Optimal information transfer in enzymatic networks: A field theoretic formulation

Himadri S. Samanta, Michael Hinczewski, and D. Thirumalai

Optimal information transfer in enzymatic networks: A field theoretic formulation

Himadri S. Samanta,[1] Michael Hinczewski,[2] and D. Thirumalai[1]

[1]*Department of Chemistry, The University of Texas at Austin, TX 78712*

[2]*Department of Physics, Case Western Reserve University, OH 44106*

(Dated: June 19, 2017)

Abstract

Signaling in enzymatic networks is typically triggered by environmental fluctuations, resulting in a series of stochastic chemical reactions, leading to corruption of the signal by noise. For example, information flow is initiated by binding of extracellular ligands to receptors, which is transmitted through a cascade involving kinase-phosphatase stochastic chemical reactions. For a class of such networks, we develop a general field-theoretic approach in order to calculate the error in signal transmission as a function of an appropriate control variable. Application of the theory to a simple push-pull network, a module in the kinase-phosphatase cascade, recovers the exact results for error in signal transmission previously obtained using umbral calculus (Phys. Rev. X., **4**, 041017 (2014)). We illustrate the generality of the theory by studying the minimal errors in noise reduction in a reaction cascade with two connected push-pull modules. Such a cascade behaves as an effective three-species network with a pseudo intermediate. In this case, optimal information transfer, resulting in the smallest square of the error between the input and output, occurs with a time delay, which is given by the inverse of the decay rate of the pseudo intermediate. Surprisingly, in these examples the minimum error computed using simulations that take non-linearities and discrete nature of molecules into account coincides with the predictions of a linear theory. In contrast, there are substantial deviations between simulations and predictions of the linear theory in error in signal propagation in an enzymatic push-pull network for a certain range of parameters. Inclusion of second order perturbative corrections shows that differences between simulations and theoretical predictions are minimized. Our study establishes that a field theoretic formulation of stochastic biological signaling offers a systematic way to understand error propagation in networks of arbitrary complexity.

INTRODUCTION:

Cell signaling involves the ability of cells to detect changes in the environment and respond to them [1–6], a fundamental necessity of living systems. Several signaling networks involve proteins, which switch between active and inactive states. By quantitatively describing how different signaling proteins are functionally linked, we can understand the behavior of signaling pathways, and the associated bandwidth that determines fidelity of information transfer [7]. In typical enzymatic networks, environmental information is transmitted into the cell interior through cascades of stochastic biochemical reactions [8]. Noise inevitably propagates through the cascade, potentially corrupting the signal. Depending on the parameters, small changes in the input can be translated into large (but noise corrupted) output variations. The amplification is essential but it must also preserve the signal content to be useful for downstream processes. The signaling circuit, despite operating in a noisy environment, needs to maintain high fidelity between output and the amplified input [9]. Over the years, concepts in information theory have been adopted to assess the fidelity of signal transmission [10, 11]. Several studies have used mutual information between input and output signals to quantify the reliability of signal transduction [12–17]. The formalism has been applied to the study of a variety of networks including cascades and networks with feedback [12]. These, and other studies have expanded our understanding of the fidelity of information transfer in biological networks in which both noise and copy number fluctuations are important.

In a recent paper [18], we considered the problem of how to extract information faithfully from noisy signals using mathematical methods developed in the context of communication theory developed over sixty years ago by Wiener [19] and independently by Kolmogorov [20]. The Wiener-Kolmogorov (WK) approach has since proven a useful tool in a variety of contexts in biological signaling [9, 21, 22]. The WK theory, reformulated by Bode and Shannon [11], assumes that the input and output are continuous variables that describe stationary stochastic processes. The goal of approach is to minimize the mean squared error between the input and output signals, but the optimization is restricted to the space of only linear noise filters. Recently, we developed an analytic formalism of general validity to overcome some of the limitations of the WK theory based on exact techniques involving umbral calculus [23]. We illustrated the efficacy of the non-linear theory with applications

to push-pull network and its variants including instances when the input is time-dependent.

The use of non-standard mathematics in the form of umbral calculus, perhaps, obscures the physics of optimal filtering in biological networks in which the effects of non-linearities in signal amplification have to be considered. Here, we develop an alternate general formalism based on a many body formulation of reaction diffusion equations introduced by Doi and Peliti [24, 25]. This formulation converts the signal optimization problem to a standard field theory, allowing us to calculate the response and correlation functions by standard methods. The advantage of this formalism is that both discrete and continuum cases can be studied easily. Non-linear contributions can be calculated using systematic diagrammatic perturbation scheme for an arbitrary network. Networks where temporal dynamics are coupled with spatial gradients in signaling activities, which regulate intracellular processes and signal propagation across the cell, can also be investigated using the present formalism. Application of the theory to a push-pull network and a simplified biochemical network recovers the exact results obtained in our previous study. We also extend the formalism to solve signal transduction in a cascade, which serves as a model for a variety of biological networks. The formalism is general and is applicable to arbitrary networks with feedback, time delay and special variations [26]. Our work exploits standard methods in physics, illustrating the usefulness of a field theoretic formulation at the interface of communication theory and biology.

THEORY:

Linear Push-Pull Network: In order to develop the many body formalism for a general signaling network, we first consider a simple model. The concepts and the general diagrammatic expansion developed in this context, lays the foundation for applications to more complicated enzymatic networks as well as signaling cascades. In a typical signaling pathway, for example the mitogen activated protein kinase (MAPK) [27, 28] pathway, external and environmental fluctuations activate a cascade of enzymatic reactions, thus transmitting information across the membrane in a sequential manner. This pathway begins with a cell-surface receptor with tyrosine kinase activity, connected to a G protein that is closely related to Ras, which is known to be an oncogene in vertebrates. This is followed by a cascade of three serine/threonine protein kinases (termed the MAP kinase cascade).

Each step involves activation of kinases by phosphorylation reaction and deactivation by phosphatases [29–33]. A truncated version of such a cascade is a single step (Fig.1), which we refer to as a push-pull network [18]. The kinase acts as a "push" element that drives the substrate to a new state, while phosphatase acts as a "pull" that returns the substrate to the original state. All information storage and transmission systems, be they natural cellular systems or human-made electronic ones, require some sort of mechanism like this to write and erase information in a controllable fashion. Enzymes remain chemically unaltered while performing their target reaction, and thus they are capable of carrying out many rounds of the reaction. Thus, catalytic behavior can result in signal amplification: a single activated enzyme molecule can, under the right conditions, generate many molecules as product. In this simple signaling network, there are only two chemical species. One is $I(t)$ (the "input") and the other is $O(t)$ (the "output") whose production depends on $I(t)$. The upstream pathway, which serves as an external signal, creates the species I by the reaction $\phi \xrightarrow{F} I$ with an effective production rate F. The output O is a result of the reaction $I \xrightarrow{R(I)} I + O$, with a rate $R(I(t))$ that depends on the input. The species are deactivated through $I \xrightarrow{\gamma_I} \phi$ and $O \xrightarrow{\gamma_O} \phi$ with rates γ_I and γ_O respectively, mimicking the role of phosphatases (Fig.1). The input varies over a characteristic time scale γ_I^{-1}, fluctuating around the mean value $\bar{I} = F/\gamma_I$. The degradation rate sets the time scale γ_O^{-1} over which $O(t)$ responds to changes in the input.

The chemical Langevin equations describing the changes in I and O are,

$$\frac{dI}{dt} = F - \gamma_I I + \eta_I, \quad \frac{dO}{dt} = R(I) - \gamma_O O + \eta_O, \tag{1}$$

where η_I and η_O are Gaussian white noise with zero mean ($\langle \eta_\alpha \rangle = 0$) and correlation $<\eta_\alpha(t)\eta'_\alpha(t')> = 2\sqrt{\gamma_\alpha \bar{\alpha}}\delta_{\alpha\alpha'}\delta(t - t')$ with $\alpha = I, O$ and $\bar{\alpha}$ is the mean population α. For small fluctuations, $\delta\alpha(t) = \alpha(t) - \bar{\alpha}$, Eq.(1) can be solved using a linear approximation for the rate function $R(I(t)) \approx R_0\bar{I} + R_1\delta I(t)$, with coefficients $R_0, R_1 > 0$. The result is

$$\delta I(t) = \int_{-\infty}^{t} dt' e^{-\gamma_I(t-t')}\eta_I(t'), \tag{2}$$

$$\delta O(t) = \int_{-\infty}^{t} dt' \frac{R_1}{G} e^{-\gamma_O(t-t')} \left[G\delta I(t') + \frac{G}{R_1}\eta_O(t') \right],$$

where in the second line an arbitrary scaling factor G has been introduced. The solution for $\delta O(t)$ has the structure of a linear noise filter equation; $\tilde{s} = \int_{-\infty}^{t} dt' H(t - t')c(t')$, with $c(t) = s(t) + n(t)$. The signal $s(t) = G\delta I(t)$ together with the noise term $n(t) \equiv GR_1^{-1}\eta_O(t)$

constitute the corrupted signal, $c(t)$. The output $\tilde{s}(t) \equiv \delta O(t)$ is produced by convolving $c(t)$ with a linear kernel $H(t) \equiv R_1 G^{-1} \exp(-\gamma_O t)$, which filters the noise. As a consequence of causality, the filtered output $\tilde{s}$ at time t depends only on $c(t')$ from the past.

The primary goal in transmitting signal with high fidelity is to devise an optimal causal filter, $H_{opt}(t)$, which renders $\tilde{s}(t)$ as close to $s(t)$ as possible. In a remarkable development, Weiner [19] and Kolmogorov [20] independently discovered a solution to this problem in the context of communication theory, which launched the modern era in signal decoding from time series. In particular, WK proposed a solution that minimizes the square of the differences between $\tilde{s}$ and $s(t)$ by seeking an optimal filter $H_{WK}(t)$ among all possible linear filters. In the push-pull network, this means having $\delta O(t)$ reproduce as accurately as possible the scaled input signal $G\delta I(t)$. For a particular $\delta I(t)$ and $\delta O(t)$, the value of the mean squared error $E = \langle (\tilde{s} - s)^2 \rangle / \langle s^2 \rangle$ is smallest when $G = \langle (\delta O)^2 \rangle / \langle \delta O \delta I \rangle$, which we identify as a gain factor. In this case, $E = 1 - \langle \delta O \delta I \rangle^2 / (\langle (\delta O)^2 \rangle \langle (\delta I)^2 \rangle)$.

The optimal causal filter H_{WK} satisfies the following Wiener-Hopf equation[9, 18],

$$C_{cs}(t) = \int_{-\infty}^{t} dt' H_{WK}(t - t') C_{cc}(t'), \quad t > 0 \tag{3}$$

where $C_{xy}(t) \equiv \langle x(t')y(t' + t) \rangle$ is the correlation between points in the time series x and y, assumed to depend only on the time difference $t - t'$. We can evaluate the correlation functions C_{cs} and C_{cc} using Eq. (2), and substituting these solutions in Eq. (3), the optimal filter function can be solved by assuming a generic ansatz, $H_{WK}(t) = \sum_{i=1}^{N} A_i \exp(-\lambda_i t)$. The unknown coefficients, A_i, and the associated rate constants λ_i are found by comparing the left and right hand sides of Eq. (3). Elsewhere [18], we showed that $H_{WK}(t) = \gamma_I(\sqrt{1 + \Lambda} - 1) \exp(\gamma_I \sqrt{1 + \Lambda} t)$. The conditions for achieving WK optimality, $H(t) = H_{WK}(t)$, are [18],

$$\gamma_O = \gamma_I \sqrt{1 + \Lambda}, \quad G = \frac{R_1}{\gamma_I(\sqrt{1 + \Lambda} - 1)}, \tag{4}$$

leading to the minimum relative error,

$$E_{WK} = \frac{2}{1 + \sqrt{1 + \Lambda}}, \quad \Lambda \equiv \frac{R_1^2}{R_0 \gamma_I}. \tag{5}$$

The fidelity between the output and input is described through a single dimensionless optimality control parameter, Λ, which can be written as $\Lambda \equiv (R_0/\gamma_I)(R_1/R_o)^2$. The first term, R_0/γ_I, is a burst factor, measuring the mean number of output molecules produced per input molecule during the active lifetime of the input molecule. The second term, $(R_1/R_o)^2$,

is a sensitivity factor, reflecting the local response of the production function $R(I)$ near $\bar{I}$ (controlled by the slope $R_1 = R'(\bar{I})$) relative to the production rate per input molecule $R_0 = R(\bar{I})/\bar{I}$.

In our recent work [18], we extended the WK approach to include non-linearity and the discrete nature of the input and output molecules I and O [18]. Both these considerations are relevant in biological circuits where $R(I)$ is non-linear and the copy numbers of I and O are likely to be small. Starting from the exact master equation, valid for discrete populations and arbitrary $R(I)$, we rigorously solved the original optimization problem for the error E between output and input using the principles of umbral calculus[23]. Umbral calculus starts by noting that $R(I)$ can be expanded as a Newton series (the finite difference analog of the Taylor series). The main results are as follows. For any arbitrary function $R(I)$ is expanded as,

$$R(I) = \sum_{n=0}^{\infty} \sigma_n v_n(I), \tag{6}$$

with $v_n(I) = \sum_{m=0}^{\infty}(n-m)!(-\bar{I})^m \binom{n}{m}\binom{I}{n-m}$ and $\sigma_n = <v_n(I)R(I)>/(\bar{I}^n n!)$, the relative error can be expressed by an exact expression,

$$E = 1 - \frac{\bar{I}\gamma_O^2\sigma_1^2}{(\gamma_I + \gamma_O)^2}\left[\gamma_O\sigma_0 + \sum_{n=1}^{\infty}\sigma_n^2 \frac{n!\gamma_O\bar{I}^n}{\gamma_O + n\gamma_I}\right]^{-1}. \tag{7}$$

The expression above is bounded from below by

$$E \geq E_{opt} \equiv \frac{2}{1 + \sqrt{1 + \tilde{\Lambda}}}, \tag{8}$$

where $\tilde{\Lambda} = \bar{I}\sigma_1^2/(\sigma_0\gamma_I)$. The equality is only reached when $\gamma_O = \gamma_I\sqrt{1 + \tilde{\Lambda}}$ and $R(I)$ is an optimal linear filter of the form, $R_{opt}(I) = \sigma_0 + \sigma_1(I - \bar{I})$, with all $\sigma_n = 0$ for $n \geq 2$. Obtaining the lower bound is important for noise reduction in biological networks as it provides insights into energy costs required to reduce the error [10]. It follows from Eq.(8) that in order to reduce error by a factor of 10 the number of signaling events has to increase by 10^4 [10]. Because signaling events are triggered by chemical reactions (in the kinase-phosphatase network it involves ATP) this requires considerable energy cost.

Field theoretic formulation: In order to generalize the results in our previous study [18] to arbitrary regulatory networks, we adopt a many body approach pioneered by Doi and Peliti[24, 25]. Such an approach has been used in the study of a variety of

reaction diffusion equations [34, 35]. Besides suggesting plausible new ways of examining how signals are transmitted in biochemical reaction networks, the current theory shows how standard field theoretic methods can be adopted for use in control theory. By way of demonstrating its utility, we rederive the exact analytical solution (Eq.(7)) for the relative error in the push-pull network. In the Doi-Peliti formalism, the configurations at time t in a locally interacting many body system are specified by the occupation numbers of each species on a lattice site i. In our case, I_i is the input population and O_i is the output population. As a consequences of the stochastic dynamics, the on-site occupation numbers are modified. Arbitrarily many particles of either population are allowed to occupy any lattice site. In other words, $I_i, O_i = 0, 1, \cdots \infty$. The master equation for the local reaction scheme that governs the time evolution of the configurational probability with I_i input and O_i output at site i at time t is obtained through the balance of gain and loss terms. The result is,

$$\frac{dP(I_i, O_i, t)}{dt} = \gamma_I[(I_i + 1)P(I_i + 1, O_i, t) - I_i P(I_i, O_i, t)] + F[P(I_i - 1, O_i, t) - P(I_i, O_i, t)] \quad (9)$$
$$+ \gamma_O[(O_i + 1)P(I_i, O_i + 1, t) - O_i P(I_i, O_i, t)] + R(I_i)[P(I_i, O_i - 1, t) - P(I_i, O_i, t)].$$

We use the Fock space representation to account for the changes in the site occupation number by integer values for the chemical reactions describing the network. Following Doi and Peliti, we introduce the bosonic ladder operator algebra with commutation relation $[a_i, a_j] = 0$, $[a_i, a_j^\dagger] = \delta_{ij}$ for the input population, allowing us to construct the input particle number eigenstates $|I_i\rangle$ obeying $a_i|I_i\rangle = I_i|I_i - 1\rangle$, $a_i^\dagger|I_i\rangle = |I_i + 1\rangle$, $a_i^\dagger a_i|I_i\rangle = I_i|I_i\rangle$. A Fock state with I_i particles on site i is obtained from the vacuum state $|0\rangle$, defined by the relation $a_i|0\rangle = 0$, and $|I_i\rangle = a_i^{\dagger I_i}|0\rangle$. Similarly, we introduce annihilation and creation operators for output particles b_i and $b_i^\dagger$ that commute with the input ladder operators: $[a_i, b_j] = 0 = [a_i, b_j^\dagger]$.

Stochastic kinetics for the entire lattice is implemented by considering the master equation for the configurational probability $P(\{I_i\}, \{O_j\}, t)$, given by a sum over all lattice points on the right hand side of Eq.(9), by noting that a general Fock state is constructed by the tensor product $|\{I_i\}, \{O_j\}\rangle = \Pi_i|I_i\rangle|O_i\rangle$. We define a time dependent formal state vector through a linear combination of all possible Fock states, weighted by their configurational

probability at time t,

$$|\psi(t)\rangle = \sum_{\{I_i, O_j\}}^{\infty} P(\{I_i\}, \{O_j\}, t)\Pi_{\{i,j\}}a_i^{\dagger I_i}b_i^{\dagger O_j}|0\rangle. \tag{10}$$

This superposition state encodes the stochastic temporal evolution. We use standard methods to transform the time dependence from the linear master equation into an imaginary time Schrödinger equation, governed by a time-dependent stochastic evolution operator H,

$$\frac{d}{dt}|\psi(t)\rangle = -H(\{a\}, \{a^\dagger\}; \{b\}, \{b^\dagger\})|\psi(t)\rangle. \tag{11}$$

We may multiply Eq.(9) by $a_i^{\dagger I_i}b_i^{\dagger O_i}|0\rangle$, and sum over all values of I_i, O_i. With the definition of the state $|\psi(t)\rangle$,

$$|\psi(t)\rangle = \sum_{I_i, O_i}^{\infty} P(I_i, O_i, t)a_i^{\dagger I_i}b_i^{\dagger O_i}|0\rangle, \tag{12}$$

the γ_I term, i.e. $\gamma_I[(I_i + 1)P(I_i + 1, O_i, t) - I_i P(I_i, O_i, t)]$, in Eq.(9) becomes,

$$\gamma_I \sum_{I_i, O_i}^{\infty} [(I_i + 1)P(I_i + 1, O_i) - I_i P(I_i, O_i)]a_i^{\dagger I_i}b_i^{\dagger O_i}|0\rangle$$

$$= \gamma_I \sum_{I_i, O_i}^{\infty} [P(I_i + 1, O_i)a_i a_i^{\dagger I_i+1}b_i^{\dagger O_i}|0\rangle - P(I_i, O_i)a_i^\dagger a_i a_i^{\dagger I_i}b_i^{\dagger O_i}|0\rangle]. \tag{13}$$

By relabeling the indices in the first sum, we arrive at the desired Hamiltonian expressed in second quantized representation as, $H_{\gamma_I} = -\gamma_I(1 - a_i^\dagger)a_i$. Similarly, terms with coefficients F, γ_O and $R(I_i)$ in Eq.(9) give the following contributions, $H_F = -F(a_i^\dagger - 1)$, $H_{\gamma_O} = -\gamma_O(1 - b_i^\dagger)b_i$, $H_{R(I_i)} = -R(a_i^\dagger a_i)(b_i^\dagger - 1)$. The total Hamiltonian H takes the following form,

$$H = H_{\gamma_I} + H_F + H_{\gamma_O} + H_{R(I)}$$

$$= \sum_i [-\gamma_I(1 - a_i^\dagger)a_i - F(a_i^\dagger - 1) - \gamma_O(1 - b_i^\dagger)b_i - R(a_i^\dagger a_i)(b_i^\dagger - 1)]. \tag{14}$$

A convenient choice for the initial configuration for the master equation describing the stochastic particle reactions is an independent Poisson distribution at each site,

$$P(\{I_i\}, \{O_i\}; 0) = \Pi_i P_0(I_i)P_0(O_i) = \Pi_i e^{-\bar{I}_0}e^{-\bar{O}_0}O_0^{-O_i}I_0^{-I_i}/I_i!O_i!. \tag{15}$$

with mean initial input and output concentrations $\bar{I}_0$ and $\bar{O}_0$. Just as in quantum mechanics, Eq.(11) can be formally solved leading to,

$$|\psi(t)\rangle = e^{Ht}|\psi(0)\rangle, \tag{16}$$

8

with the initial state $|\psi\rangle = e^{\bar{I}_0 \sum_i (a_i^\dagger - 1) + \bar{O}_0 \sum_i (b_i^\dagger - 1)} |0\rangle$.

Our goal is to compute averages and correlation functions with respect to the configurational probability $P(\{I_i\}, \{O_i\}; t)$, which is accomplished by means of the projection state $\langle\mathcal{P}| = \langle 0 | \Pi_i e^{a_i + b_i}$, for which $\langle\mathcal{P}|0\rangle = 1$ and $\langle\mathcal{P}|a_i^\dagger = \langle\mathcal{P}| = \langle\mathcal{P}|b_i^\dagger$, since $[e^{a_i}, a_j^\dagger] = e^{a_i}\delta_{ij}$. The average value of an observable $A(\{I_i\}, \{O_i\})$ is,

$$\langle A(t)\rangle = \sum_{\{I_i\}, \{O_i\}} A(\{I_i\}, \{O_i\}) P(\{I_i\}, \{O_i\}; t), \tag{17}$$

from which the statistical average of an observable can be calculated using,

$$\langle A(t)\rangle = \langle\mathcal{P}|A(\{a_i^\dagger, a_i; b_i^\dagger, b_i\})|\psi(t)\rangle \tag{18}$$

$$= \langle\mathcal{P}|A(\{a_i^\dagger, a_i; b_i^\dagger, b_i\}) e^{-H((\{a_i^\dagger\}, \{a_i\}; \{b_i^\dagger\}, \{b_i\}) t} |\psi(0)\rangle.$$

We follow a well-established route in quantum many particle theory [36], and proceed towards a field theory representation by constructing a path integral equivalent of the time dependent Schrödinger equation (Eq.(11)) based on coherent states [37]. These are defined as right eigenstates of the annihilation operators, $a_i|\alpha_i\rangle = \alpha_i|\alpha_i\rangle$ and $a_i|\beta_i\rangle = \beta_i|\beta_i\rangle$, with complex eigenvalues α_i and β_i. The coherent states satisfy $|\alpha_i\rangle = \exp(\frac{1}{2}|\alpha_i|^2 + \alpha_i\alpha_i^\dagger)|0\rangle$, the overlap integral $\langle\alpha_j|\alpha_i\rangle = \exp(-\frac{1}{2}|\alpha_i|^2 - \frac{1}{2}|\alpha_j|^2 + \alpha_j^*\alpha_i)$, and the completeness relation $\int \Pi_i d^2\alpha_i|\{\alpha_i\}\rangle\langle\{\alpha_i\}| = \pi$. After splitting the temporal evolution (Eq.(11)) into infinitesimal increments, inserting the completeness relation at each time step, and with additional manipulations leads to an expression for the configurational average,

$$\langle A(t)\rangle \propto \int \Pi_i d\alpha_i d\alpha_i^* d\beta_i d\beta_i^* A(\{\alpha_i\}, \{\beta_i\}) e^{-\mathcal{S}[\alpha_i^*, \beta_i^*, \alpha_i, \beta_i]}. \tag{19}$$

The exponential statistical weight is determined by the action,

$$\mathcal{S}[\alpha_i^*, \beta_i^*, \alpha_i, \beta_i] = \sum_i \left[\int_0^{t_f} \left\{ \alpha_i^*(t)\frac{\partial\alpha_i(t)}{\partial t} + \beta_i^*(t)\frac{\partial\beta_i(t)}{\partial t} \right\} + H(\alpha_i^*, \beta_i^*, \alpha, \beta) \right] dt. \tag{20}$$

Finally, by taking the continuum limit using $\sum_i \to a_0^{-d} \int d^d x$, a_0 is a lattice constant, $\alpha_i(t) \to \phi(x, t)$, $\beta_i(t) \to \psi(x, t)$ and $\alpha_i(t) \to a_0^d\phi(x, t)$, $\beta_i^*(t) \to a_0^d\psi^*(x, t)$, the expectation value is represented by a functional integral,

$$\langle A(t)\rangle \propto \int \Pi_i \mathcal{D}[\phi^*, \phi, \psi^*, \psi] A(\{\phi\}, \{\psi\}) e^{-\mathcal{S}[\psi^*, \phi^*, \psi, \phi]} \tag{21}$$

with an effective action

$$\mathcal{S}[\psi^*, \phi^*, \psi, \phi] = \int_0^{t_f} \left[\left\{ \psi^*(t)\frac{\partial\psi(t)}{\partial t} + \phi^*(t)\frac{\partial\phi(t)}{\partial t} \right\} + H(\psi^*, \phi^*, \psi, \phi) \right] dt. \tag{22}$$

9

In the Hamiltonian (Eq.(14)), $a^\dagger$ and $b^\dagger$ are replaced by the field variables ϕ^* and ψ^*, respectively. Similarly, a and b operators become ϕ and ψ respectively.

The action in Eq.(22) encodes the stochastic master equation kinetics through four independent fields $(\psi^*, \phi^*, \psi, \phi)$. With this formulation, an immediate connection can be made to the response functional formulation using the Janssen - De Dominicis formalism for Langevin equations [38, 39]. In this approach, the response field enters at most quadratically in the pseudo-Hamiltonian, which may be interpreted as averaging over Gaussian white noise. With this in mind, we apply the non-linear Cole-Hopf transformation [40, 41], in order to obtain quadratic terms in auxiliary fields, $\phi^* = e^{\bar{\phi}_I}$, $\phi = e^{-\bar{\phi}_I}\phi_I$, $\psi^* = e^{\bar{\psi}_O}$, $\psi = e^{-\bar{\psi}_O}\psi_O$, to the action in Eq.(22). The Jacobian for this variable transformation is unity, and the local particle densities are $\phi^*\phi = \phi_I$ and $\psi^*\psi = \psi_O$. We obtain the following Hamiltonian,

$$H = -\gamma_I(-\bar{\phi}_I + \frac{\bar{\phi}_I^2}{2})\phi_I - F(\bar{\phi}_I + \frac{\bar{\phi}_I^2}{2}) - \gamma_O(-\bar{\psi}_O + \frac{\bar{\psi}_O^2}{2})\psi_O - R(\phi_I)(\bar{\psi}_O + \frac{\bar{\psi}_O^2}{2}). \qquad (23)$$

In the above equation, the exponential term has been expanded to second order. The rate equations are obtained through $\delta\mathcal{S}/\delta\bar{\psi}\mid_{\bar{\psi}=0} = 0$ and $\delta\mathcal{S}/\delta\bar{\phi}\mid_{\bar{\phi}=0} = 0$. The terms quadratic in the auxiliary fields ($\bar{\psi}$ and $\bar{\phi}$) encapsulate the second moment of the Gaussian white noise with zero mean.

In order to obtain fluctuation corrections needed to calculate minimum error in signal transduction, we write the action in terms of fluctuating fields, $\delta\phi_I = \phi_I - \langle\phi_I\rangle$ and $\delta\psi_O = \psi_O - \langle\psi_O\rangle$ as,

$$H = \bar{\phi}_I[\gamma_I\delta\phi_I - \gamma_I\langle\phi_I\rangle\bar{\phi}_I] +$$
$$\bar{\psi}_O[\gamma_O\delta\psi_O - \{c_1\delta\phi_I + \frac{c_2}{2}\delta\phi_I^2 + \cdots\} - \gamma_O\langle\psi_O\rangle\bar{\psi}_O] \qquad (24)$$

where we have expanded $R(\phi_I)$ in a Taylor series,

$$R(\phi_I) = \sum_0^\infty \frac{c_n}{n!}(\delta\phi_I)^n, \qquad (25)$$

with constant c_n. Note this expansion differs from the one used in Eq.(6). The coefficients of $\bar{\phi}_I^2$ and $\bar{\psi}_O^2$ reflect the noise correlations in Langevin description.

In Fourier space the action becomes

$$\mathcal{S}[\tilde{\Psi}, \Psi] = \int_w \bar{\phi}_I[-iw\ \delta\phi_I + \gamma_I\delta\phi_I - \gamma_I\langle\phi_I\rangle\bar{\phi}_I] + \qquad (26)$$
$$\bar{\psi}_O[-iw\ \delta\psi_O + \gamma_O\delta\psi_O - c_1\delta\phi_I - \gamma_O\langle\psi_O\rangle\bar{\psi}_O]$$
$$+\mathcal{S}_{int}[\tilde{\Psi}, \Psi]$$

where $\tilde{\Psi}$ represents the set $\{\bar{\phi}_I, \bar{\psi}_O\}$ and Ψ denotes $\{\phi_I, \psi_O\}$. The non-linear contribution to the action is $\mathcal{S}_{int}[\tilde{\Psi}, \Psi] = \int_w \bar{\psi}_O[\frac{c_2}{2}\delta\phi_I(w_1)\delta\phi_I(w - w_1)] + \cdots$. Physical quantities can be expressed in terms of correlation functions of fields Ψ and $\tilde{\Psi}$, taken with the statistical weight $e^{-\mathcal{S}[\tilde{\Psi}, \Psi]}$,

$$\langle \Psi\tilde{\Psi} \rangle = \frac{\int \mathcal{D}[i\tilde{\Psi}] \int \mathcal{D}[\Psi] \Psi\tilde{\Psi} e^{-\mathcal{S}[\tilde{\Psi}, \Psi]}}{\int \mathcal{D}[i\tilde{\Psi}] \int \mathcal{D}[\Psi] e^{-\mathcal{S}[\tilde{\Psi}, \Psi]}}. \tag{27}$$

In order to compute the correlation function involving response fields, it is useful to introduce the generating functional,

$$\mathcal{Z}[\tilde{J}, J] = \langle \exp \int_t \sum_\alpha (\tilde{J}_\alpha(t)\tilde{\Psi}_\alpha(t) + J_\alpha(t)\Psi_\alpha(t)) \rangle \tag{28}$$

where α represents the set $\{\phi_I, \psi_O\}$, for which the required correlation functions are obtained via functional derivatives of $\mathcal{Z}$ with respect to the appropriate source fields.

The procedure is readily implemented for the Gaussian theory with statistical weight $e^{-\mathcal{S}_0[\tilde{\Psi}, \Psi]}$. In Fourier space, we can write the harmonic function as,

$$\mathcal{S}_0[\tilde{\Psi}, \Psi] = \int_w \sum_\alpha \left(\tilde{\Psi}_\alpha(-w) \ \Psi_\alpha(-w) \right) \mathcal{M} \begin{pmatrix} \tilde{\Psi}_\alpha(w) \\ \Psi_\alpha(w) \end{pmatrix} \tag{29}$$

with the Hermitian coupling, a (4,4) matrix $\mathcal{M}(w)$. With the aid of Gaussian integrals, we obtain,

$$\mathcal{Z}_0[\tilde{J}, J] = \exp \left[\frac{1}{2} \int_w \sum_\alpha \left(\tilde{J}_\alpha(-w) \ J_\alpha(-w) \right) \mathcal{M}^{-1} \begin{pmatrix} \tilde{J}_\alpha(w) \\ J_\alpha(w) \end{pmatrix} \right]. \tag{30}$$

From Eq.(30), we now directly infer the matrix of two point correlation functions in the Gaussian ensemble with the inverse of harmonic coupling matrix $\mathcal{M}$.

APPLICATIONS:

As a first application we apply the field-theoretic formalism to the push-pull network, which can be exactly solved for the error (Eq.(5)). In the process we illustrate the way the diagrammatic expansion works in the context of signaling networks, making it possible to apply the theory to more complicated systems.

A. Push-Pull network:

The calculation of the error (Eq.(5)) in terms of the control variable (the average number of phosphatase molecules per cell ($\bar{P}$)) requires the correlation functions $\langle \delta O \delta I \rangle$, $\langle \delta O^2 \rangle$ and $\langle \delta I^2 \rangle$. These can be expressed in terms of the matrix elements of $(\mathcal{M}^{-1})_{mn}$ (Eq.(29)). Subscripts m and n represent the mth row and n^{th} column, respectively. For example, $(\mathcal{M}^{-1})_{33}$ is the correlation function $\langle \delta\phi_I(-w)\delta\phi_I(w) \rangle$. Similarly we can obtain other correlation functions. Now, we can compute power spectra for the input and output molecules by evaluating the correlation functions of kinase and substrate populations using Eq. (30). We use perturbation theory for the action corresponding to the push-pull network to compute the non-linear contribution to the correlation function.

We obtain the following expressions for the power spectra,

$$\langle \delta\phi_I(-w)\delta\phi_I(w) \rangle_0 = \frac{2\gamma_I \langle \phi_I \rangle}{(-iw + \gamma_I)(iw + \gamma_I)}$$

$$\langle \delta\psi_O(-w)\delta\psi_O(w) \rangle_0 = \frac{c_1^2 2\gamma_I \langle \phi_I \rangle}{(-iw + \gamma_O)(-iw + \gamma_I)(iw + \gamma_O)(iw + \gamma_I)}$$
$$+ \frac{2\gamma_O \langle \phi_O \rangle}{(-iw + \gamma_O)(iw + \gamma_O)}$$

$$\langle \delta\phi_I(-w)\delta\psi_O(w) \rangle_0 = \frac{2c_1 \gamma_I \langle \phi_I \rangle}{(-iw + \gamma_O)(-iw + \gamma_I)(iw + \gamma_I)} \tag{31}$$

The $\langle \cdots \rangle_0$ is taken with respect to the non-interacting theory ($S_{int}[\tilde{\Psi}, \Psi] = 0$ in Eq.(26)). Using these functions, the error (E) and gain (G) are given by,

$$E = \frac{\langle (\delta\psi_O - G\delta\phi_I)^2 \rangle}{G^2 \langle \delta\phi_I^2 \rangle}, \qquad G = \frac{\langle \delta\psi_O^2 \rangle}{\langle \delta\phi_I \delta\psi_O \rangle}. \tag{32}$$

By inserting the expressions for the correlation functions in Eq.(31) into Eq.(32), and integrating over w, we obtain the minimum relative error for the linear push-pull network,

$$E = 1 - \frac{\bar{I}\gamma_O^2\sigma_1^2}{(\gamma_I + \gamma_O)^2} \left[\gamma_O\sigma_0 + \sigma_1^2 \frac{\gamma_O \bar{I}}{\gamma_O + \gamma_I} \right]^{-1}. \tag{33}$$

Higher order corrections to the power spectra $\langle \delta\psi_O(-w)\delta\psi_O(w) \rangle$ are calculated using perturbation theory by evaluating the Feynman diagrams (Fig.(2)),

$$\langle \delta\psi_O \delta\psi_O \rangle = \frac{\langle \delta\psi_O \delta\psi_O \sum_l^\infty (-\mathcal{S}_{int}[\tilde{\Psi}, \Psi])^l / l! \rangle_0}{\langle \sum_l^\infty (-\mathcal{S}_{int}[\tilde{\Psi}, \Psi])^l / l! \rangle_0}. \tag{34}$$

For example, the second order contribution to the $\langle \delta\psi_O(-w)\delta\psi_O(w)\rangle$ arising from the loop in Fig.(2) is $\Omega_2^2 \frac{2!\bar{I}^2}{\gamma_O(\gamma_O+2\gamma_I)}$ (see Appendix A for details). The coefficient Ω_2^2 is given by $\Omega_2^2 = \frac{c_2^2}{4} + \frac{c_3^2}{4} + \frac{\bar{I}}{4}c_2c_4 + \cdots$. Higher order terms have a similar structure: for example, the third order contribution to the power spectra is $\Omega_3^2 \frac{3!\bar{I}^2}{\gamma_O(\gamma_O+3\gamma_I)}$, with $\Omega_3^2 = \frac{c_3^2}{36} + \frac{c_4^2}{16} + \frac{\bar{I}}{36}c_3c_5 + \cdots$. By evaluating all the diagrams in Fig.(2), we obtain the final expression for the relative error,

$$E = 1 - \frac{\bar{I}\gamma_O^2\sigma_1^2}{(\gamma_I+\gamma_O)^2}\left[\gamma_O\sigma_0 + \sum_{n=1}^{\infty}\Omega_n^2\frac{n!\gamma_O\bar{I}^n}{\gamma_O+n\gamma_I}\right]^{-1}, \quad \text{with } \Omega_1 = \sigma_1. \tag{35}$$

The form of the result in Eq.(35) coincides with the exact expression (Eq.(7)) for the relative error previously obtained [18] by using an entirely different approach based on umbral calculus. However, the coefficients Ω_n are expressed in terms of the coefficients c_n used in the series for $R(I)$ (Eq.(25)) rather than σ_n. The two kinds of coefficients are non-trivially related through,

$$\sigma_n = \sum_{m=0}^{\infty}\sum_{p=0}^{m}\sum_{q=0}^{p}\frac{1}{m!}\binom{m}{p}\binom{q}{n}S_{pq}(-\bar{I})^{m-p}\bar{I}^{q-n}c_m, \tag{36}$$

where S_{pq} are Stirling's numbers of second kind. For all n, the leading order term $\frac{c_n^2}{n!^2}$ of Ω_n is the same as the leading order term of σ_n.
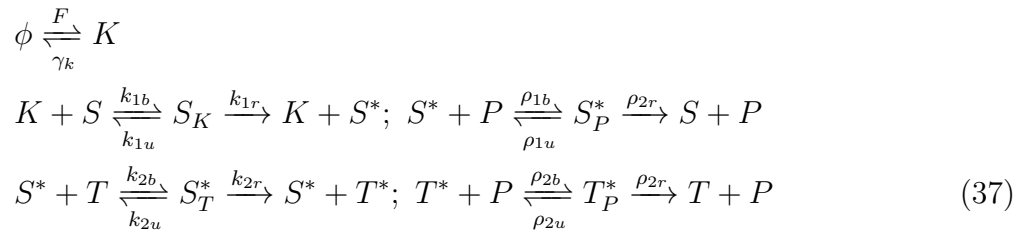
The sum within the bracket in Eq.(35) is composed of non-negative terms. The minimal sum E is obtained by setting $\Omega_n = 0$ for all $n \geq 2$. Thus, E is bounded from below by $E \geq 1 - \frac{\bar{I}\gamma_O^2\sigma_1^2}{(\gamma_I+\gamma_O)^2}\left[\gamma_O\sigma_0 + \sigma_1^2\frac{\gamma_O\bar{I}}{\gamma_O+\gamma_I}\right]^{-1}$. The term on the right hand side is minimized with respect to γ_o when $\gamma_o = \gamma_I\sqrt{1+\tilde{\Lambda}}$, with $\tilde{\Lambda} = \bar{I}\sigma_1^2/\sigma_0\gamma_I$. At the optimal γ_O, the equality becomes $E = 2/(1+\sqrt{1+\tilde{\Lambda}}) \equiv E_{opt}$. As σ_1 increases, $\tilde{\Lambda}$ becomes large which is desirable for high fidelity signal transduction. As long as $R(I)$ is approximately linear in the vicinity of $\bar{I}$, the corrections σ_n (or Ω_n) for $n > 2$ are negligible, and E is close to E_{opt}. The coefficients σ_n for $n > 2$ must be non-negligible when σ_1 is sufficiently large. Such a highly sigmoidal input-output response, known as ultra-sensitivity [1], is biologically realizable in certain regimes of signaling cascades. In the limit of a nearly step-like response, non-linearity in $R(I)$ becomes appreciable around $\bar{I}$, distorting the output signal and leading to E that is larger than E_{opt}. Because E increases with $\tilde{\Lambda}$ in this limit, the benefits of ultra-sensitivity vanish.

B. Signaling Cascades:

A natural extension is to consider a cascade created by an array of connected push-pull networks. Indeed, in some biological signaling pathways external perturbation is transmitted through a cascade of reactions involving successive activation by kinases and deactivation by phosphatases. An example is the stimulation of a receptor tyrosin kinase by epidermal growth factor, which results in downstream responses of the MAPK network [27, 42].

Because sections B, C and D are related, we explain briefly the results in order to ensure that the relationship between these sections are clear. In this section we describe the two cascade network using the field theory framework, and the coarse-graining procedure needed for obtaining an analytic expression for optimal error. In section C, we show that the two cascade network behaves as noise filter with a time delay, α^{-1}. By mapping the cascade to a push-pull network with an intermediate, we show in section D that α can be exactly calculated. Thus, the results in the three sections provide an analytic theory for optimal signaling in the two cascade network.

Consider a two step series enzymatic cascade (Fig.(3)) modeled as a sequence of two enzymatic push-pull loops stimulated by an upstream enzyme. In the first loop, an upstream enzyme, K phosphorylates the substrate, S, to produce S^*, converting it from an inactive to active state. Phosphatase (P) dephosphorylates S^* to an inactive state S. In the second loop, S^* acts as the enzyme for the phosphorylation of T and P, the corresponding phosphatases. The series of chemical reactions involved in this cascade are,

$$\phi \underset{\gamma_k}{\overset{F}{\rightleftharpoons}} K$$

$$K + S \underset{k_{1u}}{\overset{k_{1b}}{\rightleftharpoons}} S_K \xrightarrow{k_{1r}} K + S^*; \ \ S^* + P \underset{\rho_{1u}}{\overset{\rho_{1b}}{\rightleftharpoons}} S_P^* \xrightarrow{\rho_{2r}} S + P$$

$$S^* + T \underset{k_{2u}}{\overset{k_{2b}}{\rightleftharpoons}} S_T^* \xrightarrow{k_{2r}} S^* + T^*; \ \ T^* + P \underset{\rho_{2u}}{\overset{\rho_{2b}}{\rightleftharpoons}} T_P^* \xrightarrow{\rho_{2r}} T + P \tag{37}$$

where S_K, S_P^*, S_T^* and T_P^* are the reaction intermediates, and k_{ib}, k_iu, k_{ir}, ρ_{ib}, ρ_iu and ρ_{ir} , $i = 1, 2$, are the rate constants of the stochastic biochemical reactions in the cascade. The input signal $K + S_K$ is transduced into the active substrate output $T^* + T_P^*$. In an insightful article [42], a deterministic approach was used to analyze the system of chemical reactions in Eq.(37). Here we assume that the reactions are stochastic. In order to develop analytical results we only consider fluctuations of all species that deviate linearly from their mean

14

values. The validity of the asumption is established by comparing the results with kinetic Monte Carlo (KMC) simulations.

For the network in Fig.(3), the procedure outlined earlier leads to a Schrödinger-like equation with the following Hamiltonian,

$$
\begin{aligned}
\mathcal{H} =& -F(\bar{K}-1) - \gamma_k(1-\bar{K})K - k_{1b}(\bar{S}_K - \bar{K}\bar{S})KS - k_{1u}(\bar{K}\bar{S} - \bar{S}_K)S_K \\
& -k_{1r}(\bar{S}^*\bar{K} - \bar{S}_K)S_K - \rho_{1u}(\bar{S}^*\bar{P} - \bar{S}_P^*)S_P^* - \rho_{1b}(\bar{S}_P^* - \bar{S}^*\bar{P})S^*P \\
& -\rho_{1r}(\bar{P}\bar{S} - \bar{S}_P^*)S_P^* - k_{2b}(\bar{S}_T^* - \bar{S}^*\bar{T})TS^* - k_{2u}(\bar{S}^*\bar{T} - \bar{S}_T^*)S_T^* \\
& -k_{2r}(\bar{S}^*\bar{T}^* - \bar{S}_T^*)S_T^* - \rho_{2u}(\bar{T}^*\bar{P} - \bar{T}_P^*)T_P^* - \rho_{2b}(\bar{T}_P^* - \bar{T}^*\bar{P})T_P^* - \rho_{2r}(\bar{P}\bar{T} - \bar{T}_P^*)T_P^*.
\end{aligned}
$$

$$(38)$$

We can approximately map the two-step cascade into a two-species coarse-grained network, which acts like a noise filter, as described in detail in Ref. [18]. Consider a signaling pathway (Fig.(1)) with time varying input $I(t)$ and time varying output $O(t)$. These are the total populations (free and bound) of the input and output active kinases, with $I = K + S_K$ and $O = T^* + T_P^*$. The upstream pathway provides an effective production rate F of input I, while the output O results from the reaction $I \xrightarrow{R(I)} I + O$. As before, γ_I and γ_O are the degradation rates for the input and output respectively, mimicking the role of phosphatase. The input and output correlation functions, evaluated using the field theory formalism, have the approximate structure,

$$
P_I(w) = \frac{2F\gamma_I^{-2}}{1 + (w/\gamma_I)^2}
$$

$$
P_O(w) = \frac{(R_1/\gamma_o G)^2}{1 + (w/\gamma_0)^2} \left[G^2 P_I(w) + \frac{2F(G/\gamma_I)^2}{\Lambda} \right] \tag{39}
$$

where we have used a linear approximation for $R(I) \approx R_0\bar{I} + R_1(I - \bar{I})$ with $R_0, R_1 >$. Optimality is achieved when $\gamma_O = \gamma_I\sqrt{1+\Lambda}$ with gain $G = R_1/(\gamma_I(\sqrt{1+\Lambda}-1))$. Relative error with the minimum $E_{WK} = 2/(1+\sqrt{1+\Lambda})$. As before, the fidelity between output and input is controlled by single dimensionless control parameter $\Lambda = (R_1/\gamma_I)(R_1/R_0)^2$. This mapping allows us to use the general WK result for gain (G) and the minimum relative error (E_{WK}) to predict the optimality condition, allowing us to calculate the minimum possible value of E. The results for the error in terms of the mean number of phosphatase are given by the red lines in Fig.(4).

In order to test the accuracy of our theory we simulated the dynamics of the enzymatic cascade using the KMC method. The relative error E shown in Fig.(4) is in excellent

15

agreement with the theoretical predictions. Interestingly, E achieves a minimum at $\bar{P} = 10^5$ molecules/cell, which is ten times larger than the phosphatase concentration in the one step enzymatic push-pull loop using similar parameters. Fig.(4) shows that there is a well defined narrow range of phosphatase population in which the error is minimum. The range decreases as Λ decreases (Fig.(4)). The minimum value for the relative error does not reach the value predicted by the WK limit (Eq.(5)). As we show below, the additional error arises from an effective time delay as the signal passes from one cascade to another. We also demonstrate that the time delay can alternatively be mimicked by reducing the two cascade system to a coarse-grained pathway with an intermediate (Fig.(3b)).

C. Noise filtering with time delay:

In order to prove that the two-cascade loop effectively acts like a noise filter with time delay, we derive the condition for minimum error for the latter following the Bode-Shannon formulation of the WK theory [11]. In this scenario, the transmitted signal can only be recovered after a constant delay, α. The output $O(t)$ is produced by convolving the corrupted signal (input $GI(t)+$ noise $n(t)$) with a causal filter $H(t)$. In Fourier space, we obtain,

$$O(w) = H(w)c(w) = H(w)(GI(w) + n(w)) \tag{40}$$

where $x(w) = \int_{-\infty}^{\infty} dw\ x(t)e^{-iwt}$ for the time series $x(t)$. The relative error is given by [11],

$$E = \frac{\int_{-\infty}^{\infty} \frac{dw}{2\pi}[|H(w)|^2 P_n(w) + |H(w) - 1|^2 P_I(w)]}{\int_{-\infty}^{\infty} \frac{dw}{2\pi} P_I(w)} \tag{41}$$

where $P_I(w)$ and $P_n(w)$ are the power spectral densities (PSDs) of $GI(t)$ and $n(t)$ respectively. We need to minimize E in Eq.(41) over all possible $H(w)$, with the condition that $H(t) = 0$ for $t < \alpha$. The optimal causal filter has the following form [9, 11, 22],

$$H_{WK}(w) = \frac{e^{iw\alpha}}{P_c^y(w)} \left\{ \frac{P_I(w)e^{-iw\alpha}}{P_c^y(w^*)} \right\}_y. \tag{42}$$

The y super and subscript refer to two different decompositions in the frequency domain. Causality can be enforced by noting the following conditions: (i) Any physical PSD, in this case $P_c(w)$ corresponding to the corrupted signal $c(t) = GI(t) + n(t)$, can be written as $P_c(w) = |P_c^y(w)|^2$. The factor $P_c^y(w)$, if treated as a function in the complex w plane, does not have zeros and poles in the upper half-plane (Im $w \rangle 0$). (ii) We also define an additive

decomposition denoted by $\{F(w)\}_y$ for any function $F(w)$, which consists of all terms in the partial fraction expansion of $F(w)$ with no poles in the upper half-plane. By using the PSDs, $P_I(w) = \frac{2G^2\gamma_I \bar{I}}{w^2+\gamma_I^2}$ and $P_c(w) = \frac{2G^2\gamma_I \bar{I}}{w^2+\gamma_I^2} + \frac{2G^2}{\gamma_I \Lambda}$, we obtain the following optimal filter $H_{WK}(w)$,

$$H_{WK}(w) = \frac{e^{\alpha(iw-\gamma_I)}\gamma_I(\sqrt{1+\Lambda}-1)}{\gamma_I(\sqrt{1+\Lambda}-iw)}. \tag{43}$$

In the limit $\alpha \ll \gamma_I^{-1}$, the optimal error E_{WK} takes the following form [22],

$$E_{WK} = \frac{2}{1+\sqrt{1+\Lambda}} + \frac{2\Lambda\alpha\gamma_I}{(1+\sqrt{1+\Lambda})^2}, \tag{44}$$

where the second term in the above equation is the correction due to the time delay to the WK minimum value of the relative error for an instantaneous filter ($\alpha \to 0$). The correction is positive for all values of α and Λ, which implies that time delay must increase the error in signal transmission. If we add this correction to the WK minimum result for the relative error of instantaneous filter (Eq.(5)), for specific values of α calculated explicitly in the following section, we recover the minimum relative error in the signaling cascade. Thus, the two step enzymatic cascade minimizes the noise but behaves like a single step network with a time delayed filter.

D. Deriving the time delay α by mapping onto a three-species pathway with an intermediate:

Alternatively, we can derive an explicit expression for the delay parameter α by using a different mapping for the original cascade. Instead of mapping onto a two-species network of I and O with a time delay, we map onto a three-species network (Fig.(3b)) with I, M, and O. Here there is no explicit time delay, but an additional species M that will play the role of a "pseudo" intermediate mimicking the effect of the time delay. This network is governed by the reactions: $\phi \xrightarrow{F} I$, $I \xrightarrow{R_a(I)} I + M$, $M \xrightarrow{R_b(M)} M + O$, $I \xrightarrow{\gamma_I} \phi$, $M \xrightarrow{\gamma_M} \phi$ and $O \xrightarrow{\gamma_O} \phi$. The production functions have the linear form: $R_a(I) = \sigma_{a0} + \sigma_{a1}(I - \bar{I})$ and $R_b(M) = \sigma_{b0} + \sigma_{b1}(M - \bar{M})$. Earlier analysis of this network [22] has shown that it behaves like a time delayed filter, with the minimal error in the same form as Eq.(44), with $\alpha = \gamma_M^{-1}$ and effective $\Lambda = \Lambda_b\sqrt{1+\Lambda_a}$, where $\Lambda_a = \bar{I}\sigma_{a1}^2/\sigma_{ao}\gamma_I$ and $\Lambda_b = \bar{M}\sigma_{b1}^2/\sigma_{bo}\gamma_M$.

The original signaling cascade (Fig.(3a)) can be mapped onto the three-species pathway (Fig.(3b)). This involves identifying the population $S^* + S_P^* = M$ as a "pseudo" intermediate,

with an effective degradation γ_M. The mapping can be carried out by comparing PSDs between the two models. For the three-species network these are given by,

$$P_{\delta I}(w) = \frac{2\gamma_I \bar{I}}{w^2 + \gamma_I^2} \tag{45}$$

$$P_{\delta O}(w) = \frac{G^{-2}\sigma_{a1}^2\sigma_{b1}^2}{(w^2 + \gamma_M^2)(w^2 + \gamma_O^2)}\left(G^2 P_{\delta I}(w) + \frac{2G^2\gamma_M \bar{M}}{\sigma_{a1}^2} + \frac{G^2(w^2 + \gamma_M^2)2\gamma_O \bar{O}}{\sigma_{a1}^2\sigma_{b1}^2}\right)$$

Now, the PSDs for signaling cascade, calculated from the Doi-Peliti formalism, are given by,

$$P_{\delta I}(w) = \frac{\sum_{i=0}^{N-1} n_{\delta I,i}w^{2i}}{1 + \sum_{i=1}^{N} d_{\delta I,i}w^{2i}} \text{ and } P_{\delta O}(w) = \frac{\sum_{i=0}^{N-1} n_{\delta O,i}w^{2i}}{1 + \sum_{i=1}^{N} d_{\delta O,i}w^{2i}}, \tag{46}$$

where the w-independent parameters $n_{\delta I,i}$, $n_{\delta O,i}$, $d_{\delta I,i}$ and $d_{\delta O,i}$ are related to the rate co-efficients in the cascade reactions (Eq.(37)). Here, $N = 7$ corresponds to the number of independent dynamical variables ($K, S_K, S^*, S_P^*, T, T_P^*$ and T^*). By mapping Eq.(46) into Eq.(45), we can extract the degradation rate of intermediate species ($S^* + S_P^*$), γ_M in terms of coefficients in Eq.(46),

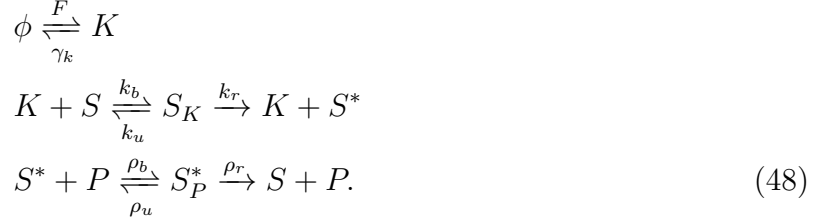$$\gamma_M^2 = \frac{1}{2}[A + \sqrt{(A^2 + 4B)}], \tag{47}$$

with $A = \frac{d_{\delta O,2}}{d_{\delta O,3}} - \gamma_I^2$ and $B = A\gamma_I^2 - \frac{d_{\delta O,1}}{d_{\delta O,3}}$. The time delay parameter $\alpha = \gamma_M^{-1}$ in the signaling cascade. With this identification for α we have a complete theory for E, with no adjustable parameter, as a function of the control parameter, the mean phosphatase levels. It is tempting to speculate that a multiple (> 2) step cascade might also be mathematically equivalent to a network with a single pseudo intermediate.

E. Enzymatic Push-Pull Loop:

In considering the cascade model, we focused on the case where fluctuations around mean populations levels were small enough that the linear approximation is valid. To study the effects of non-linearity, we will look at a simpler system (one stage of the cascade) but without any constraints on the size of the fluctuations. A microscopic model for the enzymatic push-pull network is shown in Fig.(5). The upstream enzyme, K phosphorylates a substrate S to S^*, thereby converting it from an inactive to an active state. The effective production rate in the upstream pathway for enzyme K is F. The degradation rate for K is γ_K. The enzyme is either free (K) or bound to substrate (S_K). The input I is the total enzyme population $I = K + S_K$. Phosphatase, P, on the other hand dephosphorylates the

active substrate S^* to an inactive state S. The output of the two phosphorylation cycle is $O = S^* + S_P^*$.

The biochemical reactions for the enzymatic network with the corresponding rate constants are,

$$\phi \underset{\gamma_k}{\overset{F}{\rightleftharpoons}} K$$

$$K + S \underset{k_u}{\overset{k_b}{\rightleftharpoons}} S_K \overset{k_r}{\rightarrow} K + S^*$$

$$S^* + P \underset{\rho_u}{\overset{\rho_b}{\rightleftharpoons}} S_P^* \overset{\rho_r}{\rightarrow} S + P. \tag{48}$$

In the stochastic chemical reactions that govern the phosphorylation/dephosphorylation steps, the input signal $I = K + S_K$ is transduced into the active substrate output $S^* + S_P^*$. To derive the conditions for optimality, we follow the procedure outlined in the previous section. Starting from the master equation, we can derive a Schrödinger-like equation with the following Hamiltonian,

$$\mathcal{H} = -F(\bar{K} - 1) - \gamma_k(1 - \bar{K})K - k_b(\bar{S}_K - \bar{K}\bar{S})KS$$
$$-k_u(\bar{K}\bar{S} - \bar{S}_K)S_K - k_r(\bar{S}^*\bar{K} - \bar{S}_K)S_K - \rho_u(\bar{S}^*\bar{P} - \bar{S}_P^*)S_P^*$$
$$-\rho_b(\bar{S}_P^* - \bar{S}^*\bar{P})S_P^* - \rho_r(\bar{P}\bar{S} - \bar{S}_P^*)S_P^* \tag{49}$$

After using coherent-state path integral formalism, we arrive at the expression for the action corresponding to the enzymatic push-pull loop from which we calculate the power spectra for the input and output.

As in the signaling cascade network described in the previous section, we approximately map the complete enzymatic network into a noise filter [18]. The input and output correlation functions, evaluated using field theory formalism, have the approximate structure given in Eq.(39). Starting from the full dynamical equations (Eq.(48)), we compute correlation functions using field theory by solving the Wiener-Hopf relation in Eq.(3), for the optimal function $H_{WK}(t)$.

Correlation functions of input and output calculated for enzymatic push-pull loop have the approximate form of Eq.(39), with effective values of parameters γ_I, γ_O, R_1 and Λ which have been expressed in terms of loop reaction rate parameters. This mapping allows us to use WK result for the gain (G) and minimum relative error (E_{WK}) to predict the optimality and minimum possible value of E. The results for the error in terms of the mean number of phosphatase are given by the solid lines in Fig.(6).

In order to illustrate the accuracy of the theory we performed KMC simulations by choosing the forward and backward reaction rates in Eq.(48) describing the enzymatic push-pull loop network (all units are in s^{-1}) : $k_b = \rho_b = 10^{-5}$, $k_u = 0.02$, $\rho_u = 0.5$, $k_r = 3$, $\rho_r = 0.3$, $F = 1$. The deactivation rate $\gamma_k = 0.01s^{-1}$ of enzyme K which controls the characteristic time scale over which the input signal varies, mimicking the role of phosphatase. Mean free substrate and phosphatase populations are in the ranges $\bar{S} = \bar{P} \sim 10^3 - 10^5$ molecules/cell. Fig (6) shows that E is a minimum at a particular value of phosphatase concentration $\bar{P}$, where optimality condition is satisfied i.e. $\gamma_O = \gamma_I \sqrt{1 + \Lambda}$. For a particular value of $\Lambda = 100$, we see minimum error $E = 0.18$ for the enzymatic push-pull loop. The result of the KMC simulations (purple circles) are in excellent agreement with the analytical calculation (blue line) for all Λ values.

In the parameter space used in the results in Fig.(6), a linear theory reproduces the simulation results well. However, deviations from the predictions of the linear theory are expected if the input parameters are varied. In order to investigate these deviations we first obtained the error using the parameter values, $k_b = \rho_b = 10^{-3}$, $k_u = 0.02$, $\rho_u = 0.5$, $k_r = 3$, $\rho_r = 0.3$ using KMC simulations. The relative error for $\Lambda = 100$ is shown in purple line in Fig.(7). The blue line, calculated from linear theory predictions, deviates substantially from simulations (purple line in Fig.(7)). To improve the predictions of the theory we calculated second order corrections to E. The result, displayed as green curve in Fig.(7)), shows that there is improved agreement between theory and simulations. The non-linear corrections, which are substantial, brings the theoretical predictions closer to the simulation results, especially near the values of $\bar{S}$ for which the error is a minimum (Fig.(7))). We suspect that higher order perturbative corrections will further improve the results based on the following observation. We fit the dependance of the error for ($\bar{S} > 1600$) using the function, $E(\bar{S}) = a + b(\bar{S} - \bar{S}_{\min})^{1.4} + c(\bar{S} - \bar{S}_{\min})^2$, where a, b and c are constants and $\bar{S}$ is the value of $\bar{S}$ at which $E(\bar{S}_{\min}) = a$ is a minimum. The functional form of $E(\bar{S})$ is the same for the exact simulation results, and the predictions of the linear and non-linear theory except the coefficients a, b and c are different. We, therefore, surmise that higher order terms merely renormalize the coefficients, keeping unaltered the form of relative error. Consequently, we conclude that improved estimates of a, b, c from third and higher order contributions should produce predictions in better agreement with simulation.

CONCLUDING REMARKS:

In order to assess the accuracy of signal transmission, using the mean square of the error between input and output as a fidelity measure, we have developed a field theoretic formulation that allows us to predict conditions for optimal information transfer for an arbitrary stochastic chemical reaction network. The starting point is the classical master equation for interacting particle systems, which is mapped to a non-Hermitian 'quantum' many-body Hamiltonian dynamics. Finally, the coherent-state path integral representation is utilized to arrive at a continuum field theory description that faithfully incorporates the intrinsic reaction noise and discreteness of the original stochastic processes. The formulation allows us to use standard field theory methods to compute the relative error in the information transfer using perturbation theory to all orders in non-linearity. This approach leads to an analytical expression for the minimum relative error in signal transduction. The usefulness of the general field theory formulation is illustrated through signaling networks of increasing complexity.

Detailed study of an enzymatic push pull loop, the basic unit involved in complex signaling pathways, show that it behaves like an optimal linear WK noise filter, as previously established using entirely different methods [18]. In this particular case, the joint probability $P(\delta I, \delta O)$ is approximately bivariate Gaussian, which means the error E is also directly related to the mutual information M in bits between δI and δO as $E = 2^{-2M}$ [22].

The two-stage enzymatic cascade behaves as an optimal filter without achieving the minimum predicted by the WK theory. We attribute the deviation to the time delayed response of the cascade. By mapping the cascade signaling network to a three-species push-pull like model with a pseudo intermediate state we derived an explicit expression for the time delay. We show that the time delay is associated with the degradation rate of the pseudo intermediate state in the coarse-grained representation of the two-step cascade. We also demonstrate that in those cases where the linear approximation breaks down, systematic perturbative corrections can be calculated using our theory, which minimize the difference between the findings in the simulations and theoretical predictions. The success in this example illustrates the power of the formalism. Analyzing experimental data using the framework introduced here will help decipher the design principles governing signaling networks in biology, and allow us to understand the constraints imposed by noise in information transfer.

Appendix A: Second order loop correction to the signaling error for the push-pull network:

Here, we illustrate the calculation of E arising from perturbation expansion of the field theory for the push-pull network with non-linearity explained in the text. To second order the diagram needed to compute E is,
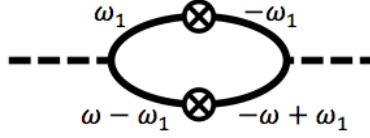
$$
\begin{aligned}
E_2 &= 2(2\gamma_I \bar{I})^2 \int \frac{dw}{2\pi} \int \frac{dw_1}{2\pi} \frac{1}{(w_1^2 + \gamma_I^2)(w^2 + \gamma_O^2)(w_1 - w)^2 + \gamma_I^2)} \\
&= 2(2\gamma_I \bar{I})^2 \int \frac{dw}{2\pi} \frac{1}{(w^2 + \gamma_O^2)} \frac{2\pi i}{2\pi} \frac{1}{2i\gamma_I} \left[\frac{1}{w(w - 2i\gamma_I)} + \frac{1}{w(w + 2i\gamma_I)} \right] \\
&= 2 \frac{(2\gamma_I \bar{I})^2}{\gamma_I} \frac{2\pi i}{2\pi} \left[\frac{1}{2i\gamma_O} - \frac{1}{4i\gamma_I} \right] \frac{1}{4\gamma_I^2 - \gamma_O^2} \\
&= \frac{2\bar{I}^2}{\gamma_O(\gamma_O + 2\gamma_I)}.
\end{aligned}
\tag{50}
$$

In the first line of the above equation, we perform the complex integration in the upper half plane by evaluating the residues at poles $w_1 = i\gamma_I$ and $w_1 = w + i\gamma_I$, respectively. Similarly, in the second line we calculate the residues at poles $w = i\gamma_O$ and $w = 2i\gamma_I$.

The coefficient, Ω_2^2 (Eq.(35)), is diagrammatically represented as,

$$
\Omega_2^2 = \frac{c_2^2}{4} + \frac{c_3^2}{4} + \frac{1}{4}c_2 c_4 \left(\bigcirc \right) + \frac{1}{24}c_2 c_6 \left(\bigcirc \right)^2 + \cdots \quad,
$$

where the expression for the loop in the first bracket is $2\gamma_I \bar{I} \int \frac{dw_1}{2\pi} \frac{1}{(w_1^2 + \gamma_I^2)} = \bar{I}$. The coefficients Ω_n in Eq.(35) are functions of c_n. In turn, Ω_ns and σ_ns are also connected by the relation between σ_n and c_n (see main text). For all n, the leading order term ($\frac{c_n^2}{n!^2}$) of Ω_n and σ_n is

23

identical. Similarly, the third order term is (third term in the Fig.(2))

$$E_3 = 6\Omega_3^2 (2\gamma_I \bar{I})^3 \int \frac{dw}{2\pi} \int \frac{dw_1}{2\pi} \frac{1}{(w_1^2 + \gamma_I^2)(w^2 + \gamma_O^2)(w_1 + w_2 - w)^2 + \gamma_I^2)(w_2^2 + \gamma_I^2)} \quad (51)$$

$$= \Omega_3^2 \frac{6\bar{I}^3}{\gamma_O(\gamma_O + 3!\gamma_I)}.$$

The generalization to the nth order term gives $E_n = \Omega_n^2 \frac{n! \bar{I}^n}{\gamma_O(\gamma_O + n!\gamma_I)}$.

Appendix B: Action for enzymatic push-pull network:

We give the form of the action here for the enzymatic push-pull network for which the chemical reaction scheme is given in Eq.(48). Despite the complexity, the action can be manipulated using Mathematica in order to obtain general expression for the error.

$$\mathcal{S}[\tilde{\Phi}, \Phi] = \int_w \bar{\phi}_K[-iw\phi_K - F + \gamma_K \phi_K + k_b \phi_K \phi_S - (k_r + k_u)\phi_{S_K}] + \bar{\phi}_{S_K}[-iwS_K - \quad (52)$$

$$k_b \phi_K \phi_S + (k_r + k_u)\phi_{S_K}] + \bar{\phi}_{S^*}[-iwS^* - k_r\phi_{S_K} - \rho_u \phi_{S_P^*} + \rho_b \phi_{S^*}\phi_P] + \bar{\phi}_S[-iw\phi_S + k_b\phi_K\phi_S$$

$$-\rho_r \phi_{S_P^*} - k_u\phi_{S_K}] + \bar{\phi}_{S_P^*}[-iw\phi_{S_P^*} + (\rho_u + \rho_r)\phi_{S_P^*} - \rho_b\phi_{S^*}\phi_P] + \bar{\phi}_P[-iw\phi_P - (\rho_u + \rho_r)\phi_{S_P^*} + \rho_b\phi_{S^*}\phi_P]$$

$$+\frac{1}{2}\bar{\phi}_K^2[-F - \gamma_K\phi_K - k_b\phi_K\phi_S - (k_r + k_u)\phi_{S_K}] + \frac{1}{2}\bar{\phi}_{S_K}^2[-k_b\phi_K\phi_S - (k_r + k_u)\phi_{S_K}] + \frac{1}{2}\bar{\phi}_{S^*}^2[-k_r\phi_{S_K}$$

$$-\rho_u\phi_{S_P^*} - \rho_b\phi_{S^*}\phi_P] + \frac{1}{2}\bar{\phi}_S^2[-k_b\phi_K\phi_S - \rho_r\phi_{S_P^*} - k_u\phi_{S_K}] + \frac{1}{2}\bar{\phi}_{S_P^*}^2[-(\rho_u + \rho_r)\phi_{S_P^*} - \rho_b\phi_{S^*}\phi_P]$$

$$+\frac{1}{2}\bar{\phi}_P^2[-(\rho_u + \rho_r)\phi_{S_P^*} - \rho_b\phi_{S^*}\phi_P -] - k_b[\bar{\phi}_{S_K}\bar{\phi}_K - \bar{\phi}_{S_K}\bar{\phi}_S + \bar{\phi}_K\bar{\phi}_S]\phi_K\phi_S - k_r[\bar{\phi}_{S^*}\bar{\phi}_K - \bar{\phi}_{S^*}\bar{\phi}_{S_K}$$

$$-\bar{\phi}_K\bar{\phi}_{S_K}]\phi_{S_K} - \rho_u[\bar{\phi}_{S^*}\bar{\phi}_P - \bar{\phi}_{S^*}\bar{\phi}_{S_P^*} - \bar{\phi}_P\bar{\phi}_{S_P^*}]\phi_{S_P^*} - \rho_r[\bar{\phi}_S\bar{\phi}_P - \bar{\phi}_{S^*}\bar{\phi}_P - \bar{\phi}_S\bar{\phi}_{S_P^*}]\phi_{S_P^*} -$$

$$\rho_b[-\bar{\phi}_{S^*}\bar{\phi}_{S_P^*} - \bar{\phi}_P\bar{\phi}_{S_P^*} + \bar{\phi}_P\bar{\phi}_{S^*}]\phi_{S^*}\phi_P - k_u[\bar{\phi}_S\bar{\phi}_K - \bar{\phi}_K\bar{\phi}_{S_K} - \bar{\phi}_S\bar{\phi}_{S_K}]\phi_{S_K}.$$

Appendix C: Network with spatial dependence:

In this article, we assume that the system is well-stirred, which amounts to neglect of spatial organization of the cells. Despite this limitation, these stochastic models provide useful insights into the signaling process. However, this is a poor approximation for certain aspects of signal transduction. For example, the cell membrane exhibits significant spatial organization, with diffusion rates approximately two orders of magnitude slower than those

in the cytosol. Thus, to unravel the complexities of signaling pathways, quantitative models must consider spatial organization as an important feature of cell signaling. In order to consider spatial dependence, we need to include the gain and loss term in the master equation, corresponding to the diffusion of input and output species. The corresponding term in the Hamiltonian for the input species is $H_{dif}^I = D\sum_{ij}(a_i^\dagger - a_j^\dagger)(a_i - a_j)$. A similar term for the output species is, $H_{dif}^O = D\sum_{ij}(b_i^\dagger - b_j^\dagger)(b_i - b_j)$. Finally, these terms will appear in the action in the form $-D\bar{\phi}_I\nabla^2\phi_I$ and $-D\bar{\psi}_O\nabla^2\psi_O$ respectively.

Appendix D: Mapping between parameters in the network and field theoretic formalism:

1). For enzymatic push pull loop, the input and output correlation functions, evaluated using the field theory formalism, have the approximate structure of Eq.(39). We extract the effective parameters γ_I, γ_O and Λ in terms of rate parameters used in the master equation. The effective parameters resulting from the mapping are,

$$\gamma_I = \frac{k_-\gamma_K}{k}, \ \gamma_O = \frac{\rho_r\rho_+}{\sqrt{\rho^2 - 2\rho_r\rho_+}}$$
$$\Lambda = \frac{k_r\kappa_+k^2\rho^2}{\gamma_K k_-[\rho^2(k^2 - k_rk_+) - \rho_r\rho_+k^2]}, \tag{53}$$

where $k_+ = k_b\bar{S}$, $k_- = k_u + k_r$, $k = k_+ + k_-$, $\rho_+ = \rho_b\bar{P}$, $\rho_- = \rho_u + \rho_r$, and $\rho = \rho_+\rho_-$.

2). For the signaling cascade, the mapping of Eq.(46) into Eq.(39) results in the following effective parameters,

$$\gamma_I = d_{\delta I,1}^{-1/2}, \ \gamma_O = d_{\delta O,1}^{-1/2} \text{ and } \Lambda = \frac{1}{\gamma_I^4\eta_{\delta O,1}}, \tag{54}$$

where, $d_{\delta O,1}$, $d_{\delta I,1}$ and $\eta_{\delta O,1}$ are complicated functions of rate parameters involved in the master equation in the signaling cascade.

[1] A. Goldbeter and D. E. Koshland. An amplified sensitivity arising from covalent modification in biological systems. *Proc. Natl. Acad. Sci.*, 78:6840, 1981.

[2] M Thattai and A van Oudenaarden. Intrinsic noise in gene regulatory networks. *Proc. Natl. Acad. Sci.*, 98(15):8614–8619, 2001.

[3] M. Thattai and A. van Oudenaarden. Attenuation of noise in ultra sensitive signaling cascades. *Biophys. J.*, 82:2943–2950, 2002.

[4] A. Eldar and M. B. Elowitz. Functional roles for noise in genetic circuits. *Nature*, 467:167–173, 2010.

[5] A. Raj and A. van Oudenaarden. Nature, nurture, or chance: Stochastic gene expression and its consequences. *Cell*, 135:216–226, 2008.

[6] N. Maheshri and E. K. O'Shea. Living with noisy genes: How cells function reliably with inherent variability in gene expression. *Ann. Rev. Biophys. Biomol. Struct.*, 36:413–434, 2007.

[7] C. G. Bowsher, M. Voliotis, and P. S. Swain. The fidelity of dynamic signaling by noisy biomolecular networks. *PLOS Computational Biology*, 9:e1002965, 2013.

[8] L. Cai, C. K. Dalal, and M. B. Elowitz. Frequency modulated nuclear localization bursts coordinate gene regulation. *Nature*, 455:485, 2008.

[9] N. B. Becker, A. Mugler, and P R. ten Wolde. Optimal prediction by cellular signaling networks. *Phys. Rev. Letts.*, 115:258103, 2015.

[10] I. Lestas, G. Vinnicombe, and J. Paulsson. Fundamental limits on the suppression of molecular fluctuations. *Nature*, 467:174–178, 2010.

[11] H. W. Bode and C. E. Shannon. A simplified derivation of linear least square smoothing and prediction theory. *Proc. IRE*, 38:417, 1950.

[12] W. H. de Ronde, F. Tostevin, and P. R. ten Wolde. Effect of feedback on the fidelity of information transmission of time-varying signals. *Phys. Rev. E*, 82:031914, 2010.

[13] E Ziv, I. Nemenman, and C. H. Wiggins. Optimal signal processing in small stochastic biochemical networks. *PLoS ONE*, 2(10): e1077, 2007.

[14] G. Tkacik, C. G. C. Jr., and W. Bialek. Information capacity of genetic regulatory elements. *Phys. Rev. E*, 78:011910, 2008.

[15] A. M. Walczak, A. Mugler, and C. H. Wiggins. A stochastic spectral analysis of transcriptional regulatory cascades. *Proc. Natl. Acad. Sci.*, 106:6529–6534, 2009.

[16] S. Mehta, H. Li, P. G. Hogan, and K. W. Cunningham. Domain architecture of the regulators of calcineurin (rcans) and identification of a divergent rcan in yeast. *Mol. Cell. Biol.*, 29(10):2777–2793, 2009.

[17] A. Mugler, A. M. Walczak, and C. H. Wiggins. Spectral solutions to stochastic models of gene expression with bursts and regulation. *Phys. Rev. E*, 80:041921, Oct 2009.

[18] M. Hinczewski and D. Thirumalai. Cellular signaling networks function as generalized wiener-kolmogorov filters to suppress noise. *Phys. Rev. X*, 4:041017, 2014.

[19] N. Wiener. *Extrapolation, Interpolation and Smoothing of Stationary Times Series*. Wiley, New York, 1949.

[20] A. N. Kolmogorov. Interpolation and extrapolation of stationary random sequences. *Izv. Akad. Nauk SSSR, Ser. Mat.*, 5:3–14, 1941.

[21] M. Hinczewski and D. Thirumalai. Noise control in gene regulatory networks with negative feedback. *J. Phys. Chem. B*, 120:6166, 2016.

[22] D. Hathcock, J. Sheehy, C. Weisenberger, E. Ilker, and M Hinczewski. Noise filtering and prediction in biological signaling networks. *IEEE Trans. Mol. Biol. Multi-Scale Commun.*, 2:16, 2016.

[23] S. Roman. *The Umbral Calculus.* Dover, New York, 2005.

[24] M. Doi. *J. Phys. A: Math. Gen.*, 9:1465, 1976.

[25] L. Peliti. *J. Phys. (France)*, 46:1469, 1985.

[26] K. P. Silva, P. Chellamuthu, and J. Q. Boedicker. Signal destruction tunes the zone of activation in spatially distributed signaling networks. *Biophys. J.*, 112:1037–1044, 2017.

[27] B. Schoeberl, C. Eichler-Jonsson, E. D. Gilles, and G. Muller. Computational modeling of the dynamics of the map kinase cascade activated by surface and internalized egf receptors. *Nat Biotech*, 20(4):370–375, 04 2002.

[28] P. Hersen, M. N. McClean, L. Mahadevan, and S. Ramanathan. Signal processing by the hog map kinase pathway. *Proc. Natl. Acad. Sci.*, 105(20):7165–7170, 2008.

[29] J. Levine, H. Y. Kueh, and L. Mirny. Intrinsic fluctuations, robustness, and tunability in signaling cycles. *Biophys. J.*, 92:4473–4481, 2007.

[30] E R Stadtman and P B Chock. Superiority of interconvertible enzyme cascades in metabolic regulation: analysis of monocyclic systems. *Proc. Natl. Acad. Sci.*, 74:2761–2765, 1977.

[31] P. B. Detwiler, S. Ramanathan, A. Sengupta, and B. I. Shraiman. Engineering aspects of enzymatic signal transduction: Photoreceptors in the retina. *Biophys. J.*, 79:2801–2817, 2000.

[32] R. Heinrich, B. G. Neel, and T. A. Rapoport. Mathematical models of protein kinase signal transduction. *Molecular Cell*, 9:957–970, 2002.

[33] W. Kolch, M. Halasz, M. Granovskaya, and B. N. Kholodenko. The dynamic control of signal transduction networks in cancer cells. *Nat Rev Cancer*, 15:515–527, 2015.

[34] B. P. Lee and J. Cardy. Renormalization group study of the a+b→0 diffusion-limited reaction. *J. Stat. Phys.*, 80:971, 1995.

[35] J. L. Cardy and Uwe C. Täuber. Field theory of branching and annihilating random walks. *J. Stat. Phys.*, 90:1, 1998.

[36] J. W. Negele and H. Orland. *Quantum many-particle systems*. Redwood City: Addison-Wesley, 1988.

[37] Uwe C. Täuber. *Critical dynamics*. Chembridge university press, 2014.

[38] De Dominicis. Techniques de renormalization de la theorie des champs et dynamique des phenomenes critiques. *J. Phys. (France) Colloq.*, C1:C247–C253, 1976.

[39] H. K. Janssen. On a lagrangian for classical field dynamics and renormalization group calculations of dynamical critical properties. *Z. Phys. B*, 23:377–380, 1976.

[40] J. D. Cole. On a quasi-linear paribolic equation occurring in aerodynamics. *Quart. Appl. Math.*, 9:225–236, 1951.

[41] E. Hopf. The partial differential equation u t + uu x = u xx. *Comm. Pure Appl. Math.*, 3:201–230, 1950.

[42] R. Heinrich, B. G. Neel, and T. A. Rapoport. Mathematical models of protein kinase signal transduction. *Mol. Cell*, 9:957, 2002.
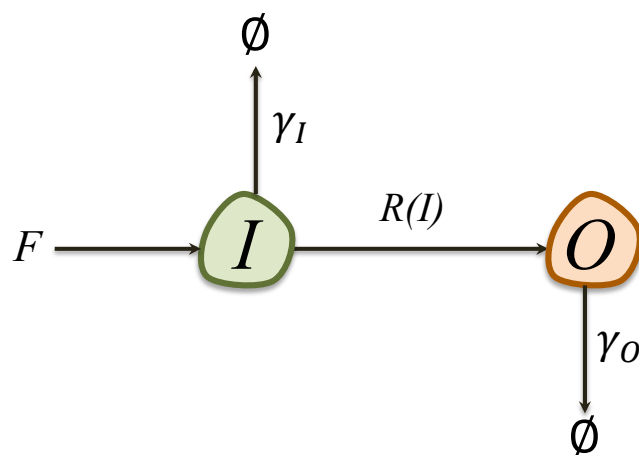
FIG. 1. Schematic of a push-pull network, involving an input species I and output species O. The production of O from I is controlled by the rate function $R(I)$. The degradation rates for I and O are γ_I and γ_O, respectively.

FIG. 2. Examples of diagrams for the correlation function $\langle \delta\psi_O(-w)\delta\psi_O(w)\rangle$. The Ω_ns are coefficients with, $\Omega_1^2 = c_1^2$, $\Omega_2^2 = \frac{c_2^2}{4} + \frac{c_3^2}{4} + \frac{\bar{I}}{4}c_2c_4 + \cdots$, $\Omega_3^2 = \frac{c_3^2}{36} + \frac{c_4^2}{16} + \frac{\bar{I}}{36}c_3c_5 + \cdots$ and so on.
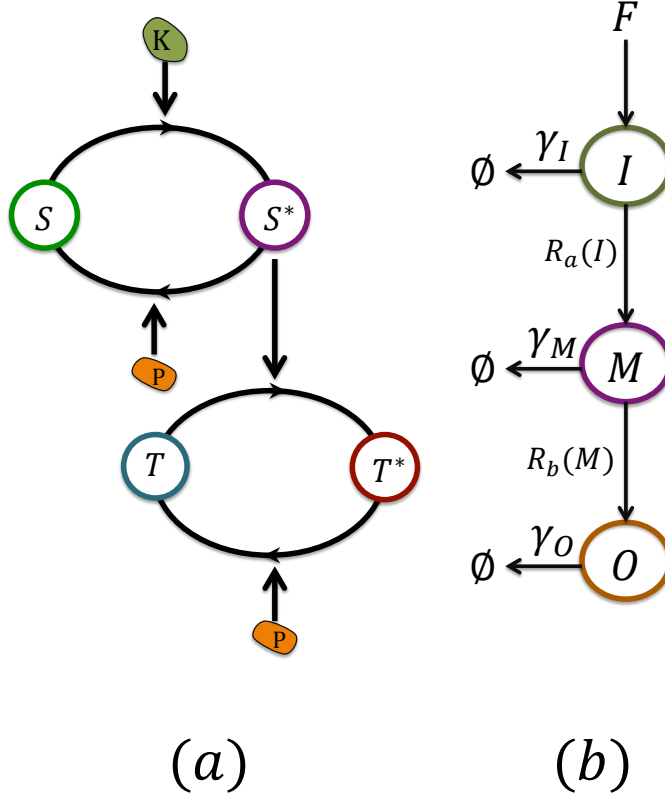
FIG. 3. (a) Enzymatic cascade that arises naturally in mitogen activated protein kinase (MAPK) networks. In a caricature of such a network, kinase (K) phosphorylates the substrate (S), leading to the formation of S^*. Deactivation is triggered by reactions with the phosphatase (P). S^* phosphorylates the substrate T, producing T^* and P reverts it to the original form through dephosphorylation. The rate parameters in the chemical reactions (Eq. (37)) used to produce numerical results (in units of s^{-1}) are : $k_{1b} = k_{2b} = \rho_{1b} = \rho_{2b} = 10^{-5}$, $k_{1u} = 0.02$, $k_{2u} = 0.3$, $\rho_{1u} = 0.5$, $\rho_{2u} = 1.0$, $k_{1r} = 3$, $k_{2r} = 5.0$, $\rho_{1r} = 0.3$, $\rho_{2r} = 0.1$, $F = 1$. The deactivation rate $\gamma_k = 0.01s^{-1}$ controlling the characteristic time scale over which the input signal varies. Mean free substrate and phosphatase populations are in the ranges $\bar{S} = \bar{P} \sim 10^4 - 10^6$ molecules/cell. (b) Three species coarse-grained signaling network with the indicated rates is intended to capture the physics of the cascade in (a). The mathematical equivalence between the networks in (a) and (b) is established in the text.
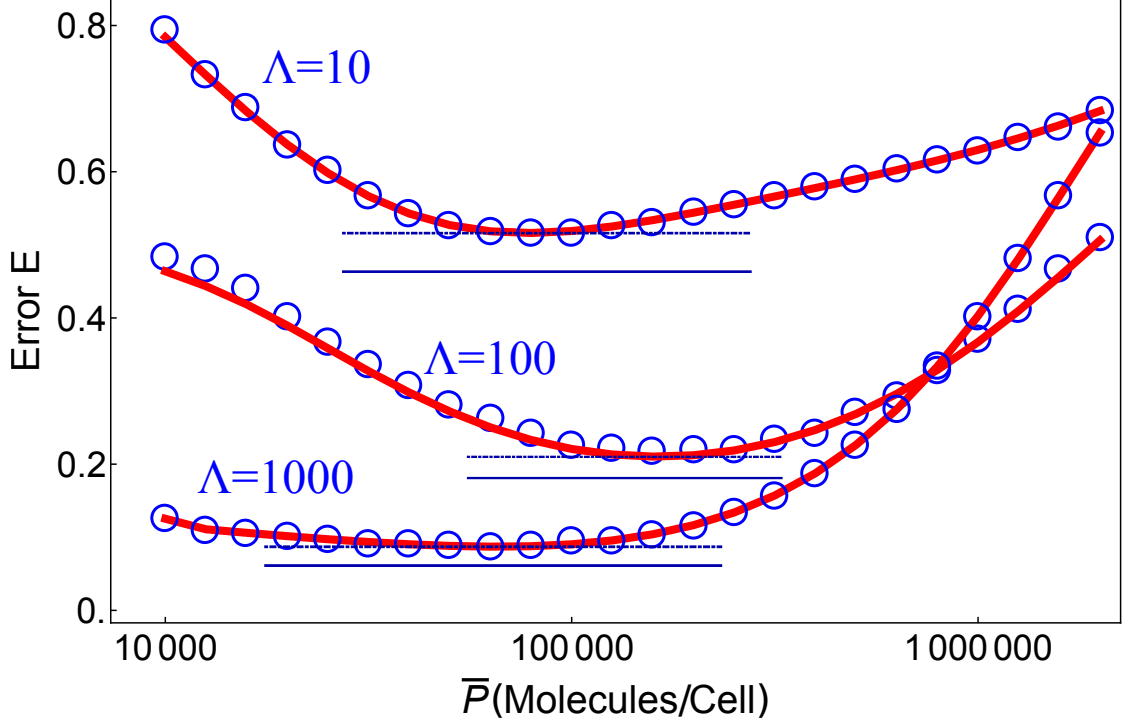
FIG. 4. Relative error E for the signaling cascade (red lines show the theoretical predictons; blue circles are obtained using the kinetic Monte Carlo simulations) for three Λ values. The blue dashed line gives the predictions (Eq.(44)) using the WK formalism for E_{WK} with time delay. The solid blue line is the minimal error corresponding to the theory without time delay in Eq.(5). The comparison shows that the two-loop cascade behaves as a push-pull network with a time delay. The time delay parameter, α, is explicitly given in Eq.(47). Thus, the theory has no adjustable parameter.
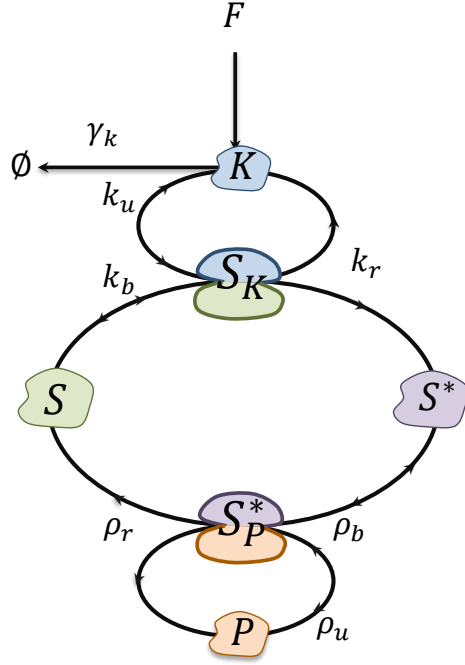
FIG. 5. Enzymatic push-pull loop showing phosphorylation of the substrate (S) by kinase (K) to produce the active form S^*. Phosphatase (P) reverts it to the original form through dephosphorylation. S_K and S_P^* represent the substrate in the complex with kinase and phosphatase, respectively. Binding, unbinding and the reaction rate constants are shown with arrows.
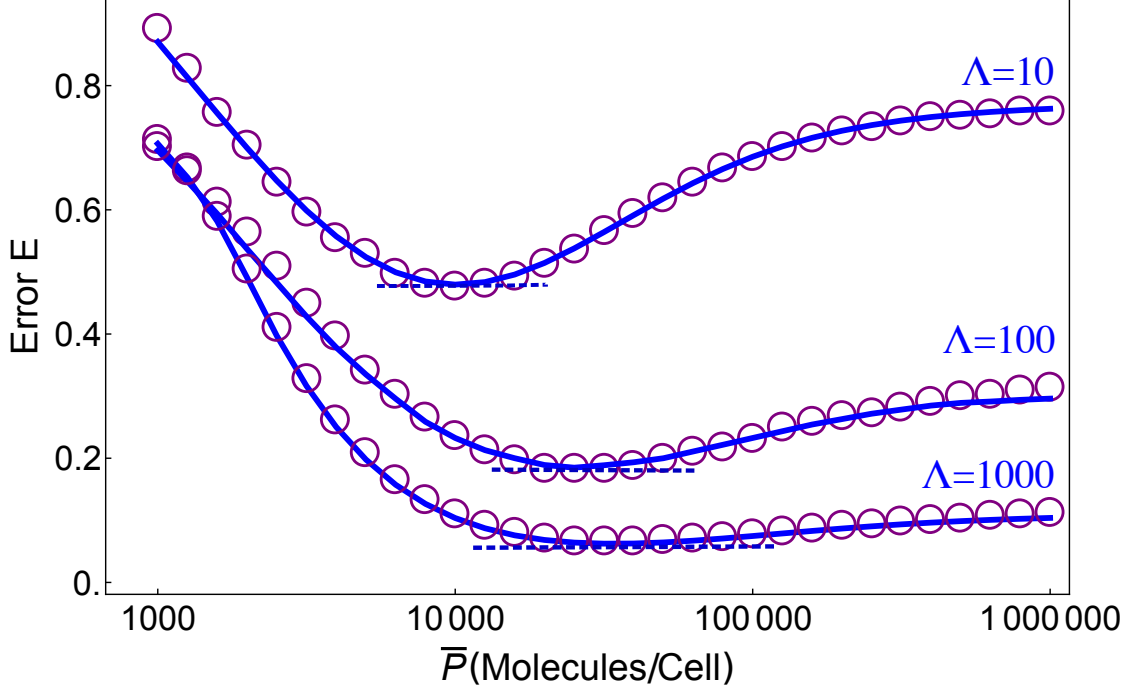
FIG. 6. Relative error E for the enzymatic push-pull loop in Fig.(5). The blue lines correspond to theoretical predictions. The KMC simulation results are given in purple circles. The dashed line is the minimal error corresponding to the WK theory (Eq.(5)). The values of the rates corresponding to the chemical reactions in the enzymatic push-pull network (Eq.(48)) is given in the main text. For the parameter values the predictions of the linear theory are very accurate.
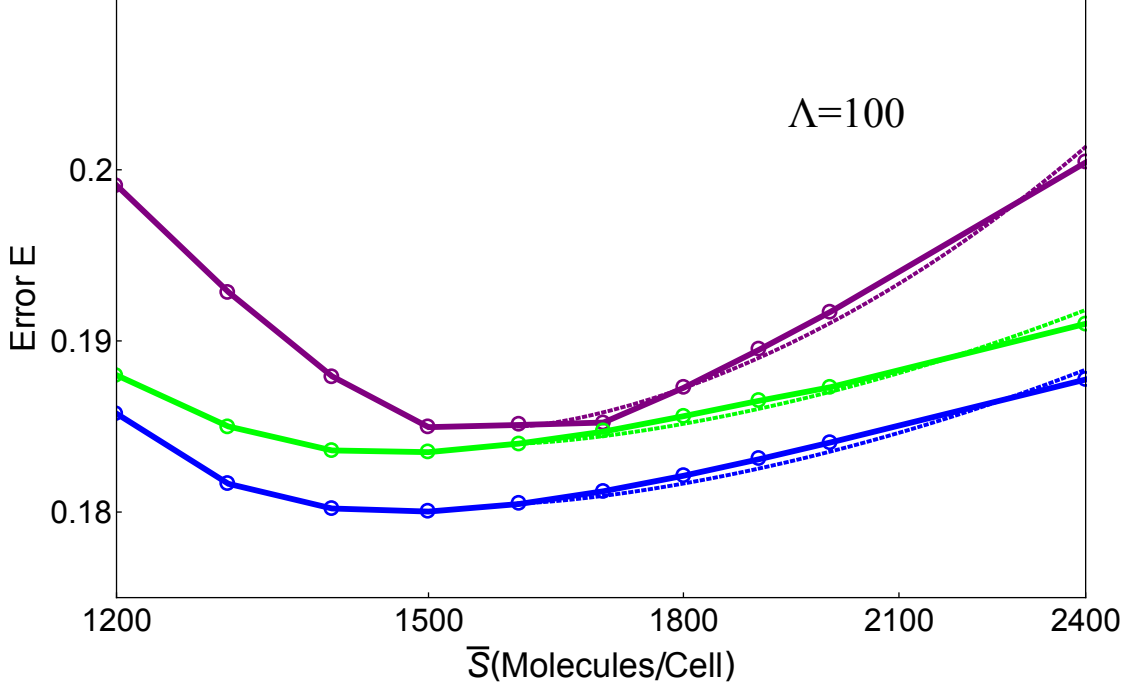
FIG. 7. Error E for the enzymatic push-pull loop for different values of $\bar{S}$ with $\Lambda = 100$. The rate parameters used in Eq.(48) are $k_b = \rho_b = 10^{-3}$, $k_u = 0.02$, $\rho_u = 0.5$, $k_r = 3$ and $\rho_r = 0.3$. The blue line is the result calculated using linear theory. The green line results from second order corrections to the error E. The KMC simulation results are given in purple line. Clearly inclusion of non-linear corrections improves the predictions of the theory in the range of $\bar{S}$ values for which E is small. Dotted lines are fit with the function $a + b(\bar{S} - 1600)^{1.4} + c(\bar{S} - 1600)^2$ where a, b, c are constants. For all the curves a, b, c values change but the functional form of E as a function of $\bar{S}$ is the same. It is likely that if the theory is extended beyond second order, there should be further improvement by bringing a, b and c values closer to the simulation results.