A universal tradeoff between power, precision and speed in physical communication

Subhaneil Lahiri,[1,*] Jascha Sohl-Dickstein,[1,2,†] and Surya Ganguli[1,‡]

[1]*Department of Applied Physics, Stanford University, Stanford CA 94305, USA*
[2]*Google Research, Mountain View CA 94043, USA*

Maximizing the speed and precision of communication while minimizing power dissipation is a fundamental engineering design goal. Also, biological systems achieve remarkable speed, precision and power efficiency using poorly understood physical design principles. Powerful theories like information theory and thermodynamics do not provide general limits on power, precision and speed. Here we go beyond these classical theories to prove that the product of precision and speed is universally bounded by power dissipation in any physical communication channel whose dynamics is faster than that of the signal. Moreover, our derivation involves a novel connection between friction and information geometry. These results may yield insight into both the engineering design of communication devices and the structure and function of biological signaling systems.

Evolution has discovered remarkably rapid, precise, and energy efficient mechanisms for biological computation. For example the human brain performs myriad computations, including complex object recognition in 150 ms [1] while consuming less than 20 watts [2]. In contrast supercomputers operate in the megawatt range, and cannot yet rival general human performance. To both understand the design principles governing biological computation, and to exploit such principles in engineered systems, it is essential to develop a general theoretical understanding of the relationship between power, precision and speed in computation. For example, what are the fundamental limits and design tradeoffs involved in simultaneously optimizing these three quantities?

Existing general theories, while powerful, often elucidate fundamental limits on at most two of these quantities. For example, information theory [3, 4] provides limits on the accuracy of communication under power constraints. But achieving such limits may require coding messages in asymptotically large blocks, thus providing no theoretical guarantees on speed (though see recent work [5, 6] on finite block length coding). Thermodynamics, through the second law, places fundamental limits on the work required to implement a physical process. But achieving such limits on thermodynamic efficiency requires quasistatic processes that unfold infinitely slowly. More recent work has elucidated the minimal energy required to perform a physical process in finite time [7, 8]), but does not address accuracy in any computation. Landauer [9–11] revealed that the erasure of information sets a lower bound on the energy consumed in computation. This observation inspired reversible computing [12], which can achieve accurate computation at asymptotically zero energy expenditure, but at the expense of requiring asymptotically infinite time in the presence of noise.

In the absence of general theories governing performance limits of computation at finite power, precision and speed, many works in systems biology have focused on tradeoffs between subsets of these quantities in very specific chemical kinetic schemes for specific computations. Fundamental work on kinetic proofreading studied two way trade-offs between energy and accuracy [13–21] or speed and accuracy [22] in the communication of genetic information. Also, many works have studied specific tradeoffs between energy and precision in cellular chemosensation [23–29]. Notably, [30] studied simultaneous tradeoffs between power, speed and accuracy, but again in a very specific scheme for sensory adaptation.

Here we derive a very general three-way performance limit on power, precision and speed in physical communication. We focus on the problem of communication as it is a fundamental prerequisite for more complex computations. Indeed, in modern parallel computing, communication between processors, not computation within processors, presents an essential bottleneck for energy efficiency [31]. Our derived performance limit applies to *any* Markovian communication channel whose internal dynamics is faster than dynamics of the external signal to be communicated. In such a scenario, the external signal drives the communication channel into a non-equilibrium regime, in which the power dissipated can be described through a thermodynamic friction tensor on a manifold of channel state distributions [32–36]. We derive a lower bound on this friction tensor in terms of Fisher information, a fundamental quantity in the geometry of information [37]. By developing a novel inequality relating friction, which governs energy dissipation, to information geometry, which governs accuracy in statistical estimation, we derive our general relation between power, precision and speed. In essence, we find that the product of precision and speed is bounded by power.

Physical channels coupled to external signals. We model the communication channel as a physical system in contact with a thermal bath at inverse temperature $\beta = 1/k_{\mathrm{B}}T$. The channel is coupled to an n dimensional signal $\boldsymbol{\lambda}$, specified by components λ^μ, $\mu = 1 \dots n$, so that the energy of the channel in microstate i is $E_i(\boldsymbol{\lambda})$. When the external signal is held at a fixed $\boldsymbol{\lambda}$, we assume the channel relaxes to an equilibrium Boltzmann distribution

$$\pi_i(\boldsymbol{\lambda}) = \mathrm{e}^{-\beta[E_i(\boldsymbol{\lambda}) - \mathcal{F}(\boldsymbol{\lambda})]}, \tag{1}$$

where $\mathcal{F}$ is the free energy. We describe the non-equilibrium dynamics of the channel by a continuous-time Markov process, where the transition rate from state i to state j is K_{ij} and $K_{ii} = -\sum_{j\neq i} K_{ij}$. We assume the dynamics satisfies detailed balance:

$$\pi_i(\lambda)K_{ij}(\lambda) = \pi_j(\lambda)K_{ji}(\lambda). \tag{2}$$

Thus the external signal modifies the channel dynamics by directly modulating the transition rates (Fig. 1A).

Under the dynamics in (2), for fixed external signal $\boldsymbol{\lambda}$, the channel state distribution relaxes to (1), yielding a manifold of equilibrium channel state distributions parameterized by $\boldsymbol{\lambda}$. However, signals varying in time at a finite speed will drive the channel state distribution off the equilibrium manifold into a non-equilibrium distribution $\mathbf{p}(t)$. This distribution will be distinct from the equilibrium distribution $\boldsymbol{\pi}(\lambda(t))$ associated with the instantaneous value of the external signal (Fig. 1B). By driving the channel at finite speed, temporally varying signals perform physical work on the channel. Some of this work contributes to a change in free energy of the channel, while the rest is irreversibly dissipated as heat into the thermal bath. Thus temporally varying signals yield a dissipation of excess power. The non-equilibrium distribution $\mathbf{p}(t)$ also contains information about the history of the signal $\boldsymbol{\lambda}(t)$. Thus a downstream observer that can measure the channel microstate could use this information to estimate the signal with some level of precision. Below we discuss in further detail the nature of signal speed, channel power dissipation, channel information geometry, and estimation precision, and we derive general relations between these quantities. Moreover, in [38, §5] we discuss an extension of our results to situations where the channel dynamics breaks detailed balance, and the manifold of equilibrium distributions ((1) and Fig. 1B) is replaced with a manifold of non-equilibrium steady states.

Power dissipation and the friction tensor. In general, because the non-equilibrium distribution $\mathbf{p}(t)$ depends on the entire history of the past temporal signal $\boldsymbol{\lambda}(t')$ for $t' < t$, the power dissipation due to a changing signal can also depend on this entire history. However, if the temporal signal $\boldsymbol{\lambda}(t)$ varies more slowly than the channel dynamics (see [38, §3] for a precise description of this slow signal regime), then the non-equilibrium channel distribution $\mathbf{p}(t)$ remains close to the equilibrium manifold in Fig. 1B, and the excess power dissipation at time t depends on the signal history only through its instantaneous value $\boldsymbol{\lambda}(t)$ and time derivative $\dot{\boldsymbol{\lambda}}(t)$ [32]:

$$\mathcal{P}_{\text{ex}} = \sum_{\mu\nu} g_{\mu\nu}(\boldsymbol{\lambda})\,\dot{\lambda}^{\mu}\dot{\lambda}^{\nu}, \tag{3}$$

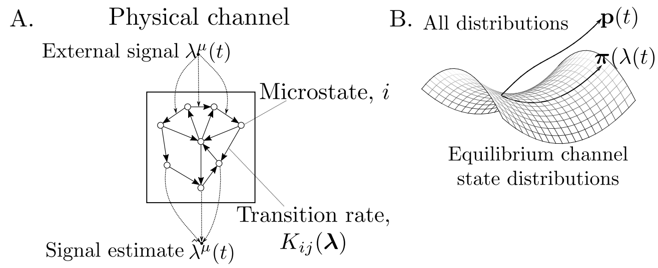

FIG. 1: Modeling physical channels coupled to external signals. A. An external signal $\lambda^\mu(t)$ modulates the transition rates of an arbitrary continuous-time Markovian dynamical system, modeling a physical channel in contact with a heat bath. A downstream receiver can construct an estimate $\hat{\lambda}^\mu$ of the instantaneous signal by observing the instantaneous microstate of the channel. B. A manifold of equilibrium channel state distributions $\boldsymbol{\pi}(\boldsymbol{\lambda})$ with intrinsic coordinates given by constant signal values $\boldsymbol{\lambda}$. Temporally varying signals $\boldsymbol{\lambda}(t)$ drive the channel microstates through a trajectory of non-equilibrium distributions $\mathbf{p}(t)$, off the equilibrium manifold.

where $g_{\mu\nu}$ is a friction tensor on the signal manifold,

$$\begin{gathered} g_{\mu\nu}(\boldsymbol{\lambda}) = (k_{\mathrm{B}}T)\int_0^\infty \mathrm{d}t' \,\left\langle \delta\phi_\mu(0)\,\delta\phi_\nu(t')\right\rangle, \\ \phi^i_\mu = -\beta\frac{\partial E_i}{\partial\lambda^\mu}, \qquad \delta\phi^i_\mu = \phi^i_\mu - \langle\phi_\mu\rangle\,. \end{gathered} \tag{4}$$

Here expectations are computed with respect to the equilibrium distribution $\boldsymbol{\pi}(\boldsymbol{\lambda})$, and derivatives are computed at the point $\boldsymbol{\lambda}$. ϕ^i_μ is the conjugate force exerted by the channel in response to changing a single signal component λ^μ when the channel is in microstate i. Thus the statistics of force fluctuations at equilibrium combined with finite signal velocity determines excess power dissipation out of equilibrium, in the slow signaling limit.

From friction to information geometry. We now derive a lower bound on the friction tensor for the models of physical channels described above (see [38, §2] for more details). First, the force correlation in (4) can be written as

$$\langle\delta\phi_\mu(0)\delta\phi_\nu(t')\rangle = \sum_{ij} p_{ij}(0,t')\delta\phi^i_\mu\delta\phi^j_\nu, \tag{5}$$

where $p_{ij}(t,t') = \pi_i\left[\exp(\mathbf{K}(\boldsymbol{\lambda})\,(t'-t))\right]_{ij}$ is the probability of being in state i at time t and in state j at a later time t', under equilibrium dynamics at at a constant external signal $\boldsymbol{\lambda}$. To simplify the matrix exponential, it is useful to employ an eigendecomposition of the rate matrix: $\mathbf{K} = -\sum_a q_a\mathbf{u}^a\boldsymbol{\eta}^a$. Here $\mathbf{u}^a$ are column vectors obeying $\mathbf{K}\mathbf{u}^a = -q_a\mathbf{u}^a$, $\boldsymbol{\eta}^a$ are row vectors obeying $\boldsymbol{\eta}^a\mathbf{K} = -q_a\boldsymbol{\eta}^a$, and they further obey the normalization condition $\boldsymbol{\eta}^a\mathbf{u}^b = \delta^{ab}$. With detailed balance, the eigenrates q_a are real and positive, ordered in increasing order, and the eigenvectors can be chosen to be real, satisfying $\eta^a_i = \pi_i u^a_i$. The slowest eigenmode is the $a=0$ stationary mode, with $q_0 = 0$, $\boldsymbol{\eta}^0 = \boldsymbol{\pi}$ and $\mathbf{u}^0 = \mathbf{e}$, a column vector of ones. We assume that the Markov dynamics is ergodic, so the 0th eigenvalue of $\mathbf{K}$ is non-degenerate.

Now inserting (5) into (4), transforming to the eigenbasis of $\mathbf{K}$, and integrating over time t' yields

$$\begin{aligned} g_{\mu\nu} &= k_{\mathrm{B}}T \sum_{a>0} \tau_a \left(\boldsymbol{\eta}^a \cdot \delta\boldsymbol{\phi}_\mu\right)\left(\boldsymbol{\eta}^a \cdot \delta\boldsymbol{\phi}_\nu\right) \\ &\geq k_{\mathrm{B}}T\,\tau_{\min} \sum_{a>0} \left(\boldsymbol{\eta}^a \cdot \delta\boldsymbol{\phi}_\mu\right)\left(\boldsymbol{\eta}^a \cdot \delta\boldsymbol{\phi}_\nu\right) \\ &= k_{\mathrm{B}}T\,\tau_{\min} F_{\mu\nu}, \end{aligned} \tag{6}$$

where $\tau_a = 1/q_a$, $\tau_{\min} = \min_{a>0} \tau_a$, and

$$F_{\mu\nu} = \sum_i \pi_i(\boldsymbol{\lambda}) \left[\partial_{\lambda_\mu} \ln \pi_i(\boldsymbol{\lambda})\right]\left[\partial_{\lambda_\nu} \ln \pi_i(\boldsymbol{\lambda})\right] \tag{7}$$

is the Fisher information. The inequality (6) means that $g_{\mu\nu} - k_{\mathrm{B}}T\,\tau_{\min}F_{\mu\nu}$ is a positive semi-definite matrix.

This bound depends on the fastest channel timescale $\tau_{\min}$, and is only tight when the channel has a single timescale. Systems with many degrees of freedom can often have extremely fast timescales. However, in practice, signals do not couple to arbitrarily fast time-scales. In this case, $\tau_{\min}$ should thought of as the fastest channel time-scale τ_a that is appreciably driven by the signal (i.e. $\boldsymbol{\eta}^a \cdot \delta\boldsymbol{\phi}_\mu$ is non-negligible). Indeed, we will see below two examples where this timescale is much slower than the channel's fastest timescale.

A power–precision–speed inequality. The previous section revealed a simple inequality relating friction to information. Here we build on this inequality to derive a general relation between power, precision and signal speed. In particular, the Fisher information in (7) is a Riemannianian metric on the manifold of equilibrium channel state distributions, describing the information geometry [37] of this manifold. This metric measures the sensitivity of the channel distribution $\boldsymbol{\pi}(\boldsymbol{\lambda})$ to changes in the signal $\boldsymbol{\lambda}$. Intuitively, the higher this sensitivity, the more precisely one can estimate the signal $\boldsymbol{\lambda}$ from an observation of the stochastic microstate i of the channel.

This intuition is captured by the Cramer-Rao theorem. For simplicity, we focus below on the case of a one dimensional signal λ. We discuss analogous results for multidimensional signals in [38, §4]. Consider a single observation of the channel microstate i, drawn from the equilibrium channel distribution $\boldsymbol{\pi}(\lambda)$. Further consider an unbiased signal estimator $\hat{\lambda}(i)$, i.e. a function of the stochastic channel microstate i whose mean over observations is equal to the true signal λ. The precision of this estimator is defined as the reciprocal of the variance of $\hat{\lambda}$ over the channel stochasticity: $\mathrm{Prec}(\hat{\lambda}) = \frac{1}{\mathrm{Var}(\hat{\lambda})}$. The Cramer-Rao [39, 40] bound states that estimator precision is bounded by Fisher information,

$$\mathrm{Prec}(\hat{\lambda}) \leq F, \tag{8}$$

for *any* unbiased estimator $\hat{\lambda}$ (here we have dropped the indices in $F_{\mu\nu}$ for the special case of scalar signals).

A potential complication in the application of the classical Cramer-Rao bound for static signals λ, to our case of time varying signals $\lambda(t)$ is that the channel microstate i is drawn from a non-equilibrium distribution $\mathbf{p}(t)$, not the equilibrium distribution $\boldsymbol{\pi}(\lambda(t))$. However, in the slow signal limit, which is related to an expansion in the temporal derivatives of $\lambda(t)$, we can show that the discrepancy between these two distributions can be neglected, as any such discrepancy only corrects higher order terms in this expansion (see [38, §6, Eq. 17]). Thus to the leading order in the slow signal expansion, in which the relation (3) between the friction tensor and power dissipation holds, we can also replace the Fisher information of $\mathbf{p}(t)$ with that of $\boldsymbol{\pi}(\lambda(t))$.

Now, with the careful analysis of the validity of the slow signal limit in hand, by simply combining the relation (3) between power, friction, and signal speed, the inequality (6) relating friction to information, and the inequality (8), relating information to precision, we derive our central result relating power, precision and speed:

$$\mathrm{Prec}(\hat{\lambda})\, V \leq \frac{\mathcal{P}_{\mathrm{ex}}}{k_{\mathrm{B}}T\,\tau_{\min}}, \tag{9}$$

where $V = \dot{\lambda}^2$ is the squared signal velocity. Thus the product of two desirable quantities, the communication precision and signal speed, is upper bounded by an undesirable quantity, the excess power dissipation. This inequality uncovers the fundamental result that any attempt to communicate faster signals at fixed precision, or with higher precision at fixed signal speed, necessarily requires greater power dissipation. Moreover, this relation applies universally to *arbitrary* physical channels.

Saturating (9) requires finding a statistically efficient unbiased estimator $\hat{\lambda}$ that saturates the Cramér-Rao bound (8). For the exponential family distributions that occur in statistical mechanics, we show how to construct such estimators for coordinates "dual" to $\boldsymbol{\lambda}$ (see [37] and [38, §7, Eqs. 18&19]).

Example channels. We now illustrate the general relations derived above in specific examples. Here we only summarize the results. Further details can be found in [38, §8], as well as more examples involving multidimensional signals and channels violating detailed balance.

Heavily damped harmonic oscillator. Consider a heavily damped particle in a viscous medium moving in a quadratic potential, where the external signal $\lambda(t)$ controls the position of the potential's minimum. The particle position x obeys a Langevin equation,

$$\zeta\dot{x} = -\kappa(x - \lambda(t)) + \sqrt{2\zeta k_{\mathrm{B}}T}\,\xi(t), \tag{10}$$

where ζ is the drag coefficient, κ is the potential's spring constant and $\xi(t)$ is zero mean Gaussian white noise with $\langle\xi(t)\xi(t')\rangle = \delta(t - t')$, reflecting fluctuations due to a thermal bath. The distribution of $x(t)$ given any signal history $\lambda(t)$ is Gaussian with moments

$$\langle x(t)\rangle = \int_0^\infty \frac{\mathrm{d}t'}{\tau}\,\mathrm{e}^{-t'/\tau}\lambda(t-t') = \sum_{n=0}^{\infty}\left[-\tau\frac{\mathrm{d}}{\mathrm{d}t}\right]^n \lambda(t),$$
$$\langle \delta x(t)\delta x(t')\rangle = \sigma^2 \mathrm{e}^{-\left|t-t'\right|/\tau}, \tag{11}$$

where $\tau = \frac{\zeta}{\kappa}$ is the channel's relaxation timescale and $\sigma^2 = \frac{k_\mathrm{B}T}{\kappa}$ is the variance of the channel's equilibrium position fluctuations. In the slow signal limit, where $\lambda(t)$ varies over timescales larger than τ, the channel's mean position approximately tracks the signal: $\langle x(t)\rangle \approx \lambda(t)$. More precisely, (11) reveals that this slow signal limit is equivalent to neglecting higher order terms in a temporal derivative expansion. This truncation is a good approximation when the temporal signal has negligible power at frequencies larger than $\frac{1}{\tau}$ [38, §3]. In this limit, a good estimator for the signal based on the channel state x is simply $\hat{\lambda} = x$, and its precision is $\mathrm{Prec}(\hat{\lambda}) = \frac{1}{\sigma^2}$. In the same slow limit, we compute power dissipation [38, §8.1]:

$$\mathcal{P}_\mathrm{ex} = \kappa\dot{\lambda}(t)\int_0^\infty \mathrm{d}t'\,\mathrm{e}^{-t'/\tau}\dot{\lambda}(t-t') \approx \zeta\dot{\lambda}(t)^2. \tag{12}$$

Intuitively, the drag force is given by $-\zeta\dot{x}$, so the rate of doing work against it is $\zeta\dot{x}^2$, and in the slow signal limit, $x(t)\approx\lambda(t)$. Finally, using the Fokker-Planck description of the channel [38, §8.1], we find the intrinsic channel eigenmode timescales are $\tau_n = \tau/n$, for $n = 1, 2, \ldots\infty$. However, λ only couples to the $n = 1$ mode, so $\tau_\mathrm{min} = \tau$.

Combining all these results yields,

$$\frac{\mathrm{Prec}(\hat{\lambda})\,V}{\mathcal{P}_\mathrm{ex}} = \frac{[\sigma^{-2}][\dot{\lambda}^2]}{[\zeta\dot{\lambda}^2]} = \frac{1}{k_\mathrm{B}T\,\tau_\mathrm{min}}, \tag{13}$$

revealing that the damped harmonic oscillator channel saturates the general bound (9). Note that the precision, and the Fisher information, are given by $\frac{1}{\sigma^2} = \frac{\kappa}{k_\mathrm{B}T}$. This means that increasing the spring constant, κ, increases precision, as it forces x to track λ more closely. However, it also speeds up the system, i.e. it decreases $\tau = \frac{\zeta}{\kappa}$, and has no net effect on the power consumption, $\mathcal{P}_\mathrm{ex} = \zeta\dot{\lambda}^2$. In contrast, increasing the drag coefficient, ζ, will increase power consumption and slow down the system, as expected, but has no effect on precision. In practice, it is not possible to make κ arbitrarily large, or ζ arbitrarily small. This will limit how small one could make τ.

Ising ring: Consider a one dimensional Ising ring with periodic boundary conditions, i.e. N spins, $\sigma_n = \pm 1$, with $\sigma_0 = \sigma_N$, all receiving a signal h, with Hamiltonian

$$H = -h\sum_n \sigma_n - J\sum_n \sigma_n\sigma_{n+1}. \tag{14}$$

We perform all computations as the signal h passes through $h = 0$ at finite velocity $\dot{h}$, and we assume Glauber dynamics [41] (see [38, §8.2]) for the spins. This channel can model cooperativity between cell surface chemical receptors [42], with $\sigma_n = \pm 1$ representing the active and inactive receptor states, the field h determined by the ligand concentration, and J controlling receptor cooperativity. Equivalently, this channel could model cooperatively in the opening and closing of voltage gated ion channels, with h reflecting time-varying voltage and the spins reflecting channel configurations.

First, although the Ising ring Glauber dynamics has a spectrum of eigenmode timescales, with the shortest being $\frac{1}{\alpha N}$, where α is the overall rate of the dynamics, the signal h couples only to a mode with a single timescale (see [38, §8.2]), yielding

$$\tau_\mathrm{min} = \frac{\mathrm{e}^{2\beta J}\cosh 2\beta J}{\alpha}. \tag{15}$$

This quantity increases with J, due to critical slowing down [43]. The slow signal limit is valid when the timescale τ_h over which h varies is much larger than τ_min. For a fixed τ_h, this limit yields an upper limit on values of J that we can analyze, which is roughly $J \ll k_\mathrm{B}T\ln\alpha\,\tau_h$.

The Fisher information is given by

$$F = N\beta^2\mathrm{e}^{2\beta J}, \tag{16}$$

which also increases with J. In essence, increasing J has two opposing effects on how well the spin statistics transmits the signal h. First the gain of the mean spin response to h (i.e. magnetic susceptibility) increases, improving coding. Second, the variance of spin response also increases with J, impairing coding. The former effect dominates over the latter, leading to increased information with cooperativity. Moreover, in [38, §8.2] we show how to construct an efficient unbiased estimator for the dual coordinate to the signal h (see [38, §7, Eqs. 18&19]).

The power dissipated when h is varied is given by

$$\mathcal{P}_\mathrm{ex} = \frac{N\beta\,\mathrm{e}^{4\beta J}\cosh 2\beta J}{\alpha}\,\dot{h}^2. \tag{17}$$

This also increases with J, partly due to increased response gain to h, and partly due to critical slowing down. Now, combining (15), (16), and (17), we find

$$\frac{F\,V}{\mathcal{P}_\mathrm{ex}} = \frac{1}{(k_\mathrm{B}T)\,\tau_\mathrm{min}}, \tag{18}$$

where $V = \dot{h}^2$ is the squared signal velocity. This implies the general bound (9) would be saturated for the Ising ring if the Cramer-Rao bound (8) could be saturated.

We note that while increasing J increases Fisher information, the dissipated power increases even faster, yielding diminishing returns in terms of Fisher information per watt. This is analogous to the diminishing returns exemplified by the concavity of the capacity-cost curve in information theory [4]. Also, in [42], the same system was analyzed in a different setting. There the signal was static while the channel was observed for an extended period, whereas here the signal is changing and the channel is observed instantaneously. There, increasing receptor cooperativity J reduced performance, since critical slowing yields fewer independent signal observations. Similarly, here we see that cooperativity unhelpfully tightens the tradeoff between power, precision and speed, as it slows down the system, decreasing the right-hand-side of (9), as verified in [38, §8.2].

Discussion. In summary, by deriving general relations between friction and information, we have shown that the product of signal speed and channel precision cannot exceed power dissipation for an extremely general class of physical communication channels. Intuitively, this three-way tradeoff arises because any increase in speed at fixed precision requires the channel state distribution to change more rapidly, leading to increased power dissipation. Similarly any increase in precision at fixed speed requires high signal sensitivity, or a larger signal dependent change in the channel equilibrium state distribution as measured by the Fisher information metric, which again leads to greater power dissipation.

Our newly discovered three-way tradeoff motivates new experiments to assess exactly how close biological systems come to simultaneously optimizing power, precision and speed. Indeed any experiment that measures only two of these three quantities fundamentally cannot assess how close evolution pushes biology to the limits set by physics in general information processing tasks. Moreover, our work opens the door to intriguing theoretical extensions. Here, we focused on tradeoffs in estimating the current value of a slowly changing signal from an instantaneous observation of a physical channel. Alternatively, we could consider estimating either the temporal history of a signal from the instantaneous channel state [44, 45], or estimating a static signal given an extended time series of channel states. The former would involve Fisher information metrics of channel states over signal trajectories, while the later would involve the Fisher information of probability distributions over channel state trajectories. It would be interesting to explore universal three-way tradeoffs between power, precision and speed in these more general dynamical scenarios. We hope that the essential ideas underlying our mathematical derivation of universal tradeoffs between power precision and speed will be of benefit in understanding even more general scenarios of communication and computation across both biology and engineering.

Acknowledgements

We thank Madhu Advani, Gavin Crooks, Dibyendu Mandal and the participants of the Lineq reading group at UC Berkeley for useful discussions. We thank Genentech, the Office of Naval Research, the Burroughs-Wellcome Fund and the Alfred P. Sloan, James S. McDonnell, Simons and McKnight Foundations for funding.

* Electronic address: sulahiri@stanford.edu
† Electronic address: jaschasd@google.com
‡ Electronic address: sganguli@stanford.edu

[1] S. Thorpe, D. Fize, and C. Marlot, Nature **381**, 520 (1996).
[2] P. Sterling and S. Laughlin, *Principles of neural design* (MIT Press, 2015).
[3] C. E. Shannon, Bell System Technical Journal **27**, 379 (1948).
[4] T. M. Cover and J. A. Thomas, *Elements of information theory* (John Wiley & Sons, 2012).
[5] M. Gastpar, B. Rimoldi, and M. Vetterli, IEEE Transactions on Information Theory **49**, 1147 (2003).
[6] Y. Polyanskiy, H. V. Poor, and S. Verdu, IEEE Transactions on Information Theory **56**, 2307 (2010).
[7] T. Schmiedl and U. Seifert, Phys. Rev. Lett. **98**, 108301 (2007).
[8] H. Then and A. Engel, Phys. Rev. E **77**, 041105 (2008).
[9] R. Landauer, IBM Journal of Research and Development **5**, 183 (1961).
[10] C. H. Bennett, IBM Journal of Research and Development **17**, 525 (1973).
[11] C. H. Bennett, International Journal of Theoretical Physics **21**, 905 (1982).
[12] T. Toffoli, in *Automata, Languages and Programming*, edited by J. de Bakker and J. van Leeuwen (Springer Berlin Heidelberg, 1980), vol. 85 of *Lecture Notes in Computer Science*, pp. 632–644.
[13] J. J. Hopfield, Proc. Nat. Acad. Sci. USA **71**, 4135 (1974).
[14] M. A. Savageau and R. R. Freter, Proc. Nat. Acad. Sci. USA **76**, 4507 (1979).
[15] M. A. Savageau and R. R. Freter, Biochemistry **18**, 3486 (1979).
[16] R. R. Freter and M. A. Savageau, Journal of Theoretical Biology **85**, 99 (1980).
[17] M. Ehrenberg and C. Blomberg, Biophysical journal **31**, 333 (1980).
[18] C. Blomberg and M. Ehrenberg, Journal of Theoretical Biology **88**, 631 (1981).
[19] M. A. Savageau and D. S. Lapointe, Journal of Theoretical Biology **93**, 157 (1981).
[20] H. Qian, Journal of molecular biology **362**, 387 (2006).
[21] A. Murugan, D. A. Huse, and S. Leibler, Phy. Rev. X **4**, 021016 (2014).
[22] A. Murugan, D. A. Huse, and S. Leibler, Proc. Nat. Acad. Sci. USA **109**, 12034 (2012).
[23] R. G. Endres and N. S. Wingreen, Phys. Rev. Lett. **103**, 158101 (2009).
[24] P. Mehta and D. J. Schwab, Proc. Nat. Acad. Sci. USA **109**, 17978 (2012).
[25] A. H. Lang, C. K. Fisher, T. Mora, and P. Mehta, Phys. Rev. Lett. **113**, 148103 (2014).
[26] A. C. Barato, D. Hartich, and U. Seifert, New Journal of Physics **16** (2014).
[27] C. C. Govern and P. R. ten Wolde, Phys. Rev. Lett. **113**, 258102 (2014).
[28] C. C. Govern and P. R. Ten Wolde, Proc. Nat. Acad. Sci. USA **111**, 17486 (2014).
[29] P. Sartori, L. Granger, C. Lee, and J. Horowitz, PLoS computational biology **10**, e1003974 (2014).

[30] G. Lan, P. Sartori, S. Neumann, V. Sourjik, and Y. Tu, Nature physics **8**, 422 (2012).
[31] W. J. Dally and B. P. Towles, *Principles and Practices of Interconnection Networks* (Elsevier, 2004).
[32] D. A. Sivak and G. E. Crooks, Phys. Rev. Lett. **108**, 190602 (2012).
[33] P. R. Zulkowski, D. A. Sivak, G. E. Crooks, and M. R. Deweese, Phys. Rev. E **86**, 041148 (2012).
[34] P. R. Zulkowski, D. A. Sivak, and M. R. DeWeese, PloS one **8**, e82754 (2013).
[35] P. R. Zulkowski and M. R. DeWeese, Phys. Rev. E **89**, 052140 (2014).
[36] D. Mandal and C. Jarzynski, arXiv (2015), 1507.06269.
[37] S.-i. Amari and H. Nagaoka, *Methods of information geometry*, vol. 191 (American Mathematical Soc., 2007).
[38] Supplementary material.
[39] H. Cramér, *Mathematical Methods of Statistics* (Princeton University Press, 1945).
[40] C. Radhakrishna Rao, Bull. Calcutta Math. Soc. **37**, 81 (1945).
[41] R. J. Glauber, J. Math. Phys. **4**, 294 (1963).
[42] M. Skoge, Y. Meir, and N. S. Wingreen, Phys. Rev. Lett. **107**, 1 (2011).
[43] P. C. Hohenberg and B. I. Halperin, Rev. Mod. Phys. **49**, 435 (1977).
[44] S. Ganguli, D. Huh, and H. Sompolinsky, Proc. Nat. Acad. Sci. USA **105**, 18970 (2008).
[45] S. Ganguli and H. Sompolinsky, in *Advances in Neural Information Processing Systems* (2010), pp. 667–675.