Rapid Communications

Modeling evolution of crosstalk in noisy signal transduction networks

Ammar Tareen,[1] Ned S. Wingreen,[2,*] and Ranjan Mukhopadhyay[1,†]
[1]*Department of Physics, Clark University, Worcester, Massachusetts 01610, USA*
[2]*Lewis-Sigler Institute for Integrative Genomics, Carl Icahn Laboratory, Washington Road, Princeton, New Jersey 08544, USA*

(Received 30 June 2017; published 8 February 2018)

Signal transduction networks can form highly interconnected systems within cells due to crosstalk between constituent pathways. To better understand the evolutionary design principles underlying such networks, we study the evolution of crosstalk for two parallel signaling pathways that arise via gene duplication. We use a sequence-based evolutionary algorithm and evolve the network based on two physically motivated fitness functions related to information transmission. We find that one fitness function leads to a high degree of crosstalk while the other leads to pathway specificity. Our results offer insights on the relationship between network architecture and information transmission for noisy biomolecular networks.

DOI: 10.1103/PhysRevE.97.020402

An essential characteristic of living cells is their ability to regulate their own behavior, based on environmental signals, to ensure survival, growth, and proliferation [1]. Reliable transmission of information about the environment along cellular signaling pathways is crucial for accurate regulation. Malfunctioning of signaling pathways underlies many pathological conditions in higher organisms, including cancers and Alzheimer's disease [2–4]. However, signaling pathways are often highly interconnected, creating signal transduction networks composed of multiple pathways [5,6]. Crosstalk between pathways accounts for many of the complex behaviors exhibited by signaling networks [7,8]. How did such complex, interconnected networks evolve and what constraints did the dynamics of evolution place on their architecture? Does crosstalk between pathways necessarily lead to reduction in the amount of information that can be reliably transmitted? This Rapid Communication describes a theoretical study of the evolution of crosstalk between signaling pathways with the aim of addressing these and related questions.

In order to understand the effect of crosstalk on the transmission of information, we draw from Shannon's work on communication theory [9] and quantify information transmission along noisy signaling pathways in terms of the mutual information (MI) between the input and output. However, rigorously computing the mutual information for noisy biochemical channels is challenging and thus often noise is assumed to be additive and Gaussian [10,11]. In this Rapid Communication, we model noisy biochemical channels using chemical stochastic Langevin equations [12], where the strength of noise nontrivially depends on the input. To this end, we introduce a method for computing mutual information in the context of such channels. Surprisingly, we find that crosstalk may not lead to a reduction in total information transmitted and that optimal information transmission need not correspond to zero crosstalk. This contrasts with the case of Gaussian channels with constant additive noise where crosstalk necessarily leads to a reduction in information transmission [13].

Modeling the evolution of biomolecular networks poses an additional challenge because evolutionary processes are governed by changes at the genotypic level, whereas selection occurs at the phenotypic level [14] and the mapping between genotype and phenotype is generally poorly understood. Currently, much of the theory related to evolution of signal transduction networks focuses on changes at the phenotypic level (e.g., direct changes to protein interactions) [15,16]. In this Rapid Communication we adapt a sequence-based evolutionary model due to Ali *et al.* [17] that allows us to map from sequence space (genotype) to rate constant space (phenotype). In biological systems, new signaling pathways can enter the genome via gene duplication and subsequent divergence [18]. Therefore, for our evolutionary study, we consider two parallel pathways that arise via gene duplication but are then allowed to diverge. We evolve our network using two biologically motivated fitness functions related to the transmission of information. For the first fitness function, we focus on a system which may have evolved to transmit the total information content along the signaling network; the fitness for this scenario is determined by the total mutual information between inputs and outputs, denoted by $\mathrm{MI}_{\mathrm{total}}$. For the second fitness function, we consider a system where inputs transmitted through their cognate signaling pathways lead to distinct responses. A natural choice of fitness function for this scenario is the sum of the mutual information of individual pathways, denoted by $\mathrm{MI}_{\mathrm{sum}}$. We find that the two fitness functions lead to very different evolutionary outcomes. In particular, evolution retains a high degree of crosstalk for the case of $\mathrm{MI}_{\mathrm{total}}$ while leading to high specificity for $\mathrm{MI}_{\mathrm{sum}}$.

In cells, there exist examples of both high degrees of crosstalk and high degrees of specificity. As an example of crosstalk, studies have shown interactions between the IGF-I and the TGF-β pathways, where in the Hep3B human hepatoma cell line, IGF-I and insulin were each shown to block TGF-β induced apoptosis, via a PI3-kinase/Akt dependent pathway [19]. In another example of crosstalk,

*wingreen@princeton.edu
†ranjan@clarku.edu

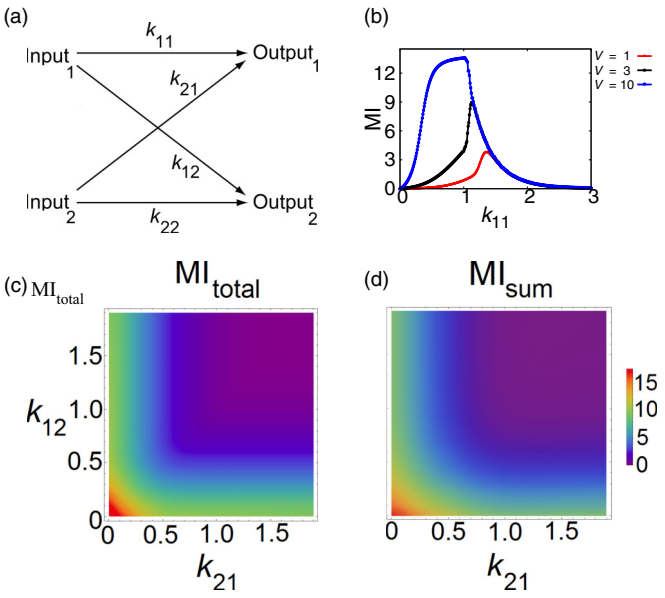

FIG. 1. (a) Signaling network showing direct and crosstalk pathways along with their associated reaction rate constants. (b) Mutual information versus k_{11} shown for several values of system volume V. The input probability distribution, $P(I)$, is chosen to be a Gaussian (mean $\mu = 0.5$ and standard deviation $\sigma = 0.1$). (c) $\mathrm{MI}_{\mathrm{total}}$ and (d) $\mathrm{MI}_{\mathrm{sum}}$ versus crosstalk rate constants k_{21} and k_{12}, with $k_{11} = k_{22} = 1$, for $V = 3$.

cyclic AMP helps regulate cell proliferation by interacting with the mitogen-activated protein kinase pathway [20]. More examples can be found in [6,21–23]. On the other hand, two-component signaling systems, which form the dominant signaling modality in bacteria, exhibit a high degree of pathway isolation and therefore a high degree of specificity [8]. Examples of specificity in signaling are found in [24–29]. Indeed, undesirable crosstalk underlies many pathological conditions in higher organisms [2–4].

In our model of a signaling pathway, we assume two layers of proteins that represent an input-output process. The first layer corresponds to a set of proteins (e.g., cell surface receptors or protein kinases) that become activated by an extracellular signal (e.g., a ligand); the activated fraction of these proteins represents the input. These input proteins, in their active form, can activate a second layer of proteins whose activated fraction represents the output. To study information transmission in this system [see Fig. 1(a)], we employ the chemical Langevin equation, which approximately models the stochastic dynamical behavior of a well-stirred mixture of molecular species that chemically interact:

$$dO_j^*/dt = A_j + B_j\xi_j(t), \tag{1}$$

where functions A and B are deterministic and stochastic parts of the Langevin equation, respectively, defined as

$$A_j = \sum_i k_{ij} I_i O_j - \alpha O_j^*,$$

$$B_j = \left[\left(\sum_i k_{ij} I_i O_j + \alpha O_j^*\right)\Big/ V\right]^{1/2}. \tag{2}$$

I_i is the strength of input i, O_j^* is the concentration of activated output protein j (aka the output), and O_j is the inactive concentration, with the total concentration of output protein held fixed, i.e., $O_{\mathrm{tot}} = O_j + O_j^*$. We assume a background deactivation rate of $\alpha = 1$ and $O_{\mathrm{tot}} = 1$, which define our units of time and volume. V represents the volume of the system, and controls the level of noise. The factors k_{ij} are reaction rate constants. ξ_j is a stochastic variable which represents Gaussian white noise with zero mean, $\langle\xi_j(t)\rangle = 0$, and is temporally uncorrelated $\langle\xi_i(t)\xi_j(t')\rangle = \delta(t-t')\delta_{ij}$.

For our evolutionary scheme, we adopt the model by Ali *et al.* [17] where the rate constants are determined by interactions between protein interfaces. We assume that input proteins possess an out-face and output proteins possess an in-face which form a pair of interaction interfaces; as in [17], we associate a binary sequence, $\vec{\sigma}_{\mathrm{in/out}}$, of hydrophobic residues (1's) and hydrophilic residues (0's) to each interface. The interaction strength between a protein (denoted by index i) and its target (denoted by index j) is determined by the interaction energy $E_{ij} = \epsilon\vec{\sigma}^{\,i}_{\mathrm{out}} \cdot \vec{\sigma}^{\,j}_{\mathrm{in}}$ between the out-face of the input protein and the in-face of the output protein. ϵ represents the effective interaction energy between two hydrophobic residues. (All energies are expressed in units of the thermal energy k_BT.) The reaction rate is

$$k_{ij} = k_0/\{1 + \exp[-(E_{ij} - E_0)]\}, \tag{3}$$

where E_0 plays the role of a threshold energy, e.g., accounting for the loss of entropy due to binding. In our calculations we varied k_0 between 1 and 20, ϵ between 0.2 and 0.6, and V between 1 and 100. We set $E_0 = 5$, and we took the length of each sequence representing an interface to be $M = 25$. This choice of parameters allowed us to vary the resulting rate constants k_{ij} over three orders of magnitude. Additionally, our range of rate constants contain the biologically relevant range for functional signaling pathways, as values of $k_{ij} > 3$ can cause the network to become saturated, resulting in extremely low values of mutual information between input and output [see Figs. 1(b)–1(d)], whereas for $k_{ij} \ll 1$, very few output proteins can become activated, leading once again to low mutual information.

For our evolutionary scheme, we assume a population sufficiently small that each new mutation is either fixed or entirely lost [30].[1] We consider only point mutations, namely, replacing a randomly chosen hydrophobic residue (1) in the in- or out-face of one protein by a hydrophilic residue (0), or vice versa. In this study, mutations are accepted if and only if they produce a fitness that is greater than or equal to the current fitness. In this work, we study two fitness functions based on the mutual information between the inputs and outputs of our

[1]The time it takes for a mutation to become fixed in a population increases with population size N, whereas the time between successive mutations goes as $1/N$ (see, for example, [31]). In the small population limit, mutations fix much more rapidly than they occur. This is the limit we have assumed for this Rapid Communication, so that an accepted mutation in our model corresponds to a mutation that gets fixed in the population.

system, with MI defined as [9]

$$\mathrm{MI}(I;O^*) = \iint P(I,O^*)\log\frac{P(I,O^*)}{P(I)P(O^*)}dI\,dO^*, \quad (4)$$

where P always represents a probability distribution function. The two fitness functions can be expressed as $\mathrm{MI}_{\mathrm{total}} = \mathrm{MI}(I_1,I_2;O_1^*,O_2^*)$ and $\mathrm{MI}_{\mathrm{sum}} = \mathrm{MI}(I_1;O_1^*) + \mathrm{MI}(I_2;O_2^*)$. For calculating MI, we chose the input probability distribution $P(I)$ to be Gaussian and used the Fokker-Planck (FP) equation corresponding to our Langevin equation [Eq. (1)] to calculate the conditional probability distributions of the output given inputs.

We first consider the simpler case of a one-input, one-output system to develop tools to address multiple input-output systems with crosstalk. For a one-input, one-output system, the resulting FP equation (in the Itô formulation [32]) is

$$\frac{\partial P}{\partial t} = -\frac{\partial}{\partial O^*}\{AP\} + \frac{1}{2}\frac{\partial^2}{\partial O^{*2}}\{B^2P\}, \quad (5)$$

where $P(O^*|I)$ represents the conditional output probability distribution given input. Note that Eq. (5) has the form of a continuity equation for probability

$$\frac{\partial P(O^*,t)}{\partial t} + \frac{\partial J(O^*,t)}{\partial O^*} = 0, \quad (6)$$

where $J = \frac{\partial}{\partial O^*}[AP - \frac{1}{2}(B^2P)]$ can be viewed as a probability current. The steady-state solution of the FP equation corresponds to a constant value of J. Imposing the boundary conditions $J = 0$ at $O^* = 0$ and at $O^* = 1$ then implies that $J = 0$ everywhere. The solution of the steady-state FP equation for zero-probability-current boundary conditions can be written as [13]

$$P(O^*|I,k_{11}) = Ne^{-2VO^*(Ik_{11}+\alpha)/\alpha - Ik_{11}} \times\left[1+\frac{(\alpha - Ik_{11})O^*)}{Ik_{11}O_{\mathrm{tot}}}\right]^{[4Ik_{11}O_{\mathrm{tot}}\alpha V/(\alpha - Ik_{11})^2]-1}, \quad (7)$$

where N is a normalization constant. Note that this conditional output probability distribution is peaked for $V = 2$ or higher. Additionally, it might appear that the right-hand side of Eq. (7) approaches ∞ as $\alpha \to Ik_{11}$; however, setting $\delta = \alpha - Ik_{11}$ and Taylor expanding around $\delta = 0$, we find that the divergent terms cancel [13]. We determine the output probability $P(O^*)$ by numerically integrating the conditional output probability over the input distribution, and thereby obtain the mutual information as a function of k_{11}, as shown in Fig. 1(b). The mutual information is nearly zero both at very small values of k_{11} because of low activation and at very large values of k_{11} because of saturated output.

We now extend the one-input, one-output system to two inputs and two outputs, and allow for crosstalk. The resulting FP equation for the joint probability distribution $P(O_1^*,O_2^*,t)$ is [33,34]

$$\frac{\partial P}{\partial t} = -\sum_i \frac{\partial}{\partial O_i^*}\{A_iP\} + \frac{1}{2}\sum_{ij}\frac{\partial^2}{\partial O_i^*\partial O_j^*}\{[B_iB_j]P\}. \quad (8)$$

The steady-state solution that satisfies the zero-probability-current boundary conditions for Eq. (8) is [13]

$$P(O_i^*|I_1,I_2) = Ne^{\{-2VO_i^*[(\alpha R_i^*+1)/(\alpha R_i^*-1)]\}} \times\left[1+\frac{(\alpha R_i^* - 1)O_i^*}{R_iR_i^*}\right]^{[4VR_iR_i^{*2}\alpha/(\alpha R_i^*-1)^2]-1}. \quad (9)$$

For notational convenience we have introduced modified rates $R_j \equiv O_{\mathrm{tot},j}(\sum_i k_{ij}I_i)$ and $R_j^* \equiv O_j^*(\sum_i k_{ij}I_i)$. Having obtained the conditional probabilities, we now obtain the two fitness functions numerically. For $V = 3$, if we set $k_{11} = k_{22} = 1$ [i.e., corresponding to values of these rate constants close to the optimum of MI for a single pathway, as seen in Fig. 1(b)], then we can depict the density plots of fitness versus crosstalk, as in Figs. 1(c) and 1(d), and observe that both fitness landscapes look similar and both have a fitness maximum at zero crosstalk (larger volumes yield qualitatively similar landscapes, see [13] for a calculation with $V = 10$). However, Figs. 1(c) and 1(d) provide only a slice through parameter space. How might an evolving system explore the full space? To answer this question we take an evolutionary approach.

We implement our evolutionary scheme as described earlier with the initial state of the system corresponding to duplicated pathways, where all the rate constants k_{ij} are equal (e.g., for all strings initialized to zero and $\epsilon = 0.2$). Figure 2(a) shows some sample runs of the evolutionary algorithm for a few different choices of initial conditions; each solid curve represents the average fitness for 100 runs for a specific set of initial strings, while the shaded regions indicate the 25th–75th fitness percentiles at that particular number of accepted mutations over all trajectories. Figure 2(a) shows results for $\mathrm{MI}_{\mathrm{total}}$; however, the results for $\mathrm{MI}_{\mathrm{sum}}$ are qualitatively similar. We can see that the final values of the rate constants do not depend critically on our choice of initial strings.

Surprisingly, evolving $\mathrm{MI}_{\mathrm{total}}$ leaves the optimized network with a high degree of crosstalk, contrary to our expectations based on Fig. 1. For example, for $\epsilon = 0.2$, if we start with low values of all k_{ij}, we typically find that all the rate constants increase simultaneously, as shown in Fig. 2(b), implying high crosstalk. Strikingly, for larger ϵ, the majority of runs exhibit bifurcations in rate constants, but still leave the optimized network with a high degree of crosstalk [see Fig. 2(c)]. In a typical bifurcation, k_{11} and k_{12} might dominate while k_{21} and k_{22} are suppressed, whereas k_{21} and k_{22} might dominate in a different run. These bifurcations yield examples of signal "fan-out" (single input, multiple outputs) and signal "fan-in" (multiple inputs, single output), found in biological systems [35]. Figure 2(d) shows a probability distribution of rate constants after rate constants have stopped changing under $\mathrm{MI}_{\mathrm{total}}$ evolution; the peaks of the histogram occur at similarly high values of the crosstalk and direct rate constants, implying a high degree of crosstalk as an evolutionary outcome for $\mathrm{MI}_{\mathrm{total}}$.

On the other hand, evolution under the fitness function $\mathrm{MI}_{\mathrm{sum}}$ leads to low crosstalk and thus isolated pathways. Figure 2(e) shows a typical run of greedy evolution under $\mathrm{MI}_{\mathrm{sum}}$. Note that in this typical run, the direct rate constant values grow [e.g., $k_{11},k_{22} \sim 1$ in the evolved network, corresponding

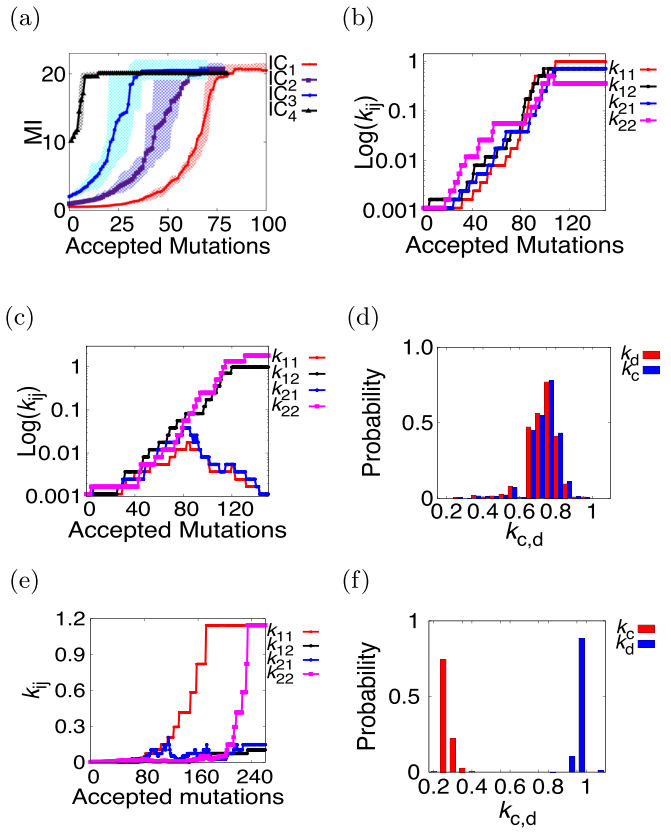

FIG. 2. (a) Fitness versus accepted mutations with four different initial conditions (labeled "IC" on the legend) [13]; solid curves represent fitness averaged over 100 simulations while shaded curves represent 25th–75th percentiles from each of the simulations at every accepted mutation. (b) Evolution under $\mathrm{MI}_{\mathrm{total}}$; $\log(k_{ij})$ versus accepted mutations. (c) Bifurcations in pairs of rate constants for $\epsilon = 0.6$. (d) Probability distribution of rate constants showing high degree of crosstalk for $\epsilon = 0.2$. k_d represents the direct rate constants; k_c represents the crosstalk rate constants. (e) Evolution under $\mathrm{MI}_{\mathrm{sum}}$ and the probability distribution of k_{ij}; rate constants versus accepted mutations. (f) Probability distribution of rate of constants showing suppression of crosstalk for $\epsilon = 0.2$; constructed from 10 000 simulations. $k_0 = 20$, $E_0 = 5$, $V = 3$.

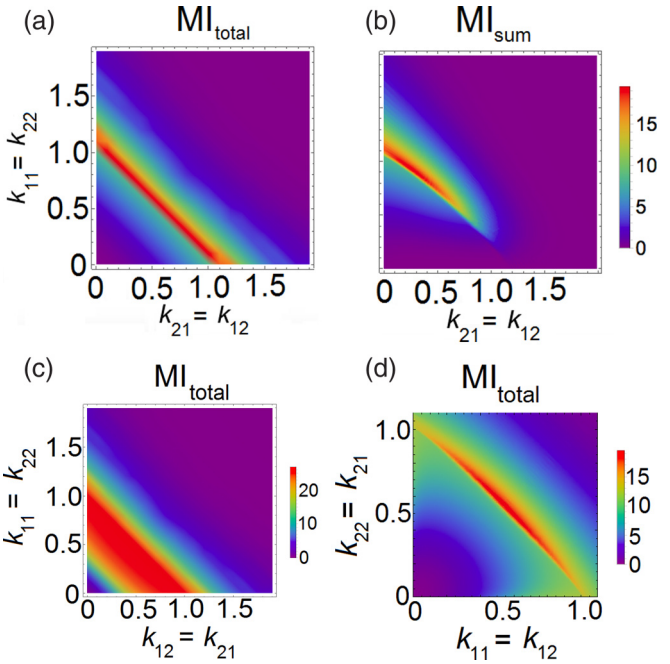

FIG. 3. Fitness landscapes plotted for $k_{21} = k_{12}$ and $k_{11} = k_{22}$ for (a) $\mathrm{MI}_{\mathrm{total}}$ and (b) $\mathrm{MI}_{\mathrm{sum}}$. $\mathrm{MI}_{\mathrm{total}}$ does not have a single global maximum associated with zero crosstalk, whereas $\mathrm{MI}_{\mathrm{sum}}$ does. (c) $\mathrm{MI}_{\mathrm{total}}$ for $k_{21} = k_{12}$ and $k_{11} = k_{22}$ for $V = 10$ displays similar behavior qualitatively. (d) $\mathrm{MI}_{\mathrm{total}}$ plotted as a function of rates $k_{11} = k_{12}$ and $k_{22} = k_{21}$ for $V = 3$.

to the optimal values in the single-input single-output case, as in Fig. 1(b)], whereas the crosstalk rate constants stay low (e.g., k_{12},$k_{21} \sim 0.1$). Figure 2(f) shows a histogram exhibiting separation of crosstalk and direct rate constants, with high values of direct rate constants and low values of crosstalk rate constants.

How can we understand this striking difference in evolutionary outcomes for the two fitness functions given that the maximum fitness depicted in Fig. 1 occurred at zero crosstalk for both functions? Although the two landscapes appeared similar, it is important to recall that the phase space of the fitness landscapes is really four dimensional and the landscapes in Fig. 1 correspond to a particular two-dimensional slice. We are then faced with the question of how to construct a lower dimensional slice of the fitness landscapes that could help us understand the difference in evolutionary outcomes. The crucial difference between evolutionary outcomes pertained to the typical ratio between direct and crosstalk rate constants; we therefore want to distinguish between the fitness dependence on the direct rate constants and crosstalk rate constants. Thus, we set $k_{11} = k_{22}$, corresponding to the direct rate constant, and $k_{12} = k_{21}$, corresponding to the crosstalk rate constant, and construct a two-dimensional slice where one axis represents the direct rate constant and the other the crosstalk rate constant. As shown in Fig. 3, the resulting fitness landscapes reveal a striking difference between the two fitness functions. In particular, we note that while $\mathrm{MI}_{\mathrm{sum}}$ is peaked at zero crosstalk (albeit with some spread to finite crosstalk), $\mathrm{MI}_{\mathrm{total}}$ is optimal over an entire band corresponding to a range of direct and crosstalk rate constants; see Figs 3(a) and 3(b) [Fig. 3(c) shows a calculation for $\mathrm{MI}_{\mathrm{total}}$ for $V = 10$, displaying similar qualitative behavior]. The existence of a single peak near zero crosstalk in the fitness landscape of $\mathrm{MI}_{\mathrm{sum}}$ and no such single peak in the landscape of $\mathrm{MI}_{\mathrm{total}}$ helps explain why evolution under $\mathrm{MI}_{\mathrm{sum}}$ leads to low crosstalk while $\mathrm{MI}_{\mathrm{total}}$ can result in high crosstalk. Lastly, to understand the bifurcations shown in Fig. 2(c), we construct another two-dimensional slice of the fitness landscape where we set $k_{11} = k_{12}$ and $k_{22} = k_{21}$ and plot the resulting $\mathrm{MI}_{\mathrm{total}}$ in Fig. 3(d). We note that while the gradient of $\mathrm{MI}_{\mathrm{total}}$ along the diagonal is positive, it can be smaller than the gradient along either axis so that $\mathrm{MI}_{\mathrm{total}}$ could increase in the transverse direction away from the diagonal. For larger ϵ, the change in the rate constants due to a mutation could be larger, which increases the likelihood for the system to take a larger step away from the diagonal and to subsequently move toward either axis, leading to a bifurcation in the magnitudes of the rate constants.

We have adapted a sequence-based protein-protein interaction model to study the evolution of crosstalk in multiple-input, multiple-output signaling networks. Evolution is driven by random mutations in sequence space, whereas selection occurs in the space of phenotypes. Using our evolutionary scheme we have shown that MI_{total} retains a high degree of crosstalk (contrary to our initial expectations based on Fig. 1), whereas MI_{sum} leads to insulated pathways with lowered crosstalk. We related the evolutionary outcomes to the fitness landscapes and showed that while MI_{sum} is optimized for zero crosstalk, MI_{total} is optimal over an entire band corresponding to a range of direct and crosstalk rate constants [see, e.g., Fig. 3(a)]. Our results pertaining to dependence of MI_{total} on crosstalk are unique to biochemical channels where the strength of the noise depends on input; these results are different from Gaussian channels with constant additive noise where crosstalk always leads to reduction in total mutual information [13].

Our work focuses on stochasticity inherent to biochemical reactions (intrinsic noise) rather than variability in cellular states (extrinsic noise) [36]. While generally both intrinsic and extrinsic noise degrade information transmitted through signaling networks, experiments show that signaling networks can mitigate, and potentially eliminate, extrinsic-noise-induced information loss [37]. Furthermore, the impact of extrinsic noise decreases with increasing network complexity [38], which justifies our focus on intrinsic noise (note, however, that owing to its simplicity, our framework can easily be generalized to incorporate extrinsic noise [39]). Our results are also robust to parameter choices. We varied our model parameters k_0, ϵ, and V such that the resulting rate constants k_{ij} spanned three orders of magnitude and observed similar outcomes in our simulations.

In order to appreciate the biological significance of our results, we note that systems for which inputs have to be integrated in order to produce output, such as quorum sensing [40], MI_{total} would be the appropriate fitness function. In such cases, our results indicate that evolution is likely to lead to high degrees of crosstalk or to fanning-in or fanning-out from inputs to outputs. In cases where distinct inputs require distinct responses from the system, we expect MI_{sum} to be the suitable quantity for fitness, in which case our results suggest an evolutionary drive to eliminate crosstalk. An example for the latter is the high osmolarity response in yeast where the pathways respond to the appropriate environmental cues in very distinct and highly precise ways [41].

In this Rapid Communication we assumed completely uncorrelated input distributions for our system; in the future, it would be interesting to explore how correlated inputs might affect evolution of crosstalk. Moreover, we focused here on two-layer signaling processes, but these can readily be extended to include multilayer cascades. Future work could also address the effects of adding feedback, a higher number of pathways, and proteins such as histidine kinases that act both as activators and deactivators [8].

This work was supported in part by the National Science Foundation, Grant No. PHY-1305525, and the National Institutes of Health, Grant No. R01 GM082938.

[1] J. D. Jordan, E. M. Landau, and R. Iyengar, Cell. **103**, 193 (2000).
[2] R. Muller, J. Cancer Res. Clin. Oncol. **130**, 429 (2004).
[3] W. Shi, A. L. Harris, and J. Mammary, Gland Biol. Neoplasia **11**, 41 (2006).
[4] D. Kalaitzidis and T. D. Gilmore, Trends Endocrinol. Metab. **16**, 46 (2005).
[5] L. Attisano and E. Labbe, Cancer Metastasis Rev. **23**, 53 (2004).
[6] M. A. Schwartz and M. H. Ginsberg, Nat. Cell Biol. **4**, E65 (2002).
[7] S. M. Hill, Anat. Rec. **253**, 42 (1998).
[8] M. A. Rowland and E. J. Deeds, Proc. Natl. Acad. Sci. USA **111**, 5550 (2014).
[9] C. E. Shannon, Bell Syst. Tech. J. **27**, 379 (1948).
[10] S. M. Lyons and A. Prasad, PLoS ONE **7**, e34488 (2012).
[11] J. Walters-Williams and Y. Li, Lect. Notes Comput. Sci. **5589**, 389 (2009).
[12] D. T. Gillespie, J. Chem. Phys. **113**, 297 (2000).
[13] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.97.020402 for details regarding solving the Fokker-Planck equation(s), some initial conditions, and validation checks.
[14] S. J. Maynard, *The Theory of Evolution* (Cambridge University Press, Cambridge, England, 1993).
[15] O. S. Soyer and S. Bonhoeffer, Proc. Natl. Acad. Sci. USA **103**, 16337 (2006).
[16] M. Mobashir, B. Schraven, and T. Beyer, PLoS ONE **7**, e50905 (2012).
[17] M. Z. Ali, N. S. Wingreen, and R. Mukhopadhyay, arXiv:1706.08499.
[18] V. Prince and F. Pickett, Nat. Rev. Genet. **3**, 827 (2002).
[19] R. H. Chen, Y. H. Su, R. L. Chuang, and T. Y. Chang, Oncogene **17**, 1959 (1998).
[20] M. Saxena, S. Williams, K. Tasken, and T. Mustelin, Nat. Cell Biol. **1**, 305 (1999).
[21] T. Hunter, Mol. Cell **28**, 730 (2007).
[22] Y. Yan, C. L. Wei, W. R. Zhang, H. P. Cheng, and J. Liu, Acta. Pharmacol. Sin. **27**, 821 (2006).
[23] H. Nishi, E. Demir, and A. R. Panchenko, J. Mol. Biol. **427**, 511 (2015).
[24] M. N. McClean, A. Mody, J. R. Broach, and S. Ramanathan, Nat. Gen. **39**, 409 (2007).
[25] M. Behar, H. G. Dohlman, and T. C. Elston, Proc. Natl. Acad. Sci. USA **104**, 16146 (2007).
[26] L. Bardwell, Biochem. Soc. Trans. **34**, 837 (2006).
[27] W. Kolch, Nat. Rev. Mol. Cell. Biol. **6**, 827 (2005).
[28] L. J. Flatauer, S. F. Zadeh, and L. Bardwell, Mol. Cell. Biol. **25**, 1793 (2005).
[29] N. Dard and M. Peter, BioEssays **28**, 146 (2006).
[30] P. A. P. Moran, Proc. Cambridge Philos. Soc. **54**, 60 (1958).
[31] M. M. Desai, D. S. Fisher, and A. W. Murray, Curr. Biol. **17**, 385 (2007).
[32] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry*, 3rd ed. (North Holland, Amsterdam, 2007).
[33] J. L. Garcia-Palacios, arXiv:cond-mat/0701242.

[34] D. T. Gillespie, Am. J. Phys. **64**, 1246 (1996).
[35] Y. Z. Chen and J. Qiu, Mol. Cell. Biol. Res. Commun. **2**, 145 (1999).
[36] J. M. Raser and E. K. O'Shea, Science **309**, 2010 (2005).
[37] J. Selimkhanov, B. Taylor, J. Yao, A. Pilko, J. Albeck, A. Hoffmann, L. Tsimring, and R. Wollman, Science **346**, 1370 (2014).
[38] L. Cardelli, A. Csikasz-Nagy, N. Dalchau, M. Tribastone, and M. Tschaikowski, Sci. Rep. **6**, 20214 (2016).
[39] P. S. Swain, M. B. Elowitz, and E. D. Siggia, Proc. Natl. Acad. Sci. USA **99**, 12795 (2002).
[40] T. Long, K. C. Tu, Y. Wang, P. Mehta, N. P. Ong, B. L. Bassler, and N. S. Wingreen, PLoS Biol. **7**, e1000068 (2009).
[41] M. A. Schwartz and H. D. Madhani, Annu. Rev. Genet. **38**, 725 (2004).