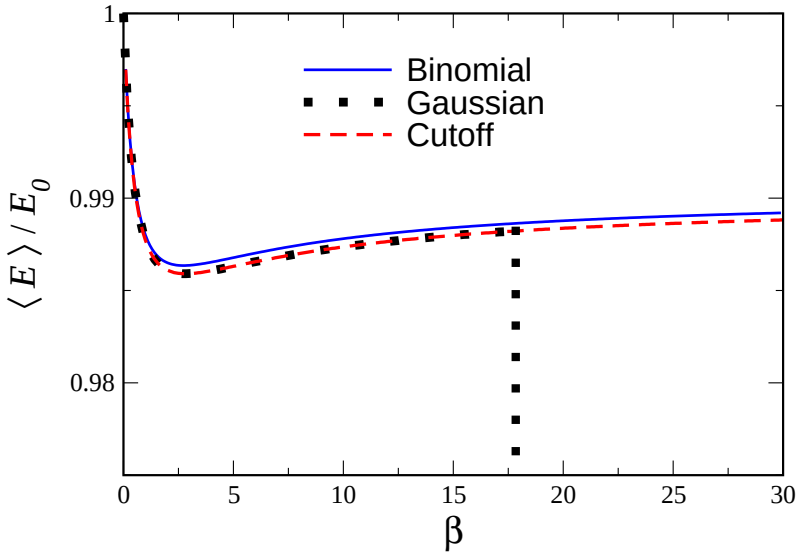

FIG. 5: (color online).Variation of $\langle E \rangle$ with β for the Gaussian model, Eq. 7, (black dotted line) and the corresponding binomial model (blue solid line), derived from averaging 10^4 simulations. Here $N = 10$ in the binomial model, and similarly the Gaussian model has $\sigma^2 = 4/90$, so that the Gaussian $\langle E \rangle$ diverges at $\beta = 17.84$. The other parameters are as in Fig. 4. These results are compared to that of the cutoff model, defined later in the text (red dashed line).

III. SUDDEN/ADIABATIC (S/A) APPROXIMATION

Unfortunately, the binomial model does not admit an analytic solution for finite N (which is of course the interesting case, since otherwise the effect of the noise is infinitesimal). We can however make use of the natural ordering of time scales in the problem to construct a solvable limit. For the circuit to show a significant transient response, it is necessary for the time scale of the A dynamics to be much faster than the B dynamics, otherwise the system adapts too rapidly to the changing signal and the transient response is aborted. Furthermore, the time scale of the E dynamics, $1/(A_0 + B_0)$, should be intermediate to those of the A and B fields. If the E dynamics is too fast, then the system is limited in any case to respond no quicker than A, and if the E dynamics is too slow, A and B have reequilibrated by the time E starts to respond, and again there is no transient response. Also, the time scale of the noise dynamics should be intermediate. Too fast noise would just get averaged away, and too slow noise would be adapted away. Thus, we are led to consider the limit where the A dynamics is much faster than all the other processes, and the B dynamics much slower. This limit is analytically treatable, as we now proceed to show.

Formally, we define our approximate theory, which we denote the Sudden/Adiabatic (S/A) theory, by taking both $\beta \to 0$ and $\delta \to 0$, with a fixed ratio $B_p = (\beta/\delta)$. Since the time scale of the B dynamics is so long, the noise in the signal is completely averaged over and we can just set $B = B_0 = B_p(k_{\rm on}c_0/(k_{\rm on}c_0 + k_{\rm off}))$, its average value. In addition, we take the limit of large α and γ (with fixed ratio A_p) which guarantees that the activator dynamics is fast enough to precisely follow the noise, i.e. $A(t) = S(t)A_p$.

We first examine the case $N = 1$. To proceed, we decompose the equation for the probability distribution of E into $P_{0,1}(E)$, the joint probability of E and the input signal S being 0 or 1, respectively, so that $P(E,t) = P_0(E,t)+P_1(E,t)$. We immediately derive that, in steady state,

$$\begin{aligned} \frac{\partial P_1}{\partial t} &= k_{\rm on}c_0P_0 - k_{\rm off}P_1 + \frac{\partial}{\partial E}\left([(B_0 + A_p)E - A_p]\,P_1\right) = 0\ , \\ \frac{\partial P_0}{\partial t} &= -k_{\rm on}c_0P_0 + k_{\rm off}P_1 + \frac{\partial}{\partial E}\left(B_0EP_0\right) = 0\ . \end{aligned} \tag{16}$$

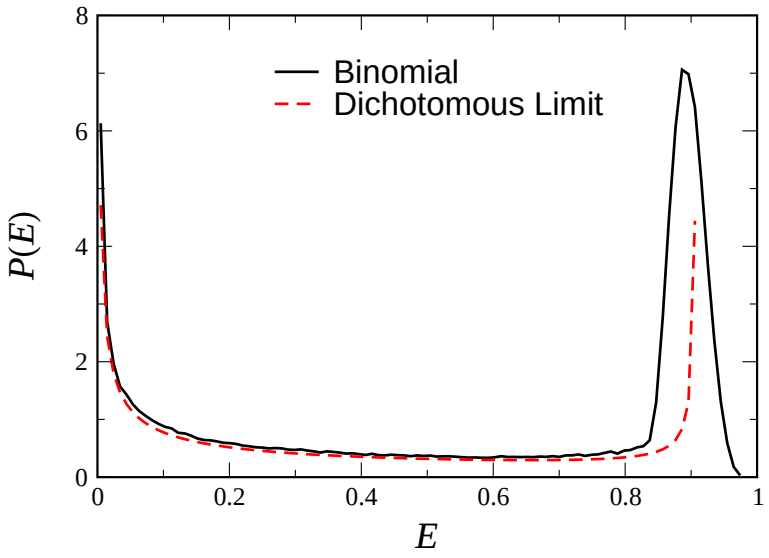

FIG. 6: The probability distribution function, $P(E) = P_+(E) + P_-(E)$ as a function of E for the $N = 1$ binomial model with $\alpha = 100$, $\gamma = 10$, $\beta = 0.2$, $\delta = 0.1$, $k_{\rm on}c_0 = k_{\rm off} = 0.4$, compared with the $N = 1$ S/A model, with parameters $A_p = 10$, $B_0 = 1$.

Adding the two equations and integrating gives

$$\frac{E_p - E}{B_0}\, P_1(E) = \frac{E}{A_p + B_0}\, P_0(E)\ , \tag{17}$$

where $E_p \equiv A_p/(A_p + B_0)$. Substituting this back into Eqs. 16 and defining $r_+ = k_{\rm off}/(A_p + B_0)$, $r_- = k_{\rm on}c_0/B_0$, we can find the normalized probabilities defined on the interval $0 < E < E_p$,

$$\begin{aligned} P_1 &= \frac{\Gamma(r_+ + r_- + 1)}{\Gamma(r_+)\Gamma(r_-)} E_p^{-(r_+ + r_-)} \frac{B_0}{k_{\rm on}c_0 + k_{\rm off}} E^{r_-}\,(E_p - E)^{r_+ - 1}\ , \\ P_0 &= \frac{\Gamma(r_+ + r_- + 1)}{\Gamma(r_+)\Gamma(r_-)} E_p^{-(r_+ + r_-)} \frac{A_p + B_0}{k_{\rm on}c_0 + k_{\rm off}} E^{r_- - 1}\,(E_p - E)^{r_+}\ . \end{aligned} \tag{18}$$

The total probability $P(E)$ has the interesting behavior of switching from being peaked at the interval center to the interval endpoints, as the parameters are varied [14]. For $r_+ < 1$, the probability density diverges at E_p, and for $r_- < 1$, the density diverges at 0. The above expression immediately gives the prediction

$$\langle E\rangle = E_p \frac{r_- + k_{\rm on}c_0/(k_{\rm on}c_0 + k_{\rm off})}{r_+ + r_- + 1}\ . \tag{19}$$

To compare this to the full binomial model, we conducted a simulation with $\alpha = 100$, $\gamma = 10$ (giving $A_p = 10$), $\beta = 0.2$, $\delta = 0.1$, and $k_{\rm on}c_0 = k_{\rm off} = 0.4$, (giving $B_0 = 1$). Here, the above theory predicts $\langle E\rangle/E_0 = .684$, where the deterministic $E_0 = k_{\rm on}c_0A_p/(k_{\rm on}c_0A_p + (k_{\rm on}c_0 + k_{\rm off})B_0)$. The simulation gave $\langle E\rangle/E_0 = .691$ and indeed the histogram is peaked at the endpoints, as predicted by the analysis. This is seen in Fig. 6.

In particular, our limiting theory predicts that as a function of B_0, i.e. β, $\langle E\rangle/E_0$ starts at a value of unity at $B_0 = 0$, which is reasonable since the system is saturated and so $\langle E\rangle$ is unity independent of the noise. For small B_0, $\langle E\rangle/E_0$ falls with an an initial slope of $k_{\rm off}/(k_{\rm on}c_0(k_{\rm on}c_0 + k_{\rm off}))$. As a function of B_0, $\langle E\rangle/E_0$ reaches a minimum at $B_0 = \sqrt{A_pk_{\rm on}c_0}$ and then turns back up, approaching unity at large B_0. This qualitative behavior is in accord with what we saw in Fig. 4, even though there the parameters are far from fulfilling the separation of scales assumed in the analysis. For $k_{\rm on}c_0 = k_{\rm off}$, the value of $\langle E\rangle/E_0$ at the minimum is

$$(\langle E\rangle/E_0)_{min} = \frac{(r+2)^2}{2(r^2 + 2r + 2)}; \qquad r \equiv \sqrt{A_p/k_{\rm on}c_0}\ . \tag{20}$$

This decreases from unity for small r to a value of $1/2$ at large r. Thus the larger $A_p/k_{\rm on}c_0$, the larger the noise-induced relative suppression of $\langle E\rangle$, since the effective noise amplitude increases as $k_{\rm on}c_0$ decreases. When $k_{\rm off} \ll k_{\rm on}c_0$, the value at the minimum approaches unity, so there is no suppression. On the other hand, when $k_{\rm off} \gg k_{\rm on}c_0$, the value at the minimum approaches $k_{\rm off}/(k_{\rm off}+A_p)$, which indicates the maximal suppression occurs at $A_p \gg k_{\rm off} \gg k_{\rm on}c_0$.

A. Moment Equations

One cannot extend the above approach to compute the an exact closed-form expression for the steady-state distributions for $N > 1$. However, one can make progress by recognizing that the moment equations take a particularly simple form. Consider the steady-state master equation for the case of general N. We have, in an obvious notation,

$$\begin{aligned}
0 &= k_{\rm on}c_0 P_{N-1} - 2k_{\rm off}P_N + \frac{d}{dE}\Big(\big[(B_0+A_p)E - A_p\big]P_2\Big)\ , \\
0 &= -jk_{\rm on}c_0P_j - (N-j)k_{\rm off}P_j + (N-j+1)k_{\rm on}c_0P_{j-1} + (j+1)k_{\rm off}P_{j+1} \\
&\qquad + \frac{d}{dE}\left(\left[\left(B_0+\frac{jA_p}{N}\right)E - \frac{jA_p}{N}\right]P_j\right)\ , \qquad\qquad j = 2\ldots N-1 \\
0 &= -Nk_{\rm on}c_0P_0 + k_{\rm off}P_1 + + \frac{d}{dE}\Big(B_0EP_0\Big)\ .
\end{aligned} \tag{21}$$

Because of the form of these equations, we can get a closed linear system for the moments $z_n \equiv \int dE P_n(E)E$:

$$0 = -(N-j)k_{\rm on}c_0z_j - jk_{\rm off}z_j + (N-j+1)k_{\rm on}c_0z_{j-1} + (j+1)k_{\rm off}z_{j+1} - \left(B_0+\frac{jA_p}{N}\right)z_j - \frac{jA_p}{N}\,\Pi_j\ . \tag{22}$$

where $\Pi_n \equiv \int dE P_n(E)$ are just the binomial occupation probabilities and $z_{-1} = z_{N+1} \equiv 0$. This $(N+1)\times(N+1)$ linear system can immediately be solved for the z's for any given N. In Fig. 7, we plot the results obtained by solving these equations to determine $\langle E\rangle = \sum_{n=0}^{N} z_n$.

For large N, the distribution Π_j becomes highly peaked around its mean, $j = Nk_{\rm on}c_0/(k_{\rm on}c_0+k_{\rm off}) \equiv N\bar{x}$, and so does z_j. Thus, to leading order, we can approximate jA_p/N by its mean, $A_0 \equiv A_p\bar{x}$, allowing us to solve the resulting system,

$$z_j \approx z_j^{(0)} = A_0/(A_0+B_0)\Pi_j = E_0\Pi_j\ . \tag{23}$$

Thus, $\langle E\rangle = E_0$, and moreover the mean value of E, conditioned on the value of the input j is in fact independent of j, so that adaptation become perfect for large N. To investigate the finite N effects, we expand around z_j, $z_j = z_j^{(0)} + \Delta_j$ using the fact that for all the important modes, $j/N - \bar{x}$ is small, of order $O(N^{-1/2})$. To solve the resultant system, we can approximate it by an ODE for $\Delta(y)$ thought of as a function of the continuous variable $y = \sqrt{N}(n/N - \bar{x})$. This calculation is presented in Appendix B, and leads to the result

$$\langle E\rangle \approx E_0 - \frac{A_p^2B_0\sigma^2}{N(A_0+B_0)^2(A_0+B_0+\omega)}\ , \tag{24}$$

where $\omega = 1/\tau = k_{\rm on}c_0 + k_{\rm off}$. Thus,

$$1 - \langle E\rangle/E_0 \approx \frac{A_0B_0(1-\bar{x})}{N\bar{x}(A_0+B_0)(A_0+B_0+\omega)}\ . \tag{25}$$

Thus again $\langle E\rangle/E_0$ starts at 1 for $B_0 = 0$ but here falls initially with slope $(1-\bar{x})/(N\bar{x}(A_0+\omega))$. Note that this is quite different from the $B = 0$ slope for $N = 1$, which was independent of A_0. Again, $\langle E\rangle/E_0$ has a minimum as a function of B_0, here at $B_0 = A_0\sqrt{1+\omega/A_0}$, with a value at the minimum of $(1-\bar{x})/\left(N\bar{x}\left(1+\sqrt{1+\omega/A_0}\right)^2\right)$. Thus again the suppression is largest for $k_{\rm off} \ll k_{\rm on}c_0$. Note that for small $k_{\rm on}c_0$, the suppression appears to grow without bound, although it is in fact bounded above by unity, indicating that the smaller $\bar{x}$, the larger N has to be in order for the large N results to be valid. Nevertheless, the suppression is still strong in this limit.

As mentioned above, our system of reactions, Eq. (1) was recently found to describe the RAS activation response of Dictyostelium cells to global cAMP stimulation. This is the first verified case of the incoherent feedforward structure in an actually biological process. The kinetic parameters derived for this model [10] are roughly in the range where

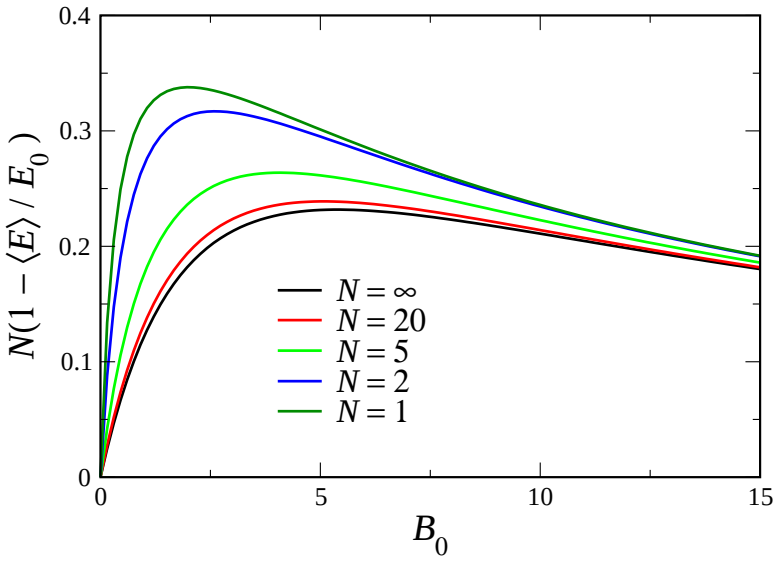

FIG. 7: The scaled suppression of the equilibrium $\langle E\rangle$, $N(1-\langle E\rangle)$ as a function of B_0 in the reduced model, for various $N = 1$, 2, 5, and 20 along with the asymptotic large-N result (in ascending order, with $N = 1$ at the bottom and the large N result, denoted $N = \infty$ at the top). The parameters are $A_0 = 10$, $k_{\rm on}c_0 = k_{\rm off} = 0.4$.

our approximations apply, namely the inhibition timescale is larger than the noise correlation time which in turn is larger than that of the activation. In that system, the downstream effector is actually fast; this is not the most general possibility, but our methodology is able to handle this limiting case without any issues. Perhaps more importantly, the RAS system couples to a huge number of receptors (then of thousands) and hence stochastic effects cannot play any role for the global response of interest here [16] . Direct testing of our results will have to await the discovery of a gene expression problem which utilizes this network structure and where the number of "receptors" (here promoter regions on the DNA) involved in signal sensing can be as small as one.

IV. GAUSSIAN MODEL – LARGE N LIMIT

We have seen that the binomial model becomes Gaussian in the large N limit. It is interesting to check that our results for the S/A limit of the binomial model agree with the corresponding limit of the Gaussian model for small noise. For small noise, the Gaussian model gives a finite answer, since our criterion is automatically satisfied. Indeed, upon expanding Eq. 7 to linear order in σ_S^2 and performing the integral we get

$$\langle E\rangle \approx E_0\left[1-\sigma_S^2\frac{\delta\gamma^2\tau\beta(\gamma-\delta)(\gamma\delta+S_0\beta+S_0\alpha)(\alpha S_0\delta\tau+S_0\beta\gamma\tau+\delta^2\gamma\tau+\delta\gamma^2\tau+\gamma\delta)}{(\alpha S_0\delta+S_0\beta\gamma+\gamma\delta^2)(\delta\tau+1)(\gamma+\delta)(\alpha S_0\delta+S_0\beta\gamma+\gamma^2\delta)(\alpha S_0\delta\tau+S_0\beta\gamma\tau+\gamma\delta)(\gamma\tau+1)S_0}\right] \quad (26)$$

and then taking α and γ to ∞ with $\alpha/\gamma = A_0/\bar{x}$ and taking β and δ to zero with $\beta/\delta = B_0/\bar{x}$, and setting $S_0 = \bar{x}$, $\sigma_S^2 = \sigma^2/N$, we indeed reproduce our large N result, Eq. 25. In addition, the general result confirms that the suppression is maximized in our distinguished limit $\gamma \gg \delta$.

Thus, the Gaussian model is perfectly acceptable in the small noise limit. One can ask if there is way to extend it beyond this limit. Clearly just increasing the noise amplitude leads to problems, as we have seen. We have seen in Fig. 4 that the problem is not restricted just to noise levels bigger than the critical value. Rather, for this case of large noise, the Gaussian answer is not accurate even when we are not close to the critical value. The problem is of course the already demonstrated large negative excursions. For small or intermediate noise levels, the problematic negative excursions of A and B are actually quite rare. To understand this better, consider the distribution of $E(t)$, for some given t. For short times this is well-behaved, but when one exceeds the time-scale of the E dynamics, the

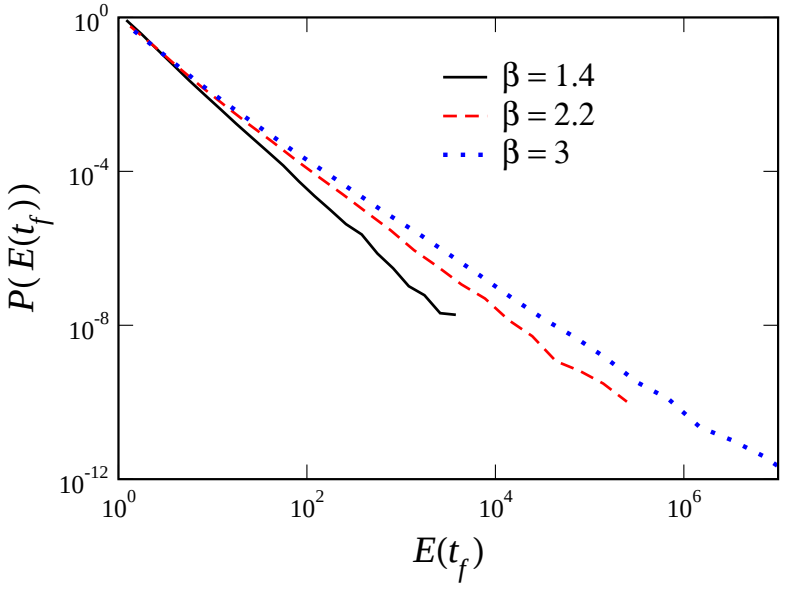

FIG. 8: (color online) Gaussian Model: Probability Distribution Function for $1 - E(t_f)$, conditioned on $E(t_f) < 0$, plotted in log-log scale, for $\alpha = 1.4$ (solid black line), 2.2 (dashed red line) and 3.0 (dotted blue line). The tails approach straight lines for large $|E(t_f)|$, indicating asymptotic power-law distributions with approximate exponents of 2.2, 1.8 and 1.6, respectively. $N = 0.5, t_f = 20$ All other parameters are as in Fig. 1. For each data set, 10^6 runs were performed, yielding 146,000, 203,000 and 234,000 data points satisfying $E(t_f) < 0$, respectively and the data was binned logarithmically using 25 bins.

distribution develops a power-law tail for large negative $E(t)$; this can be seen in Fig. 8. If $\beta < \beta_c$, the exponent of this distribution is greater than 2 in magnitude, and so the first moment is finite. Of course, there is a range for which the first moment is finite but the second and higher are already divergent. Since the exact formula shows that there is no divergence of any moment at finite t, there must be a cutoff in the power-law tail at some extremely large t dependent value, which however is very difficult to see from the numerics [15]. For $\beta > \beta_c$, on the other hand, the power decreases and the first moment diverges as well (subject to the same extremely large cutoff). As the noise level decreases, the above picture still holds. The probability of $E(t) < 0$, however, decreases exponentially with N. Thus, while for $\beta > \beta_c$, the first moment diverges, it becomes exponentially more difficult to see this in a simulation at intermediate noise levels. For small noise, it is well-nigh impossible. Thus, simulating the Gaussian model will, for small and intermediate noise give perfectly physical answers, which is, however, the incorrect answer for the true ensemble average, which is dominated by extremely rare huge events. Of course, if one wishes to rely on simulations, one can simulate directly the binomial noise model.

This line of reasoning leads to an alternate approach for extending the Gaussian model, loosely motivated by the successful use of simple cutoffs in regularizing reaction-diffusion equations which arise from the large N limit of Markov processes and which overemphasize the growth at small concentrations by missing the essential role of particle number discreteness [17–19]. (In fact, it has been recently proven that adding a cutoff to the Fisher equation [20–22] exactly reproduces the anomalous front velocity correction for asymptotic large particle number N [23]). Here, we cut off the integral in Eq. 7 at the point where either the integrand becomes (unphysically) negative or (unphysically) increasing as u increases. Fig. 4 presents results for $N = 1$ showing that this method not only prevents blow-up (by construction) but also does a decent job in capturing the true answer. It is clear that this method becomes more and more accurate as N increases, since for large N the unphysical neglected integrand is exponentially small in N (even though diverging as $u \to \infty$); this can be seen in Fig. 5. This is reminiscent of the situation we encountered with the Gaussian simulations, where the part of the distribution that has diverging mean has exponentially small probability.

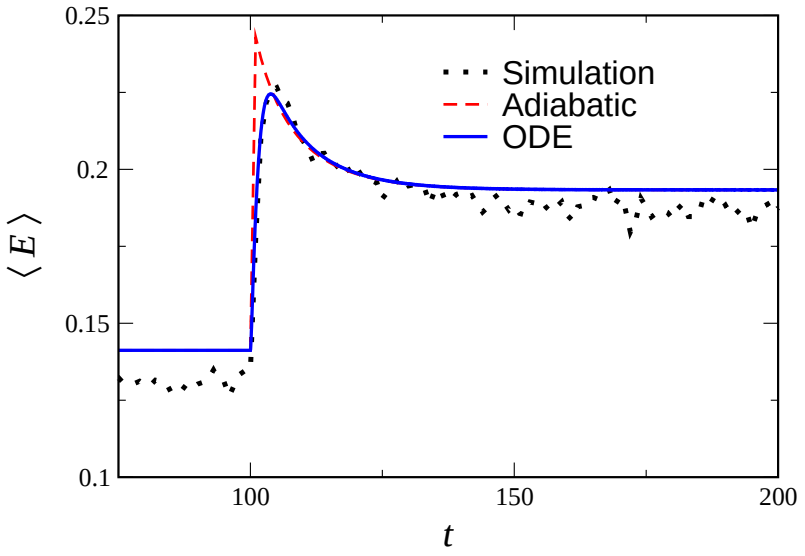

FIG. 9: (color online) The ensemble-averaged $\langle E(t)\rangle$ for the binomial model when c_0 is suddenly increased by a factor of 2 at $t_p = 100$. The average over 10^4 runs is denoted by the curve labeled Simulation (black dotted line). The adiabatic approximation where the equilibrium value of E is shown with B_0 taken to be its ensemble average $B0(t) = B_0' - (B_0' - B_0(t_p))e^{-\delta(t-t_p)}$, and the other parameters post-pulse are taken to be their new values is shown by the trace labeled Adiabatic (red dashed line). The numerical solution of the moment equations with $B_0(t)$ taken as above and $\Pi_1(t) = \bar{x}' - (\bar{x}' - \bar{x})e^{-\delta(t-t_p)}$ is marked by ODE (blue solid line). The parameters are $N = 1$, $\alpha = 100$, $\beta = 1$, $\gamma = 10$, $\delta = 0.1$, and $k_{\rm off} = 0.8$. The initial value of $k_{\rm on}c_0 = 0.08$, and post-pulse $k_{\rm on}c_0 = 0.16$. Thus $\bar{x} = 1/11$ and $\bar{x}' = 1/6$. The deterministic equilibrium value of E is 1/2, both pre- and post-pulse.

V. DYNAMIC BEHAVIOR

Up to this point, we have focused exclusively on the steady-state properties of the model. It is also interesting to investigate the dynamic behavior of the model, especially since in the deterministic limit adaptation implies that the steady-state behavior is independent of the environment, and the only nontrivial behavior is the transient response of the system to changes in the input. In our distinguished S/A limit where the A dynamics is fast and the B dynamics is slow, we can get a complete picture of the dynamics. Imagine that the input signal suddenly changes to a new value. The A field will immediately respond to this change. The B field will only relax slowly to the change. As before, the relatively fast noise is averaged over by B, and so the B dynamics can be taken to be a deterministic exponential relaxation to its new equilibrium value B_0', on the slow time scale. Thus, we can consider the changes in B as adiabatic, with the system in quasi-equilibrium with the current value of B_0. Thus, the value of E, averaged over the relatively quick fluctuations will be that given by our above solution for $\langle E\rangle$ with $B_0 = B_0(t) = B_0' + (B_0' - B_0(t_p))e^{-\delta(t-t_p)}$, where t_p is the time of the pulse, and $k_{\rm on}c_0$ is given by its new value, with A_p and $k_{\rm off}$ unchanged. We see this plotted for the case of $N = 1$ in Fig. 9, together with the full simulation, averaged over 10^4 runs. We see that this treatment captures very well the dynamics on the long $1/\delta$ time scale of the B dynamics. We see that for this strong a noise, the transient response of the system is swamped by the shift in the equilibrium value of $\langle E\rangle$. One can in fact do better by solving the time-dependent moment equations, with $\Pi_1(t) = \bar{x}' - (\bar{x}' - \bar{x})e^{-\delta(t-t_p)}$. This resolves the initial period of the rise in $\langle E(t)\rangle$, during which the noise equilibrates to its new statistics. This is also shown in the figure.

VI. SUMMARY

In summary, we have shown that the Gaussian noise model in the context of nonlinear downstream reactions fails in principle beyond a critical value of β, the strength of the inhibitor production. This renders a full analytic treatment impossible. The cause of this can be traced to the interaction of the signaling nonlinearity with the fact that Gaussian noise does not respect the constraint that chemical signals must be positive. One can get rid of this effect by linearizing the reaction equation, but this completely eliminates the possibility of investigating the extent to which the perfect adaptation seen in the deterministic limit is undermined by signal stochasticity.

Even though it cannot be solved exactly, one can make analytic progress for the most interesting range of parameters, namely where the activation is fast and the inhibition slow, as compared to the noise time scale . These can be used as a guide to determining the actual deviation of $\langle E\rangle$ from its signal-independent mean-field value. Also, it is possible to devise a cutoff procedure which accurately predicts the results of the binomial model for intermediate and large N

Acknowledgments

We thank Yu-hai Tu for stimulating discussions. This work is supported by the NSF Center for Theoretical Biological Physics Grant No. PHY-0822283. The work of N. Bostani is supported in part by the National Natural Science Foundation of China, under grant numbers 11133002, 10821061, 11050110113, and by a research fellowship for international young scientists of the Chinese Academy of Sciences.

Appendix A: Gaussian Noise

We present here the formulas for Φ and $\Delta E(u)$ appearing in the expression for the average E for the Gaussian noise model.

$$\begin{aligned}\Phi(t) \;=\; & -\frac{\tau\alpha}{\gamma^2(1-\gamma^2\tau^2)}\left(\frac{\alpha}{\gamma}+\frac{2\beta}{\gamma+\delta}\right)\left(1-e^{-\gamma t}\right)\\ & -\frac{\tau\beta}{\delta^2(1-\delta^2\tau^2)}\left(\frac{\beta}{\delta}+\frac{2\alpha}{\gamma+\delta}\right)\left(1-e^{-\delta t}\right)\\ & -\tau^2\left(\frac{\alpha}{\gamma+1/\tau}+\frac{\beta}{\delta+1/\tau}\right)\left(\frac{\alpha}{\gamma-1/\tau}+\frac{\beta}{\delta-1/\tau}\right)\left(1-e^{-t/\tau}\right)\\ & +\tau\left(\frac{\alpha}{\gamma}+\frac{\beta}{\delta}\right)^2 t\end{aligned} \tag{A1}$$

$$\begin{aligned}\Delta_E(t) \;=\; & \frac{\tau\alpha}{\gamma^2(1-\gamma^2\tau^2)}\left(1-e^{-\gamma t}\right)+\frac{2\tau\beta}{\delta(\gamma+\delta)(1-\delta^2\tau^2)}\left(1-e^{-\delta t}\right)\\ & +\frac{\tau}{\gamma+1/\tau}\left(\frac{\alpha}{\gamma-1/\tau}+\frac{\beta}{\delta-1/\tau}\right)\left(1-e^{-t/\tau}\right)\end{aligned} \tag{A2}$$

Appendix B: Large N Limit

We begin by writing the exact equation satisfied by $\Delta_j \equiv z_j - E_0\Pi_j$:

$$0=-(N-j)k_{\rm on}c_0\Delta_j-jk_{\rm off}\Delta_j+(N-j+1)k_{\rm on}c_0\Delta_{j-1}+(j+1)k_{\rm off}\Delta_{j+1}-\left(B_0+\frac{jA_p}{N}\right)\Delta_j+\frac{j-N\bar{x}}{N}A_p(1-E_0)\Pi_j\ . \tag{B1}$$

where $\bar{x}\equiv k_{\rm on}c_0/(k_{\rm on}c_0+k_{\rm off})$. To proceed, we transform this difference equation into an ODE in terms of the variable $y\equiv(j-N\bar{x})/\sqrt{N}$, and writing $\Delta_j = N^{-1}z^{(-1)}(y)+N^{-3/2}z^{(2)}(y)$. To leading order in N we get ($\omega\equiv k_{\rm on}c_0+k_{\rm off}$):

$$\omega\left(\frac{d}{dy}yz^{(1)}(y)+\sigma^2\frac{d^2}{dy^2}z^{(1)}(y)\right)-(A_0+B_0)z^{(1)}=-\frac{A_pB_0}{(A_0+B_0)\sqrt{2\pi\sigma^2}}ye^{-y^2/2\sigma^2}\ , \tag{B2}$$

the solution of which is

$$z^{(1)}(y) = \frac{A_p B_0}{(A_0+B_0+\omega)(A_0+B_0)\sqrt{2\pi\sigma^2}} y e^{-y^2/2\sigma^2} \ . \tag{B3}$$

This correction reflects the breakdown of perfect adaptation for finite N, but due to its asymmetry, does not lead to a correction in $\langle E \rangle$. To obtain the leading order correction for this quantity, we have to go to next order:

$$\omega\left(\frac{d}{dy} y z^{(2)} + \sigma^2 \frac{d^2}{dy^2} z^{(2)}\right) - (A_0+B_0) z^{(2)} = y A_p z^{(1)} - \frac{\omega(1-2\bar{x})}{2} \frac{d^2}{dy^2} y z^{(1)} + \frac{y^2(1-2\bar{x})}{2\sigma^2\sqrt{2\pi\sigma^2}} \left(1 - \frac{y^2}{3\sigma^2}\right) \frac{A_p B_0}{A_0+B_0} e^{-y^2/2\sigma^2} \ . \tag{B4}$$

The solution is then

$$z^{(2)}(y) = -\frac{A_p B_0 e^{-y^2/2\sigma^2}}{(A_0+B_0)(A_0+B_0+\omega)\sqrt{2\pi\sigma^2}} \left[\frac{A_p \sigma^2}{(A_0+B_0+2\omega)} \left(\frac{y^2}{\sigma^2} + \frac{2\omega}{A_0+B_0}\right) + (\bar{x} - 1/2) \left(\frac{y^4}{3\sigma^4} - \frac{y^2}{\sigma^2}\right) \right] \ . \tag{B5}$$

The second term does not contribute to $\langle E \rangle$, interestingly enough, and

$$\langle E \rangle \approx E_0 - \frac{A_p^2 B_0 \sigma^2}{N(A_0+B_0)^2(A_0+B_0+\omega)} \ . \tag{B6}$$

[1] M.B. Elowitz, A.J. Levine, E.D. Siggia, and P.D. Swain, Science **207**, 1183 (2002); P.S. Swain, M.B. Elowitz, and E.D. Siggia, Proc. Natl. Acad. Sci. U.S.A. **99**, 12795 (2002).
[2] H.C. Berg and E.M. Purcell, Biophys. J. **20**, 193 (1977).
[3] B. Hu, W. Chen, W.J. Rappel, and H. Levine, Phys. Rev. Lett. **105**, 048104 (2010).
[4] Gašper Tkačik, T. Gregor, and W. Bialek, PLoS ONE **3**, e2774 (2008).
[5] Y. Tu and G. Grinstein, Phys. Rev. Lett. **94**, 208101 (2005).
[6] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry*, Elsevier (2007).
[7] W. Ma, A. Trusina, H. El-Samad, W. A. Lim, C. Tang, Cell **138**, 760 (2009).
[8] C.A. Parent and P.N. Devreotes, Science. **284**, 765 (1999).
[9] A. Levchenko, P. A. Iglesias, Biophys. J. **82**, 50 (2002)
[10] K. Takeda et a;, "Incoherent feedforward control governs adaptation of activated Ras in a eukaryotic chemotaxis pathway", submitted to Sci. Sig. (2011)
[11] A GEF (guanine nucleotide exchange factor) converts the inactive form of RAS (RAS-GDP) to its active form by changing the GDP to a GTP. A GAP (GTPase Activating Protein) catalyzes the reverse reaction.
[12] W.-J. Rappel and H. Levine, Proc. Natl. Acad. Sci. U.S.A. **105**, 19270 (2008); B. Hu, D. Fuller, W.F. Loomis, H. Levine, and W.-J. Rappel, Phys. Rev. E **81**, 031906 (2010).
[13] R.G. Endres and N.S. Wingreen, Proc. Natl. Acad. Sci. U.S.A. **105**, 15749 (2008); R.G. Endres and N.S. Wingreen, Phys. Rev. Lett. **103**, 158101 (2009).
[14] J. Ohkubo, N. Shnerb and D. A. Kessler, J. Phys. Soc. Japan **77**, 44002 (2008).
[15] D. A. Kessler and E. Barkai, Phys. Rev. Lett. **105**, 120602 (2010).
[16] Noise can be important for RAS activation by a spatial gradient; see D. Fuller, W. Chen, M. Adler, A. Groisman, H. Levine, W. J. Rappel and W. F. Loomis, "External and internal noise limits to eukaryotic chemotaxis", PNAS **107**, 9656 (2010).
[17] E. Brener, H. Levine, and Y. Tu, Phys. Rev. Lett. **66**, 1978 (1991).
[18] T. B. Kepler and A. S. Perelson, Proc. Natl. Acad. Sci. U.S.A. **92**, 8219 (1995).
[19] L. S. Tsimring, H. Levine and D. A. Kessler, Phys. Rev. Lett. **76**, 4440 (1996)
[20] E. Brunet and B. Derrida, Phys. Rev. E **56**, 2597 (1997)
[21] D. A. Kessler, Z. Ner, and L. M. Sander, Phys. Rev. E **58**, 107 (1998).
[22] L. Pechenik and H. Levine, Phys. Rev. E **59**, 3893 (1999).
[23] C. Mueller, L. Mytnik and J. Quastel, Invent. Math **184**, 405 (2011)