Optimal Prediction by Cellular Signaling Networks

Nils B. Becker,[1,*] Andrew Mugler,[2] and Pieter Rein ten Wolde[3]

[1]*Bioquant, Universität Heidelberg, Im Neuenheimer Feld 267, 69120 Heidelberg, Germany*
[2]*Department of Physics, Purdue University, West Lafayette, Indiana 47907, USA*
[3]*FOM Institute AMOLF, Science Park 104, 1098 XG Amsterdam, The Netherlands*

Living cells can enhance their fitness by anticipating environmental change. We study how accurately linear signaling networks in cells can predict future signals. We find that maximal predictive power results from a combination of input-noise suppression, linear extrapolation, and selective readout of correlated past signal values. Single-layer networks generate exponential response kernels, which suffice to predict Markovian signals optimally. Multilayer networks allow oscillatory kernels that can optimally predict non-Markovian signals. At low noise, these kernels exploit the signal derivative for extrapolation, while at high noise, they capitalize on signal values in the past that are strongly correlated with the future signal. We show how the common motifs of negative feedback and incoherent feed-forward can implement these optimal response functions. Simulations reveal that *E. coli* can reliably predict concentration changes for chemotaxis, and that the integration time of its response kernel arises from a trade-off between rapid response and noise suppression.

The ability to respond and adapt to changing environments is a defining property of life. Single-celled organisms employ a range of response strategies, tailored to the environmental fluctuations they encounter. Gradual changes in osmolarity, pH, or available nutrients are sensed and responded to adiabatically. In this regime, the sensory performance as measured by the mutual information between stimulus and response, limits the achievable growth rate [1–3]. In contrast, when environmental changes are rapid and unpredictable, sensing may be futile, since any response would come too late. Here, phenotypic heterogeneity can help by providing a subpopulation of preadapted cells [4]. An intermediate regime exists where environmental fluctuations occur with some regularity, on the cellular response time scale. It is then possible and desirable for the cell to *predict* the future environment, in order to initiate a response ahead of time. When the cellular response takes a finite time τ to become effective, the *predictive* mutual information between the current sensory output and the environment τ later, limits growth [5]. Sensing strategies that leverage correlations of a stimulus with future environmental changes have indeed been observed, and reevolved experimentally [6,7].

This raises the question of what makes a cellular network an optimal predictor, rather than instantaneous reporter, of the environment. Intuitively, to predict, one should rely on the most up-to-date information, i.e., respond to the current input. However, cells often sense non-Markovian (NM) signals, whose past trajectories could add useful information. Intriguingly, in such cases, sensory networks often react not instantaneously but instead more slowly, on the time scale of the signal [8,9].

A slow network time integrates the input signal, which may dampen the response, but can also enhance the estimate of the current input signal by filtering noise from, e.g., receptor-ligand binding [10–15]. Moreover, a slow response may enhance prediction by building a memory of the signal history which is informative about the future signal. What features of signal and response then make a noninstantaneous response beneficial for prediction?

Here, we study how the accuracy of prediction depends on the noise and correlations in the input, the forecast interval, and the design of the response system. We find that single-layer responders, such as push-pull networks, can improve prediction by responding slowly. This not only allows noise averaging, but also enables reading out past signals that are more correlated with the future signal than the current signal is. Multilayer networks can further enhance prediction via nonmonotonic response functions tailored to the input. They can optimally predict low-noise signals by exploiting the signal derivative, and high-noise signals by coherently summing informative past signal values. This can be implemented via negative feedback. Finally, we perform simulations of *E. coli* bacteria that chemotax in spatially varying concentration fields. The simulations reveal that *E. coli* chemotaxis relies on predicting future concentration changes. They suggest that the optimal integration time of the kernel arises as a compromise between the benefit of responding quickly to the most recent concentration values, and the need to filter input noise.

Consider a general sensory network that responds to a time-varying extracellular signal by binding ligand molecules, relaying the signal via intermediate species, and

finally producing an output species [Fig. 1(a)]. Its pre-
diction capability depends on both responder and input
properties. Concerning the input, prediction fundamentally
requires that past inputs contain information about the
future; i.e., the signal's conditional probability density
$p(s(t + \tau)|s(t), s(t'), \ldots)$ really depends on the signal
values at $t > t' > \cdots$. For Markovian input, the only
dependence is on $s(t)$, and perfect instantaneous readout
of $s(t)$ would in fact be the optimal prediction strategy for
all future $s(t + \tau)$ [5]. However, in the presence of input
noise ξ, arising from, e.g., receptor-ligand binding, the
responder senses the degraded signal $\ell(t) = s(t) + \xi(t)$.
Then even for Markovian s, the added noise makes
$p(s(t + \tau)|\ell(t), \ell(t'), \ldots)$ dependent on past values
$\ell(t'), \ldots$, since they help determine the current input $s(t)$
by averaging over the noise ξ, and then from $s(t)$ the future
$s(t + \tau)$. Thus a slow response can help prediction of
any noisy signal via the mechanism of time integration
(which also improves accuracy for constant, noisy signals
[10–17]). As detailed below, for NM signals, another
prediction mechanism exists: a responder with memory
enables readout of additional information from past
signals $s(t'), \ldots$, improving predictions by exploiting signal
correlations.

We take the input signal $s(t)$ to be stationary Gaussian,
characterized by $\langle s(t')s(t' + t)\rangle = \sigma_s^2 r_s(t)$ where r_s denotes
the normalized autocorrelation function, and σ_s, the signal
amplitude. For Markovian processes, $r_s(t) = \exp(-\lambda t)$. A
family of NM signals can be generated via a harmonic
oscillator defined by $\partial_{\omega t} q = p, \partial_{\omega t} p = -q - \eta p + \sqrt{2\eta}\psi$
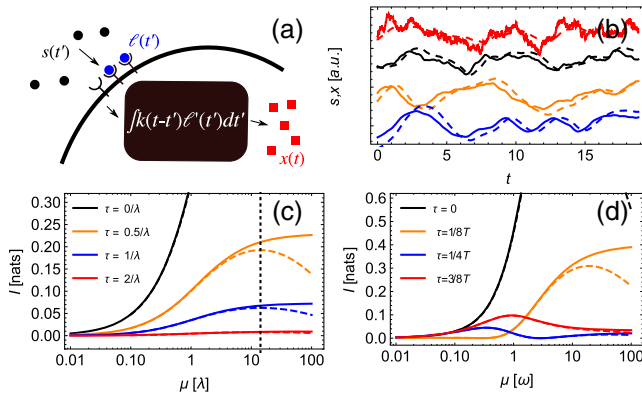

FIG. 1 (color online).   Biochemical prediction. (a) A sensory
network (black box) with output x (red) responding to an
extracellular ligand s (black) via noisy ligand-bound receptors
ℓ (blue). (b) Example traces (solid lines) of Markovian ($\lambda = 1$,
red) and non-Markovian (NM) signals s ($\omega = 1$, $\eta = 4, 2, 0.5$,
downwards), yielding outputs x of exponential responders
with $\mu = 1$ (dashed). Predictive information $I[x, s_\tau]$ in nats
(ln 2nats = 1bit) for Markovian signals (c) and NM signals with
damping $\eta = 1/2$ (d), as a function of the speed μ of an
exponential responder, for different prediction intervals τ, and
for noise level $\vartheta = 0$ (solid lines) and $\vartheta = 0.1$ (dashed lines).
Dotted line in (c) denotes μ_{ti}.

with unit white noise ψ, by letting $s \equiv q$; see Fig. 1(b).
The damping parameter η controls the signal statistics:
in the overdamped regime $\eta > 2$, $r_s(t)$ is monotonically
decreasing, while for $\eta < 2$ it is oscillatory with a period
approaching $T = 2\pi/\omega$; in both cases, the signal s obeys
Gaussian statistics. This family of signals allows analytical
results and interpolates from Markovian to non-Markovian,
long-range correlated, oscillatory signals. We model input
noise as white, $\langle \xi(t)\xi(t')\rangle = \sigma_s^2 \vartheta^2 \delta(t - t')$, where ϑ is the
relative noise strength.

Concerning the responder, we focus on linear signal-
ing networks [12,18] which afford analytical results and
often describe information transmission remarkably well
[19–21]. Because we are interested in how prediction
depends on the correlations and noise in the input, we
consider responders in the deterministic limit. The output
$x(t) = \int_{-\infty}^{t} k(t - t')\ell(t')dt'$ of the network is then deter-
mined by its linear response function $k(t)$.

The predictive power of a signal-responder system is
measured in a rigorous and biologically relevant way [5] by
the predictive mutual information $I[x, s_\tau] = \langle\log\{p(x, s_\tau)/$
$[p(x)p(s_\tau)]\}\rangle$ between the current output $x(t)$ and the
future input $s_\tau \equiv s(t + \tau)$. Because x is jointly Gaussian
with the input, the predictive information reduces to a
function $I[x, s_\tau] = -\frac{1}{2}\log(1 - r_{xs_\tau}^2)$ of the input-output
correlation coefficient

$$r_{xs_\tau} = \frac{\Psi(\tau)}{[\Sigma + \Xi]^{1/2}}. \qquad (1)$$

The overlap integral $\Psi(\tau) \equiv \int_0^\infty k(t)r_s(t + \tau)dt$ is the part
of the normalized output variance σ_x^2/σ_s^2 that is correlated
with the prediction target s_τ. The denominator splits σ_x^2/σ_s^2
into contributions from past signal, $\Sigma \equiv \int_0^\infty k(t)r_s(t - t') \times$
$k(t')dtdt'$, and past noise $\Xi \equiv \vartheta^2 \int_0^\infty k(t)^2 dt$ [5].

We first consider a push-pull network, consisting of a
single layer in which the output x is directly activated
by the receptor. It is characterized by an exponential kernel
$k(t) \propto \exp(-\mu t)$ with response speed μ. Figure 1(c) shows
how accurately such a network can predict Markovian
signals, as measured by the predictive information I,
obtained analytically from Eq. (1) [5]. Without input noise
($\vartheta \to 0$), the fastest responders maximize the accuracy I,
as expected. When including input noise, there exists an
optimal response speed $\mu_{\mathrm{ti}} = (2\lambda/\vartheta^2 + \lambda^2)^{1/2}$, independent
of τ, and approaching $\mu_{\mathrm{ti}} \to \lambda$ for high noise [22]. The
optimum arises from a trade-off between rapid tracking of
the input and noise averaging [5].

Figure 1(d) shows I for exponential responders predict-
ing oscillatory ($\eta = 0.5$) NM signals. As before, input noise
disfavors the fastest responders. Interestingly, however, a
finite response speed can be optimal even when there is no
input noise ($\vartheta = 0$): For prediction intervals above about a
quarter period, frequency-matched responders with $\mu_* \simeq \omega$
(obtained numerically [5]) perform best.

The optimal μ_* is not an effect of simple time integration but rather results from exploiting the oscillatory signal correlations. When the forecast interval $\tau \ll T$, $r_{xs_\tau}^2$ is maximized by increasing the overlap $\Psi(\tau)^2$ via a short kernel k that samples high values of the input correlation function r_s; see Figs. 2(a) and 2(b). The optimal kernels never become instantaneous, however, since that would strongly increase Ξ. As τ increases, μ_* initially increases: $k(t)$ decays faster so that it continues to overlap with the positive lobe of $r_s(t + \tau)$; input and output remain positively correlated. Surprisingly, beyond a critical prediction interval $\tau_c \simeq 0.22T$, μ_* drops discontinuously [Fig. 2(b), solid to dashed line]. The response now integrates the negative lobe of r_s, anticorrelating output and input [Fig. 2(a)]. Effectively, the output x lags behind the input s by an amount Λ, so that the current output $x(t)$ reflects the past input $s(t - \Lambda)$ rather than the current input $s(t)$. This enhances prediction, because the past signal $s(t - \Lambda)$ is more (anti)correlated with, and hence more informative about, the future $s(t + \tau)$ than the present signal $s(t)$ is, as shown by the nonmonotonic signal autocorrelation function: $r_s(\Lambda + \tau)^2 > r_s(\tau)^2$ (cf. Fig. S2 in [5]). The optimal response speed μ_* is such that $\tau + \Lambda \simeq T/2$; the response kernel $k(t)$ then probes r_s around its minimum, maximizing the squared overlap $\Psi(\tau)^2$ between them [Fig. 2(a)]. As τ increases further, increasing μ keeps the kernel localized in the negative lobe of r_s, until another transition at higher $\tau \simeq 0.75T$ focuses the response on the next positive lobe of r_s. Simulations confirmed this mechanism also for nonlinear responders and various input waveforms and noise strengths [5].

Signaling networks typically consist of more than one layer [23], generating complex kernels. To explore the design space, we maximize the predictive information over all kernels. For Gaussian signals, this is equivalent [5] to finding the optimal kernel k_* that minimizes the mean squared prediction error $\langle (x - s_\tau)^2 \rangle$, as in Wiener-Kolmogorov filter theory [12,22,24,25], used below.

The resulting optimal kernel remains exponential for input signals that are Markovian [5], so that $k_*^M(t) \propto \exp(-\mu_{ti}t)$, with μ_{ti}, as before, implementing time integration. Hence, a single, slowly responding, push-pull network layer is enough to perform globally optimal predictions of noisy Markovian signals; additional network layers cannot enhance prediction.

For NM but overdamped signals ($\eta > 2$), optimal kernels k_* have an almost exponential shape, which is insensitive to the prediction interval, Fig. 2(c) (see [5]). This indicates a prediction strategy based mainly on time integration to determine the current $s(t)$.

In contrast, oscillatory NM signals with $\eta < 2$ yield optimal kernels that are oscillatory, Fig. 2(d). Their shape depends on the prediction interval τ, and on the correlations and noise in the input [Fig. 3(a)]. At low noise, optimal kernels integrate only a short time window. They consist of a sharply peaked positive lobe followed by an undershoot, effectively estimating the future signal value from its current value and derivative [5]. This strategy of linear extrapolation avoids including past signals, which are inherently less correlated with the future. The capability to take derivatives enables a rapid response even when the current signal value carries no predictive information, $r_s(\tau) = 0$; in contrast, in this situation exponential responders would need to respond slowly, to pick up past, informative signals (Fig. 2).
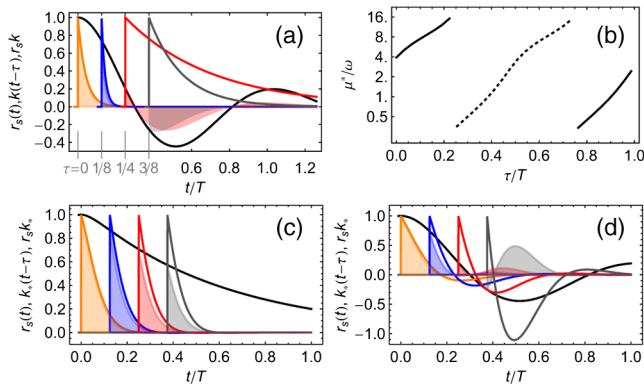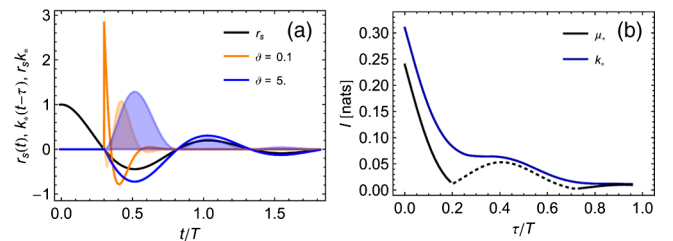

FIG. 2 (color online). Prediction by optimizing correlations. (a) NM autocorrelation r_s at $\eta = 0.5$ (black). Optimally predictive exponential kernels $k(t - \tau)$ at noise level $\vartheta = 0.25$ for $\tau/T = 1/8, 1/4, 3/8$ as indicated (orange, blue, red, gray, respectively), with corresponding overlap integrands $r_s(t)k(t - \tau)$ (shaded). (b) Optimal μ_* vs prediction interval for these parameters. Solid, dashed lines: positive, negative sx correlation, respectively. (c) As (a) at the same τ values but for the globally optimal kernels k_* and overdamped signal $\eta = 4$. (d) As (c) but for underdamped signal at $\eta = 0.5$.

FIG. 3 (color online). (a) The optimal kernel depends on the correlations and the noise ϑ in the input signal. Autocorrelation function (black) and optimal kernels k_* at $\tau = 0.3$ for oscillatory NM signals with $\eta = 0.5$, for two different noise levels ϑ. At low noise, the kernel consists of a positive lobe followed by a single undershoot. This corresponds to prediction based on linearly extrapolating the current signal. In contrast, at high noise, the kernel echoes the signal correlation function, exploiting signal values in the past that correlate with the future. (b) The optimal kernels k_* strongly improve prediction over optimal exponential kernels μ_* [as in Fig. 2(b)] around $\tau = \tau_c \approx 0.2T$; $\vartheta = 1.0$; for low noise, see Fig. S3 [5].

As noise levels ϑ rise, noise averaging becomes increasingly important, which demands longer kernels. However, to avoid signal damping, the optimal kernel must *coherently* sum past signal values. For oscillatory signals, this requires an oscillatory kernel, which integrates the signal with alternating signs. The prediction enhancement of globally optimal kernels over optimal exponential kernels is indeed largest for oscillatory input signals and, for $\tau \approx \tau_c$, it can reach up to 400% [Fig. 3(b)]. Interestingly, in the limit $\vartheta \to \infty$, maximizing Eq. (1) gives the simple result $k_*(t) \propto r_s(t+\tau)$ [5], showing that at high noise, the optimal kernel mimics the input correlations.

Optimal oscillatory kernel shapes like k_* in Fig. 3 can be implemented via negative feedback [5], a common motif in gene networks and signaling pathways [26–28]. Another common motif, incoherent feed-forward [23], only allows kernels with a positive lobe followed by a single undershoot [5]. Our results show that this is useful for predicting low-noise non-Markovian signals, but suboptimal at high noise.

In summary, accurate prediction requires capitalizing on past signal features that are correlated with the future signal, while minimizing transmission of uncorrelated past signals and noise. Single-layer networks suffice to predict Markovian signals optimally by noise averaging. Multilayer networks predict oscillatory signals optimally, by fast linear extrapolation at low noise, and by coherent summation at high noise. In the high noise limit, the optimal network response mimics the input: $k_*(t) \propto r_s(\tau+t)$.

To explore the importance of predictive power in cellular behavior, we have studied *E. coli* chemotaxis. *E. coli* moves by alternating straight *runs* with *tumbles*, which randomly reorient it. In a spatially varying environment, this motion is biased via a signaling pathway, whose output $x(t)$ controls the propensity $\alpha(t)$ that a running bacterium will tumble. We have performed simulations of chemotaxing bacteria in static concentration fields $c(\vec{r})$ in two dimensions, using the measured response kernel k [29–31]. At low concentrations, the signaling noise is dominated by the input noise. As in our theory, we therefore ask how the predictive power depends on the kernel and the input noise, ignoring intrinsic noise [32]. The tumbling propensity is then given by $\alpha(t) = \alpha_0[1 - x(t)]$, where $\alpha_0 = 1/\text{s}$ is the basal tumbling rate and $x(t) = \int_{-\infty}^{t} k(t-t')\ell(t')dt'$. The input $\ell(t') = s(t') + \xi(t')$ depends on the concentration signal $s(t) = c[\vec{r}(t)]$ and the input noise $\xi(t)$ of relative strength θ, arising e.g., from receptor-ligand binding or receptor conformational dynamics. The kernel $k(t)$ is adaptive, i.e., it integrates to 0, which allows the bacterium to respond to a wide range of background concentrations [29,30]. We compare adaptive kernels of varying range defined by $k_\nu(t) \equiv \nu^2 k(\nu t)$, where ν defines the response speed (see also [5]).

The sensory output modulates the *delay* $\sim 1/\alpha(t)$ to the next tumble. This suggests that high chemotactic efficiency requires accurate signal prediction. However, it is less
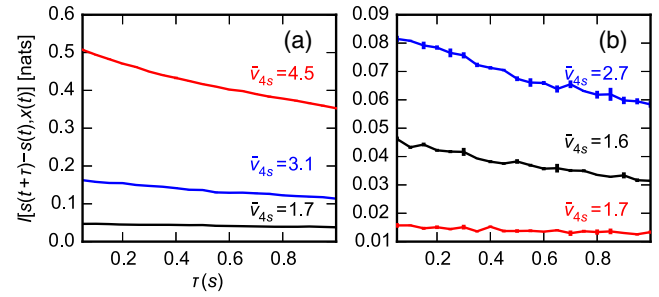

FIG. 4 (color online). The predictive power of *E. coli* in a sinusoidal concentration field with period $L = 400~\mu\text{m}$, generating a nonoscillatory input signal [5], as a function of the forecast interval τ. Information is shown for the wild-type *E. coli* kernel with $\nu = 1$ (blue), for a faster kernel with $\nu = 3$ (red) and a slower kernel with $\nu = 0.5$ (black), for noise levels $\theta = 0$ (a) and 2 (b). Corresponding chemotactic speeds $\bar{v}_{\delta t = 4s}$ are given in $\mu\text{m}/\text{s}$. Bacteria run at $20~\mu\text{m}/\text{s}$.

obvious what feature of the signal the system actually predicts: The future concentration? Or the change in concentration? More generally, what are the relevant input and output variables that control chemotaxis? Only for these variables can we expect that chemotactic performance is correlated with predictive information.

To address this question, we performed simulations for three different kernels, $\nu = 0.5, 1, 3$ where $\nu = 1$ corresponds to the measured kernel, and for two input noise levels, $\theta = 0, 2$. As our performance measure, we use the mean chemotactic speed $\bar{v}_{\delta t} = \langle [\vec{r}(t + \delta t) - \vec{r}(t)] \cdot \nabla c[\vec{r}(t)] / \{ \delta t \| \nabla c[\vec{r}(t)] \| \} \rangle$; similar results are obtained for the mean concentration $\langle c[\vec{r}(t)] \rangle$ [5]. We find that $\bar{v}_{\delta t}$ is poorly correlated with the predictive information $I[x, s_\tau]$ between current output and future concentration [5]. In contrast, it is well correlated with the predictive information $I[x, s_\tau - s]$ between current output and future concentration *change*, as Fig. 4 shows. Hence, the search strategy of *E. coli* is not based on predicting the future concentration but rather its trend, in accordance with the observation that the bilobed kernel k takes a time derivative of the signal. If this is positive, *E. coli* "expects" that the concentration will continue to rise and will extend its run.

Figure 4 also shows that the optimal kernel that maximizes the information and hence chemotactic speed depends on the input noise θ. A fast kernel emphasizes up-to-date information about recent concentration changes, enabling an accurate and rapid response at low noise. At high noise, its performance drops because it cannot filter the input noise and hence cannot reliably predict future concentration changes. The optimal kernel range then arises from a trade-off between agility and robustness.

Lastly, how far must *E. coli* look into the future for efficient chemotaxis? To anticipate concentration changes, the prediction horizon, i.e., the time over which predictive information extends, should exceed the response time. According to Eq. (1) the prediction horizon is bounded

by the signal correlation time, which is determined by the length scale of the concentration field and by the motile behavior, to be explored in future work. Already, Fig. 4 indicates that the prediction horizon of *E. coli* is indeed longer than the response time, as $I(\tau)$ decays slower than $1/\alpha_0 = 1$s. Our results thus suggest that *E. coli* can indeed reliably anticipate concentration changes.

---

*nbecker@bioquant.uni-heidelberg.de

[1] C. T. Bergstrom and M. Lachmann, arXiv:q-bio/0510007.
[2] S. F. Taylor, N. Tishby, and W. Bialek, arXiv:0712.4382.
[3] W. Bialek, *Biophysics: Searching for Principles* (Princeton University Press, Princeton, NJ, 2012).
[4] N. Q. Balaban, J. Merrin, R. Chait, L. Kowalik, and S. Leibler, Science **305**, 1622 (2004).
[5] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.115.258103 for detailed derivations and additional simulation results.
[6] A. Mitchell, G. H. Romano, B. Groisman, A. Yona, E. Dekel, M. Kupiec, O. Dahan, and Y. Pilpel, Nature (London) **460**, 220 (2009).
[7] I. Tagkopoulos, Y.-C. Liu, and S. Tavazoie, Science **320**, 1313 (2008).
[8] U. B. Kaupp, N. D. Kashikar, and I. Weyand, Annu. Rev. Physiol. **70**, 93 (2008).
[9] J. Valencia S., K. Bitou, K. Ishii, R. Murakami, M. Morishita, K. Onai, Y. Furukawa, K. Imada, K. Namba, and M. Ishiura, Genes Cells **17**, 398 (2012).
[10] H. C. Berg and E. M. Purcell, Biophys. J. **20**, 193 (1977).
[11] W. Bialek and S. Setayeshgar, Proc. Natl. Acad. Sci. U.S.A. **102**, 10040 (2005).
[12] C. C. Govern and P. R. ten Wolde, Phys. Rev. Lett. **109**, 218103 (2012).
[13] P. Mehta and D. J. Schwab, Proc. Natl. Acad. Sci. U.S.A. **109**, 17978 (2012).
[14] K. Kaizu, W. de Ronde, J. Paijmans, K. Takahashi, F. Tostevin, and P. R. ten Wolde, Biophys. J. **106**, 976 (2014).
[15] C. C. Govern and P. R. ten Wolde, Proc. Natl. Acad. Sci. U.S.A. **111**, 17486 (2014).
[16] C. C. Govern and P. R. ten Wolde, Phys. Rev. Lett. **113**, 258102 (2014).
[17] G. Aquino, L. Tweedy, D. Heinrich, and R. G. Endres, Sci. Rep. **4**, 5688 (2014).
[18] R. Heinrich, B. G. Neel, and T. A. Rapoport, Mol. Cell **9**, 957 (2002).
[19] S. Tănase-Nicola, P. B. Warren, and P. R. ten Wolde, Phys. Rev. Lett. **97**, 068102 (2006).
[20] E. Ziv, I. Nemenman, and C. H. Wiggins, PLoS One **2**, e1077 (2007).
[21] W. H. de Ronde, F. Tostevin, and P. R. ten Wolde, Phys. Rev. E **82**, 031914 (2010).
[22] M. Hinczewski and D. Thirumalai, Phys. Rev. X **4**, 041017 (2014).
[23] U. Alon, *An Introduction to Systems Biology: Design Principles of Biological Circuits*, 1st ed. (Chapman and Hall, CRC, Boca Raton, 2006).
[24] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series* (Wiley, New York, 1950).
[25] A. N. Kolmogorov, *Selected Works of A. N. Kolmogorov: Probability Theory and Mathematical Statistics* (Springer Science & Business Media, Netherlands, 1992).
[26] T. Kondo, T. Mori, N. V. Lebedeva, S. Aoki, M. Ishiura, S. S. Golden, Science **275**, 224 (1997).
[27] B. N. Kholodenko, Eur. J. Biochem. **267**, 1583 (2000).
[28] S. Sasagawa, Y.-i. Ozaki, K. Fujita, and S. Kuroda, Nat. Cell Biol. **7**, 365 (2005).
[29] J. E. Segall, S. M. Block, and H. C. Berg, Proc. Natl. Acad. Sci. U.S.A. **83**, 8987 (1986).
[30] S. M. Block, J. E. Segall, and H. C. Berg, Cell **31**, 215 (1982).
[31] A. Celani and M. Vergassola, Proc. Natl. Acad. Sci. U.S.A. **107**, 1391 (2010).
[32] H. Berg, *E. Coli in Motion*, Biological and Medical Physics Biomedical Engineering (Springer, New York, 2004).