

Information propagation in time through allosteric signaling

Tushar Modi^{*} and S. Banu Ozkan[†]

Department of Physics, Arizona State University, Tempe, Arizona 85287, USA

Steve Pressé[‡]

*Department of Physics, Arizona State University, Tempe, Arizona 85287, USA
and School of Molecular Sciences, Arizona State University, Tempe, Arizona 85287, USA*



(Received 24 June 2019; accepted 20 April 2020; published 19 June 2020)

Naively, one expects the information communicated by an enzyme downstream within a signaling network, in which the enzyme is embedded, to grow monotonically with the enzyme's rate of product formation. However, here we observe that this does not necessarily hold true for allosterically regulated enzymes, often observed in signaling networks. In particular, we find that the mutual information between the catalytic sites of an allosterically regulated enzyme and a receiver protein downstream in the signaling pathway depends on the transition kinetics between the different allosterically regulated states of the enzyme and their respective rates of product formation. Surprisingly, our work implies that allosteric down-regulation of an enzyme's rate of product formation may not only be used as a way to silence itself, as one would normally expect. Rather, down-regulation may also be used to increase the information communicated by this enzyme to a receiver protein downstream in a signaling pathway.

DOI: [10.1103/PhysRevResearch.2.023367](https://doi.org/10.1103/PhysRevResearch.2.023367)

I. INTRODUCTION

Despite the fact that only a small fraction of a cell is composed of proteins (e.g., proteins constitute 17% of *E. Coli* [1]), proteins not only mediate the key processes in cells, but they also give rise to spatiotemporal signaling to control a cell's response to its local environment underlying all critical decision-making [2] such as a cell's development and metabolism [3], motility [4], immunity, and cell-death (apoptosis) [5].

These spatiotemporally coordinated events are often achieved by proteins exhibiting allostery—a phenomenon by which the binding of a molecule at one site of a protein changes the binding affinity or catalytic activity at another distant site. Several models of allostery have previously been explored to study how allosteric interactions within one enzyme (i.e., proteins that act as catalysts) modulate its activity [6–8]. However, allostery goes beyond just remote modulation. It also provides “circuit components” from which nature builds up complex signaling networks. Here we employ an information theoretic framework to quantify how allostery propagates through protein signaling pathways in order, downstream, to modulate cellular response.

Information theory has already been proven to be useful in studying various biological phenomena such as modeling protein interaction networks [9–11], evolution [12–14], and single molecule experiments [15,16]. Moreover, it has also been instrumental in studying the effect of allostery on the transfer of information between the allosteric regulator and the catalytic site of allosteric proteins, particularly enzymes [17,18]. However, these studies primarily explored the thermodynamics of substrate binding through allostery, focusing solely on the concentration of products generated. On the other hand, the activity of the receiver protein varies in time according to the availability and on/off binding of the product produced by the enzyme. Therefore, information conveyed is encoded not only in the total concentration of the enzymatic product generated by “sender enzymes,” but also in its time-varying activity. Here, we apply information theory to quantify the role of allostery in a sender enzyme while it communicates with a receiver protein (a protein to which product molecules may bind).

To apply information theory, we first consider a simple master equation model capturing the essence of allostery in the arrival of products at the receiver protein. Next, this model is simplified to a two-state model that is studied in a discrete time domain with the help of hidden Markov models (HMMs) [19,20]. We compute the probability of observing product arrival events accounting for all possible latent allosteric states (“hidden trajectories”) of the sender enzyme. We then calculate the mutual information (MI) [21] over the joint distribution of the state of the sender enzyme and the state of the receiver protein. The joint distribution encodes both the allosterically induced state-switching transitions and the product formations by the sender enzyme's catalytic site.

Our work illustrates how allostery directly impacts the transfer of information within signaling pathways. It shows

^{*}tmodi@asu.edu

[†]Corresponding author: Banu.Ozkan@asu.edu

[‡]Corresponding author: spress@asu.edu

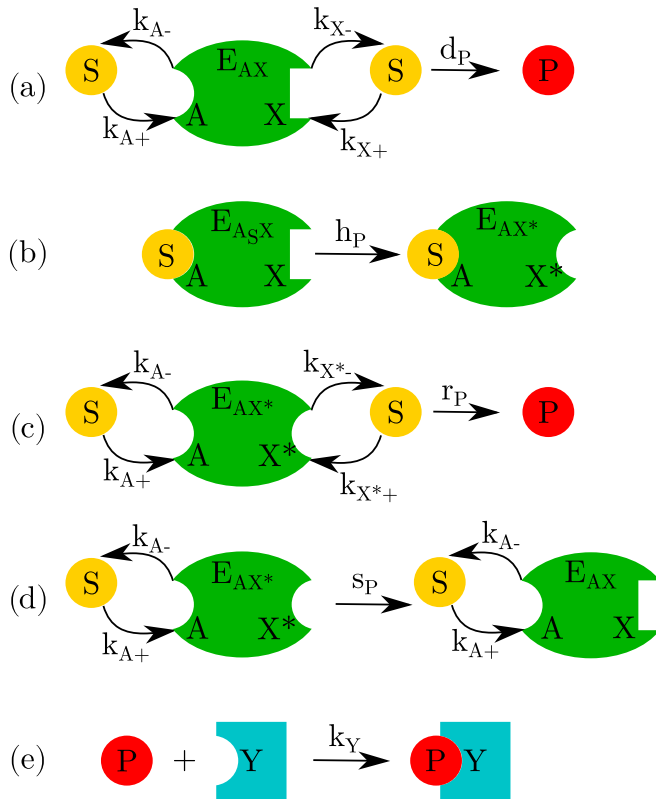


FIG. 1. Schematic diagram of the model of an allosterically regulated sender enzyme. It shows E_{AX} , a sender enzyme, with catalytic site X interacting with substrates S to make product P and allosteric regulator site A , which can interact with the substrates to change the state of the sender enzyme to E_{AX^*} with catalytic site X^* . The catalytic site X^* can then also interact with the substrate to make products P , however with different rate constants. The products generated then bind to a receiver protein Y downstream in the signaling pathway.

that the communication by a sender enzyme in a signaling pathway is not merely modulated by the number of products generated, but also significantly depends on the time signature of product arrivals at other receiver proteins present downstream. The analysis also suggests a broader role for allosteric regulation: a way to increase the information communicated within signaling pathways even, counterintuitively, via down-regulation.

II. MODEL

We start with a minimal model of a sender enzyme with an allosteric regulator site (Fig. 1) inspired by the KNF (Koshland-Nmethyl-Filmer) model of protein allosteric regulation as described in Refs. [8]. The model features a sender enzyme, E_{AX} , which functions as a sender of information to another protein downstream in the signaling pathway (a receiver protein) in the form of products. It has a catalytic site X that can bind or unbind to a substrate S to make complex E_{AXS} , with rate constants k_{X+} and k_{X-} , respectively. In addition, the sender enzyme also contains of an allosteric regulator site A . The sender enzyme with the bound complex at the catalytic site subsequently degrades to generate product,

P , with a rate constant of d_P releasing its catalytic site X back to its original unbound form.

When the allosteric regulator site A reversibly binds to substrate, it creates a complex E_{ASX} with the rate constants k_{A+} and k_{A-} , respectively; see Fig. 1(a). The bound complex at the allosteric regulator can exploit the sender enzyme's network of interactions to influence the activity of the catalytic site [see Fig. 1(b)]. This allosteric interaction occurs between a bound allosteric regulator site and an unbound catalytic site with a rate constant of h_P . The modified catalytic site (X^*) now performs its functions with different activity such that its rate constants for binding and unbinding with the substrates change to k_{X^*+} and k_{X^*-} , respectively. Moreover, the rate constant of the degradation of the complex at X^* with substrate to create product also changes to r_P ; see Fig. 1(c). It should be noted that these changes are only localized to the catalytic site, whereas the allosteric regulator does not show any changes in its dynamics of interaction, i.e., the rate constants of binding and unbinding of substrate at the allosteric regulator site do not change as the catalytic site alters its state.

Finally, the catalytic site (X) after being allosterically modified to (X^*) can relax back to its original state in a process with a rate constant of s_P . This process can occur regardless of the state of the allosteric regulator (i.e., whether it is bound or unbound); see Fig. 1(d).

The products P produced in processes (a) and (c) in Fig. 1 can interact with the receiver protein Y downstream in the signaling pathway with a rate constant of k_Y . This protein acts as a receiver for the signal generated by the sender enzyme in the signaling pathway in the form of products. During the interaction of product P with the receiver Y , a binding event signifies a successful transfer of signal from the allosterically regulated sender enzyme, E , to the receiver protein, Y .

For illustrative purposes, the reactions described in the processes (a)–(d) in Fig. 1 can be simulated in a straightforward fashion using the Gillespie algorithm [22]. Here, we simulate the product arrival events described in Fig. 1 with the help of a stochastic simulation for a sender enzyme in a system with a fixed large number of substrates. In the case of an allosterically up-regulated sender enzyme (i.e., when the rate constant/s k_{X^*+} or/and d_P is/are greater than rate constant/s k_{X+} or/and r_P), we observed “bursts” of higher rates of product formation events in the midst of a lower rate of product formation; see Fig. 2. On the other hand, without any allosteric regulation, the waiting times between product arrivals are approximately exponentially distributed (as expected). Put differently, the ratio of the mean squared to the variance in the waiting times between product arrival events, which is ≈ 1.00 , for the choice of parameters specified in the caption of Fig. 2, as expected for exponentially distributed waiting times. This ratio strongly deviates from unity (≈ 17.40) in the presence of up-allosteric regulation in the sender enzyme; see Fig. 2.

Moreover, the stochastic simulations are also able to distinguish the variation of product formation rates of sender enzymes with the amount of substrate present in the presence of allosteric regulation, and in its absence (see Fig. 1 in the Supplemental Material [23]). Analogous results are observed in the presence of down-regulation in sender enzymes. In

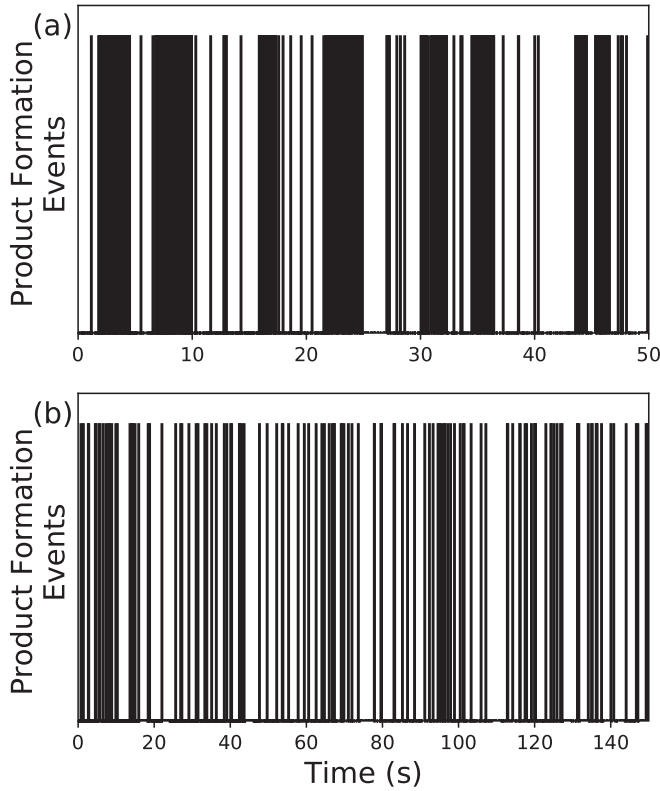


FIG. 2. Time series of product arrivals stochastically simulated using Gillespie's algorithm from coupled chemical reactions shown in Fig. 1 with an allosterically regulated (a) and unregulated (b) protein. The product arrivals are represented by vertical lines. The model with allostery exhibits a “bursty behavior.” This behavior is not observed in the absence of allostery. Parameter values used are $k_{X+} = 50 \text{ s}^{-1}$, $k_{X-} = 25 \text{ s}^{-1}$, $d_p = 0.9 \text{ s}^{-1}$, $k_{A+} = 50 \text{ s}^{-1}$, $k_{A-} = 25 \text{ s}^{-1}$, $k_{X^*+} = 75 \text{ s}^{-1}$, $k_{X^*-} = 25 \text{ s}^{-1}$, $r_p = 50 \text{ s}^{-1}$, $s_p = 70 \text{ s}^{-1}$, $h_p = 50 \text{ s}^{-1}$ for (a) and all the same except $h_p = 0 \text{ s}^{-1}$ for (b). For both the simulations we have an enzyme present in an excess of substrate (50 molecules).

addition, our stochastic simulations also suggests that the model is able to replicate a “bursty” product formation in the presence of both K -type (i.e., when allosteric regulation manifests itself by modulating the catalytic sites' binding affinity with the substrate, Fig. 2 of the Supplemental Material [23]) and V -type (i.e., when allosteric regulation affects the rate by which the bound complex of the catalytic site reduces to give products, Fig. 3 of the Supplemental Material [23]) allosteric regulations.

The results of stochastic simulation suggest that in the presence of allosteric up/down regulation (usually referred to as positive/negative cooperativity), the time of arrival of products exhibits a more complicated behavior as the catalytic site can exist in more than one type of state. Next, the products generated in bursts may bind to the receiver protein downstream transmitting the signal. Due to the bursty product formation, the amount of information encoded in the up-signal, propagated by the binding events between the products and the receiver protein, is a complicated function that depends on the time-dependent interaction between the

allosteric regulator site and the catalytic site of the sender enzyme.

To study the information encoded in the bursts of products in allosterically regulated sender enzymes, the model in Fig. 1 can be further simplified under the condition of detailed balance in steps (a) and (c) in Fig. 1 as shown in the section below.

A. Coupled chemical reactions for allosteric regulation reduce to a two-state system

As the system reaches detailed balance, i.e., when the forward and backward reaction of complex formation between substrate and the sender enzyme are equal, the concentrations of bound complexes do not change with time. Therefore, for reactions in step (a) in Fig. 1, when the rate of formation of $[E_{AX_S}]$ (and $[E_{A_S X_S}]$) and its rate of degradation are equal,

$$\frac{d[E_{AX_S}]}{dt} = k_{X+}[E_{AX}]_{\text{eq}}[S]_{\text{eq}} - k_{X-}[E_{AX_S}]_{\text{eq}} = 0, \quad (1)$$

$$\frac{d[E_{A_S X_S}]}{dt} = k_{X+}[E_{A_S X}]_{\text{eq}}[S]_{\text{eq}} - k_{X-}[E_{A_S X_S}]_{\text{eq}} = 0, \quad (2)$$

respectively. Here, E_{AX_S} represents the sender enzyme E with the unbound allosteric site A and the catalytic site X bound to a substrate, and the subscript “eq” stands for the concentration of the corresponding biochemical species in detailed balance. Hence, we can now write

$$\frac{[E_{AX_S}]_{\text{eq}}}{[E_{AX}]_{\text{eq}}[S]_{\text{eq}}} = \frac{k_{X+}}{k_{X-}}, \quad \frac{[E_{A_S X_S}]_{\text{eq}}}{[E_{A_S X}]_{\text{eq}}[S]_{\text{eq}}} = \frac{k_{X+}}{k_{X-}}. \quad (3)$$

Similarly, when the rate of formation of the bound allosteric site $[E_{A_S X}]$ (and $[E_{A_S X_S}]$) and its rate of degradation are equal, we can write

$$\frac{[E_{A_S X}]_{\text{eq}}}{[E_{AX}]_{\text{eq}}[S]_{\text{eq}}} = \frac{k_{A+}}{k_{A-}}, \quad \frac{[E_{A_S X_S}]_{\text{eq}}}{[E_{AX_S}]_{\text{eq}}[S]_{\text{eq}}} = \frac{k_{A+}}{k_{A-}}. \quad (4)$$

Likewise, for reaction (c) in Fig. 1 when then rate of formation of complex between the catalytic site in its alternate state (X^*) and the substrate, i.e., $[E_{AX_S^*}]$ (and $[E_{A_S X_S^*}]$), and its rate of degradation are equal, we can write

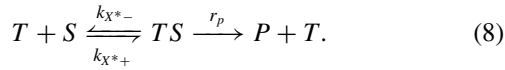
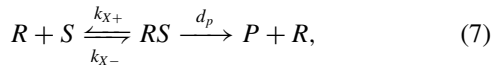
$$\frac{[E_{AX_S^*}]_{\text{eq}}}{[E_{AX^*}]_{\text{eq}}[S]_{\text{eq}}} = \frac{k_{X^*+}}{k_{X^*-}}, \quad \frac{[E_{A_S X_S^*}]_{\text{eq}}}{[E_{A_S X^*}]_{\text{eq}}[S]_{\text{eq}}} = \frac{k_{X^*+}}{k_{X^*-}}. \quad (5)$$

From hereon, for simplicity and clarity, we would refer to the sender enzymes with its unbound catalytic site in its original state (X) as R and the sender enzyme with its unbound catalytic in its alternate state (X^*) as T . Therefore, RS and TS would represent their bound forms, respectively. In addition, given the initial concentration of the sender enzyme, $[E]_{\text{initial}}$, mass balance implies an additional condition on the concentrations of the chemical species as shown below:

$$\begin{aligned} [E]_{\text{initial}} &= [E_{AX}]_{\text{eq}} + [E_{AX_S}]_{\text{eq}} + [E_{A_S X}]_{\text{eq}} + [E_{A_S X_S}]_{\text{eq}} \\ &\quad + [E_{AX^*}]_{\text{eq}} + [E_{AX_S^*}]_{\text{eq}} + [E_{A_S X^*}]_{\text{eq}} + [E_{A_S X_S^*}]_{\text{eq}}, \\ [E]_{\text{initial}} &= [R]_{\text{eq}} + [RS]_{\text{eq}} + [T]_{\text{eq}} + [TS]_{\text{eq}}. \end{aligned} \quad (6)$$

Combining Eq. (6) with Eqs. (3), (4), and (5) with the condition that substrate is present in excess, i.e., $[S]_{\text{eq}} \approx$

$[S]_{\text{initial}}$, we can reduce the model to a two-state model as shown below:



Due to the detailed balance in reaction (7), the rate of formation of RS and its degradation will be equal, i.e.,

$$\frac{d[RS]_{\text{eq}}}{dt} = k_{X+}[R]_{\text{eq}}[S]_{\text{eq}} = k_{X-}[RS]_{\text{eq}}. \quad (9)$$

Therefore,

$$[RS]_{\text{eq}} = \frac{k_{X+}}{k_{X-}}[R]_{\text{eq}}[S]_{\text{eq}}. \quad (10)$$

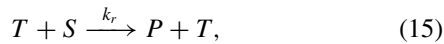
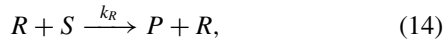
Also, the rate of production formation from step (a) in Fig. 1 can be expressed using reaction (7) and Eq. (10) as

$$\frac{d[P]}{dt} = d_p[RS]_{\text{eq}}, \quad (11)$$

$$\frac{d[P]}{dt} = \frac{k_{X+}d_p}{k_{X-}}[R]_{\text{eq}}[S]_{\text{eq}}, \quad (12)$$

$$\frac{d[P]}{dt} = k_R[R]_{\text{eq}}[S]_{\text{eq}}, \quad (13)$$

where k_R can be treated as the effective forward rate for the production of products in reaction (7). Following a similar logic, reactions (7) and (8) can be expressed in the two-state system as



where $k_R = \frac{k_{X+}d_p}{k_{X-}}$ and $k_T = \frac{k_{X*+}r_p}{k_{X*-}}$ are the effective product formation rate constants for the generation of products from the sender enzyme in the two states (R and T , respectively). In addition, as shown in Fig. 1, the switching between the states of the sender enzyme is facilitated by processes (b) and (d).

B. Hidden Markov model representation of the coupled chemical reactions

This model can be further simplified with the help of HMMs. In the language of HMMs, the state of the sender enzyme can be represented by a variable (s) that can take any value between 1 (for state R) and 2 (for state T). Further on, in order to represent the state of the sender enzyme as a Markov chain, we discretize time into intervals of δt . The time interval δt is selected to be small enough such that only one or no product can be produced in any time interval regardless of the state of the sender enzyme. This choice of small δt also ensures that the sender enzyme retains its state during δt . This approximation is also backed by several studies that suggest that allostery manifests itself on timescales varying from the order of 10 ps to several nanoseconds [24]. As a result, timescales pertaining to the rates of transitions between the states of sender enzyme (i.e., R to T and T to R) are several

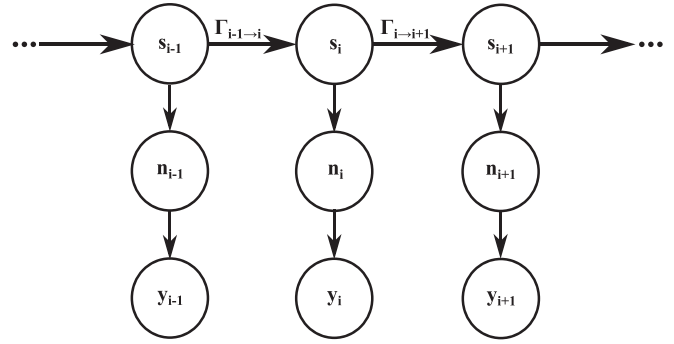


FIG. 3. Graphical model describing a two-state system of a sender enzyme with Poissonian emission and a receiver protein. Here the observable, n_i , is the number of products produced by a catalytic site at time level i in the hidden state s_i . $\Gamma_{i-1 \rightarrow i}$ represents the transition matrix of the states between time steps $i-1$ to i , and y_i shows the status of the receiver protein at time interval i .

orders slower than the timescales involved in the formation of products in a given state (see Fig. 3).

Therefore, the probability distribution of the states of the sender enzyme during time interval i [$P(s_i)$] can be expressed as a function of the probability distribution of states during the time interval $i-1$ and the dynamics of the catalytic site and the allosteric site in the sender enzyme as [25]

$$P(s_i|s_{i-1}) = \Gamma_{i-1 \rightarrow i}, \quad (16)$$

where, $\Gamma_{i-1 \rightarrow i}$ represents the transition matrix whose elements describe the probability of transition to a state s_i at time interval i given the state of the sender enzyme at time interval $i-1$. For a sender enzyme in state 1 ($s_i = 1$), the probability of remaining in state 1, $P_{1 \rightarrow 1}$, is proportional to $[R]_{\text{eq}}[S]_{\text{eq}}k_R$ from Eq. (14), whereas the probability for switching its state from state 1 to 2, $P_{1 \rightarrow 2}$, will be proportional to h_p times the probability of having a bound allosteric site, Fig. 1(b), i.e., $h_p \frac{[E_{AX}]_{\text{eq}}[S]_{\text{eq}}k_{A+}}{k_{A-}}$. On the other hand, if the sender enzyme is in state 2, the probability to remain in the state $P_{2 \rightarrow 2}$ is proportional to $[T]_{\text{eq}}[S]_{\text{eq}}k_T$ and the probability to switch, $P_{2 \rightarrow 1}$, is proportional to s_p from Fig. 1(d). Using these, the transition probabilities can be explicitly calculated using appropriate normalization conditions as

$$\begin{aligned} P_{1 \rightarrow 1} &= \frac{[R]_{\text{eq}}[S]_{\text{eq}}k_R}{[R]_{\text{eq}}[S]_{\text{eq}}k_R + h_p \frac{[E_{AX}]_{\text{eq}}[S]_{\text{eq}}k_{A+}}{k_{A-}}}, \\ P_{1 \rightarrow 2} &= \frac{h_p \frac{[E_{AX}]_{\text{eq}}[S]_{\text{eq}}k_{A+}}{k_{A-}}}{[R]_{\text{eq}}[S]_{\text{eq}}k_R + h_p \frac{[E_{AX}]_{\text{eq}}[S]_{\text{eq}}k_{A+}}{k_{A-}}}, \\ P_{2 \rightarrow 1} &= \frac{s_p}{[T]_{\text{eq}}[S]_{\text{eq}}k_T + s_p}, \\ P_{2 \rightarrow 2} &= \frac{[T]_{\text{eq}}[S]_{\text{eq}}k_T}{[T]_{\text{eq}}[S]_{\text{eq}}k_T + s_p}. \end{aligned} \quad (17)$$

Moreover, once the detailed balance is achieved in steps (a) and (c) in Fig. 1, the dynamics of the above described process becomes independent in time. Therefore, $\Gamma_{i+1 \rightarrow i} = \Gamma: \forall i \in 1, 2, \dots, M$ for M time steps, where $M \rightarrow \infty$.

The HMM model in Fig. 3 describes the dynamics of a catalytic site existing in state s_i at time step i and produces products n_i as described by the Poisson process given below:

$$P(n_i|s_i) = \frac{e^{-\lambda_{s_i}} \lambda_{s_i}^{n_i}}{n_i!}, \quad (18)$$

where λ_{s_i} is the average number of products generated by the sender enzyme in state s_i at time interval i . This parameter is closely related to the rate constants for chemical processes described in Fig. 1:

$$\lambda_{s_i} = \begin{cases} d_p[RS]_{\text{eq}}\delta t & \text{if } s_i = 1, \\ r_p[TS]_{\text{eq}}\delta t & \text{if } s_i = 2. \end{cases} \quad (19)$$

Here, $[RS]_{\text{eq}}$ and $[TS]_{\text{eq}}$ are the number of bound complexes between the substrate and the catalytic site in states 1 and 2, respectively, when detailed balance exists in steps (a) and (c) in Fig. 1. It should be noted that here we have used a number of bound complexes as opposed to the traditional use of concentration. This change is also reflected in the units of the rate constants. Therefore, using this HMM framework, the probability of generating n_i products during time interval i can be written as

$$P(n_i) = \sum_{s_i=1}^2 P(n_i|s_i)P(s_i). \quad (20)$$

Finally, we must also describe how n_i changes the state of the receiver protein. As we select a time interval for the event to occur, at most one product can be produced in any time interval by the catalytic site. Consistent with assumptions inherent to the Gillespie simulation, we assume that diffusion occurs on timescales vastly exceeding the rate of any chemical reaction (including the product production rate). The product can then either bind to the receiver protein or disappear, by diffusing to a sink (such as an off-pathway, receiver), but it does not accumulate. Hence, the probability of the state of the receiver protein downstream in the signaling pathway whether the receiver protein is bound to the product at time interval i or not) can be given as

$$P(y_i = 1|n_i) = \begin{cases} 0 & \text{if } n_i = 0, \\ p_Y & \text{if } n_i = 1 \end{cases} \quad (21)$$

or

$$P(y_i = 0|n_i) = \begin{cases} 1 & \text{if } n_i = 0, \\ 1 - p_Y & \text{if } n_i = 1, \end{cases} \quad (22)$$

where the state variable y_i reflects a binding event at the receiver protein during time interval i such that $y_i = 1$ for a successful binding event and $y_i = 0$ otherwise. Here, p_Y is a real number less than 1, which depicts the probability of activation of the receiver in the presence of the product. Moreover, for simplicity, the dynamics of the receiver are not included in the model. Hence, at the beginning of each time interval, the receiver is assumed to be in the refreshed state waiting to receive a new product. The probabilities shown above describe the dynamics of the model, which will be used for calculating the amount of information encoded by the sender enzyme while communicating with the receiver protein downstream in the signaling pathway.

Therefore, at time interval i , the products generated by the sender enzyme, which acts as a sender of a message,

are received by another receiver protein downstream in the signaling pathway, which encodes the information onto the y_i .

III. RESULTS AND DISCUSSION

In the preceding section, we reduced the dynamical model of allosteric regulation of a sender enzyme to a simplified two-state model that can be represented by an HMM in discrete time. The model consists of a sender enzyme with a catalytic site and an allosteric regulator site. The sender enzyme binds with substrate, which leads to the generation of products. The sender enzyme can exist in two different states as R (when the catalytic site exists as X) and T (when the catalytic site exists as X^*). The two states differ in their dynamics of substrate binding and the rate of product formation by the catalytic site.

The sender enzyme communicates with a receiver protein downstream in the signaling pathway with the help of products generated in time. The binding of the product with the receiver protein represents a successful reception of the signal. Finally, we can express the dynamics of switching of states of the sender protein as a function of allosteric interaction between the catalytic site and the allosteric regulator.

Now, exploiting tools from information theory, we define the amount of information conveyed by the sender enzyme to the receiver protein downstream at time interval i as the mutual information (MI) [21] between them,

$$\text{MI} = \sum_{s_i=1}^2 \sum_{y_i=0}^1 P(s_i, y_i) \ln \frac{P(s_i, y_i)}{P(s_i)P(y_i)}, \quad (23)$$

where $P(s_i, y_i)$ represents the joint probability distribution between s_i and y_i at time interval i . The MI can be computed for any i (see the Supplemental Material for more details [23]). Here we choose a large value of i , $i = 500$, to ensure that the detailed balance is achieved in steps (a) and (c) in Fig. 1 and that the MI no longer varies with i .

In the model, the modulation of the information conveyed through product arrivals can be achieved by two different mechanisms: (i) by only changing the average number of products generated by the sender enzyme in the state R or T , and (ii) by altering the switching probabilities between these two states.

We first investigated the effect of varying the product formation rate of the first state R , λ_1 , while keeping the product formation rate of the second state T and the switching probabilities between states constant, Fig. 4. First, as a sanity check, we observe that without any allosteric regulation, i.e., when the second T state is never visited by the sender enzyme [setting the transition probability of the sender enzyme from state R to T ($P_{1 \rightarrow 2}$) to zero], the MI between the state of the sender enzyme and the binding events at the receiver protein is zero. This is an expected result, as in this case the sender protein is restricted to a single state, and alternate states can no longer influence the binding of product at the receiver protein. Therefore, no information is conveyed by the products from the sender enzyme to the receiver protein, regardless of the rate of product formation by the sender.

Next, when $\lambda_1 < \lambda_2$, i.e., in the case of up-regulation, the MI between the sender enzyme and the receiver protein is higher, in contrast to the case when allosteric regulation is

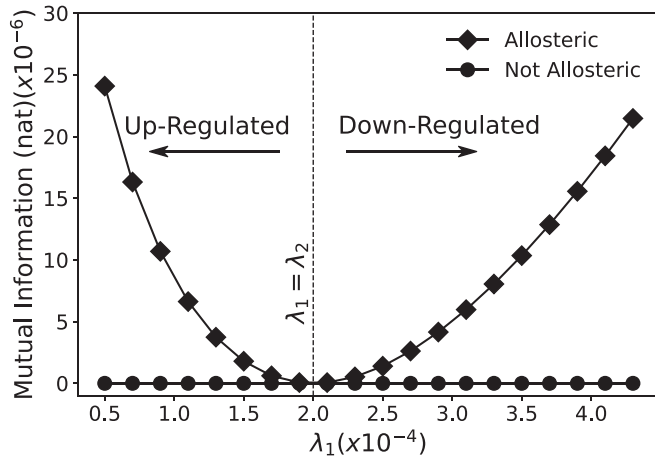


FIG. 4. Effect of allostery, as captured by λ_1 , on the information communicated by the sender enzyme. With allosteric up-regulation, i.e., when $\lambda_1 < \lambda_2$, the time series of product arrivals encodes a larger amount of information. The dotted line represents the value of λ_1 for which $\lambda_1 = \lambda_2$ and where the benefit of allostery on the MI vanishes. With allosteric down-regulation, i.e., when $\lambda_1 > \lambda_2$, the sender enzyme still communicates more information than in the absence of allostery. For the case of the allosteric sender enzyme, we used $\lambda_2 = 2 \times 10^{-4}$, and the transition probability of the sender enzyme from state R to T , $P_{1 \rightarrow 2}$ and vice versa, $P_{2 \rightarrow 1}$ as 0.2 and for the case of nonallosteric sender enzyme, $P_{1 \rightarrow 2} = 0$.

absent (i.e., when $P_{1 \rightarrow 2} = 0$). However, as λ_1 increases and comes closer to the value of λ_2 , the difference between the product formation rates in the two states of the sender enzyme decreases and so does the MI until it drops down to zero when $\lambda_1 = \lambda_2$. This observation is counterintuitive as it suggests that the MI between the sender enzyme and the receiver protein decreases with an increase in the average rate of product formation. Equally counterintuitively, the information communicated by the sender to the receiver rises once again when $\lambda_1 > \lambda_2$, i.e., for the case of down-regulation.

Second, we analyzed the variation of the information transmitted with allosteric regulation by varying the probability of switching of the state of the sender enzyme from R to T , i.e., $P_{1 \rightarrow 2}$, while holding other parameters (i.e., λ_1 , λ_2 , and $P_{2 \rightarrow 1}$) fixed, Fig. 5. As a sanity check, we first observed the amount of information conveyed in the absence of allosteric regulation, i.e., when $\lambda_1 = \lambda_2$. As expected (similar to the case observed above), the signal no longer carries any MI when the effect of allostery is absent, and it is also insensitive to the value of $P_{1 \rightarrow 2}$. This implies that the binding events at the receiver protein would be independent of the state of the sender enzyme, and as a result there is no MI between them. On the other hand, for up- and down-regulated proteins, the sender enzyme conveys minimum information while being restricted to only one state (i.e., when $P_{1 \rightarrow 2} = 0$). However, as $P_{1 \rightarrow 2}$ grows, with up-regulation the sender enzyme is able to access a state with higher average production and therefore conveys a larger amount of information to the receiver protein. However, unexpectedly, with down-regulation the sender enzyme is still able to transmit a larger amount of information for a range of switching probabilities. This counterintuitive observation illustrates once more that information transmitted

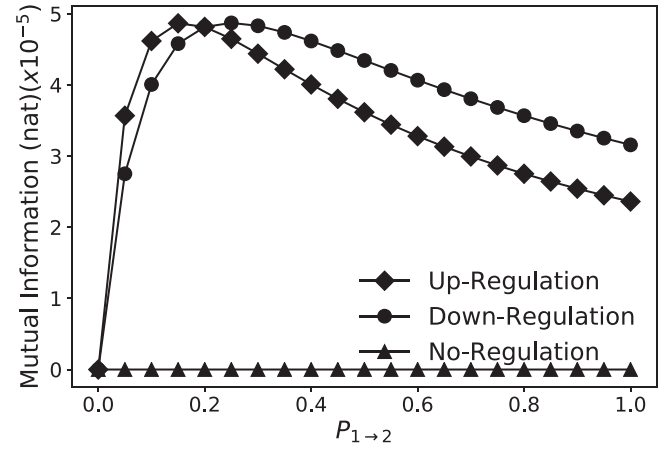


FIG. 5. Effect of changing the switching probability from state R to T , $P_{1 \rightarrow 2}$ on the information communicated by the sender enzyme for different types of allosteric regulations. We observe that without any allosteric regulation, the information encoded in the product time series does not vary with switching probabilities. However, with up-regulation, i.e., when $\lambda_1(1 \times 10^{-4}) < \lambda_2(4 \times 10^{-4})$, the sender enzyme communicates more information as the probability of it switching to an up-regulated state increases. Counterintuitively, a down-regulated sender enzyme, i.e., when $\lambda_1(4 \times 10^{-4}) > \lambda_2(1 \times 10^{-4})$, is still able to communicate a larger amount of information for a wide range of switching probabilities. For all plots, we fix the transition probability from state T to R , $P_{2 \rightarrow 1}$ to 0.08.

is not a mere function of the amount of product available to the receiver protein but also depends on the switching kinetics between the allosteric states of the sender enzyme.

IV. CONCLUSION

Here we investigated a minimal model of allostery and quantified the information communicated by the catalytic site of an allosterically regulated sender enzyme to another receiver protein further down in the signaling pathway through time-varying product formation rates. Our choice of a two-state model is a matter of convenience; our formalism readily generalizes to more complex kinetic models. More importantly, we found that an allosterically regulated enzyme may convey a larger amount of information as compared to an enzyme with no allosteric regulation by decreasing its net rate of product formation. This suggests that allostery may provide the means to control the information encoded in the time of arrival of products in a way that goes beyond the energetically demanding “more product, better signal” exploitative paradigm. That is, allostery may provide “lower signal but more information.” The possibility of parameter fine-tuning to communicate more information is especially relevant given allostery’s key role in protein evolution [26–28]. It opens the possibility that nature may fine-tune allosteric parameters (including switching rates between states as well as production rates) to adapt/evolve its signaling pathways in response to external stimuli warranting exploratory (high information/low signal) or exploitative (low information/high signal) strategies [29].

ACKNOWLEDGMENTS

This project is supported by ARO Grant No. W911NF-17-1-0162 on “Multi-Dimensional and Dissipative Dynamical Systems: Maximum Entropy as a Principle for Modeling

Dynamics and Emergent Phenomena in Complex Systems.” S.B.O. and T.M. also acknowledge the NSF-MCB Award 1715591 and Moore Foundation. We also thank Ioannis Sgouralis and Sean Seyler for helpful discussions.

- [1] R. Milo, What is the total number of protein molecules per cell volume? A call to rethink some published values, *BioEssays* **35**, 1050 (2013).
- [2] S. J. Wodak, E. Paci, N. V. Dokholyan, I. N. Berezovsky, A. Horovitz, J. Li, V. J. Hilser, I. Bahar, J. Karanicolas, G. Stock, P. Hamm, R. H. Stote, J. Eberhardt, Y. Chebaro, A. Dejaegere, M. Cecchini, J.-P. Changeux, P. G. Bolhuis, J. Vreede, P. Faccioli, S. Orioli, R. Ravasio, L. Yan, C. Brito, M. Wyart, P. Gkeka, I. Rivalta, G. Palermo, J. A. McCammon, J. Panecka-Hofman, R. C. Wade, A. D. Pizio, M. Y. Niv, R. Nussinov, C.-J. Tsai, H. Jang, D. Padhorny, D. Kozakov, and T. McLeish, Allostery in its many disguises: From theory to applications, *Structure* **27**, 566 (2019).
- [3] H. Link, D. Christodoulou, and U. Sauer, Advancing metabolic models with kinetic information, *Curr. Opin. Biotechnol.* **29**, 8 (2014), cell and pathway engineering.
- [4] S. Tafoya and C. Bustamante, Molecular switch-like regulation in motor proteins, *Philos. Trans. R. Soc. London, Ser. B* **373**, 20170181 (2018).
- [5] A. B. Herr, Evolution of an allosteric off switch in apoptotic caspases, *J. Biol. Chem.* **293**, 5462 (2018).
- [6] J. Monod, J. Wyman, and J.-P. Changeux, On the nature of allosteric transitions: A plausible model, *J. Mol. Biol.* **12**, 88 (1965).
- [7] A. Cooper and D. T. Dryden, Allostery without conformational change. A plausible model, *Eur. Biophys. J.* **11**, 103 (1984).
- [8] D. E. Koshland, G. Nmethy, and D. Filmer, Comparison of experimental binding data and theoretical models in proteins containing subunits, *Biochemistry* **5**, 365 (1966).
- [9] T. Lenaerts, J. Ferkinghoff-Borg, F. Stricher, L. Serrano, J. W. Schymkowitz, and F. Rousseau, Quantifying information transfer by protein domains: Analysis of the Fyn SH2 domain structure, *BMC Struct. Biol.* **8**, 43 (2008).
- [10] M. V. LeVine and H. Weinstein, NbIT - A new information theory-based analysis of allosteric mechanisms reveals residues that underlie function in the leucine transporter LeuT, *PLoS Comput. Biol.* **10**, e1003603 (2014).
- [11] M. Voliotis, R. M. Perrett, C. McWilliams, C. A. McArdle, and C. G. Bowsher, Information transfer by leaky, heterogeneous, protein kinase signaling systems, *Proc. Natl. Acad. Sci. USA* **111**, E326 (2014).
- [12] C. Adami, The use of information theory in evolutionary biology, *Ann. N.Y. Acad. Sci.* **1256**, 49 (2012).
- [13] H. P. de Vlarar and N. H. Barton, The contribution of statistical physics to evolutionary biology, *Trends Ecol. Evol.* **26**, 424 (2011).
- [14] O. Weiss, M. A. Jimnez-montao, and H. Herzog, Information content of protein sequences, *J. Theor. Biol.* **206**, 379 (2000).
- [15] S. Turkcan and J.-B. Masson, Bayesian decision tree for the classification of the mode of motion in single-molecule trajectories, *PLoS One* **8**, 1 (2013).
- [16] M. Tavakoli, J. N. Taylor, C.-B. Li, T. Komatsuzaki, and S. Press, Single molecule data analysis: An introduction, *Advances in Chemical Physics* (Wiley, New York, 2017), Chap. 4, pp. 205–305.
- [17] M. Komorowski and D. S. Tawfik, The limited information capacity of cross-reactive sensors drives the evolutionary expansion of signaling, *Cell Syst.* **8**, 76 (2019).
- [18] S. Marzen, H. G. Garcia, and R. Phillips, Statistical mechanics of Monod-Wyman-Changeux (MWC) models, *J. Mol. Biol.* **425**, 1433 (2013).
- [19] L. Rabiner and B. Juang, An introduction to hidden Markov models, *IEEE ASSP Mag.* **3**, 4 (1986).
- [20] K. Seymore, A. McCallum, and R. Rosenfeld, Learning hidden Markov model structure for information extraction, in *AAAI 99 Workshop on Machine Learning for Information Extraction* (AAAI, New York, 1999), pp. 37–42.
- [21] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley Series in Telecommunications and Signal Processing) (Wiley-Interscience, New York, 2006).
- [22] D. T. Gillespie, Exact stochastic simulation of coupled chemical reactions, *J. Phys. Chem.* **81**, 2340 (1977).
- [23] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevResearch.2.023367> for detailed expression for calculation of the mutual information and other supplemental figures.
- [24] L. R. McDonald, M. J. Whitley, J. A. Boyer, and A. L. Lee, Colocalization of fast and slow timescale dynamics in the allosteric signaling protein chey, *Allosteric Interactions and Biological Regulation* (Part II), *J. Mol. Biol.* **425**, 2372 (2013).
- [25] N. G. V. Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier, Amsterdam, 1992).
- [26] M. Tushar, H. Jonathan, G. Kingshuk, and O. S. Banu, Ancient thioredoxins evolved to modern-day stability function requirement by altering native state ensemble, *Philos. Trans. R. Soc. London B* **373**, 20170184 (2018).
- [27] R. Nussinov, C.-J. Tsai, and J. Liu, Principles of allosteric interactions in cell signaling, *J. Am. Chem. Soc.* **136**, 17692 (2014).
- [28] P. D. Townsend, T. L. Rodgers, L. C. Glover, H. J. Korhonen, S. A. Richards, L. J. Colwell, E. Pohl, M. R. Wilson, D. R. W. Hodgson, T. C. B. McLeish, and M. J. Cann, The role of protein-ligand contacts in allosteric regulation of the Escherichia coli catabolite activator Protein, *J. Biol. Chem.* **290**, 22225 (2015).
- [29] T. T. Hills, P. M. Todd, D. Lazer, A. D. Redish, I. D. Couzin, and Cognitive Search Research Group, Exploration versus exploitation in space, mind, and society, *Trends Cognit. Sci.* **19**, 46 (2015).