Information transmission in a two-step cascade: interplay of activation and repression

Tuhin Subhra Roy1 · Mintu Nandi2 · Ayan Biswas1 · Pinaki Chaudhury2 · Suman K. Banik1

Received: 31 March 2021 / Accepted: 14 September 2021 / Published online: 5 October 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract
We present an information-theoretic formalism to study signal transduction in four architectural variants of a model two-step cascade with increasing input population. Our results categorize these four types into two classes depending upon the effect of activation and repression on mutual information, net synergy, and signal-to-noise ratio. Using the Gaussian framework and linear noise approximation, we derive the analytic expressions for these metrics to establish their underlying relationships in terms of the biochemical parameters. We also verify our approximations through stochastic simulations.

Keywords Gene regulatory network · Two-step cascade · Linear noise approximation · Mutual information · Signal-to-noise ratio · Partial information decomposition

Introduction
Signal transduction is an essential process for a living system to survive in a dynamic environment. By the mechanism of signal transduction, a living entity senses the changes in its immediate surroundings. While doing so, a single or a multicellular species utilizes biochemical pathways to process the environmental information Alon (2006, 2007). Information processing within the purview of the gene regulatory network (GRN) has attracted lots of attention recently. Starting from simple motif to complex networks, researchers have addressed key issues related to information processing capacity, fidelity, and functionality of the network of interest (Cheong et al. 2011; Bowsher and Swain 2014; Mc Mahon et al. 2014; Selimkhanov et al. 2014; Hansen and O’Shea 2015; Jost 2020).

Quantification of information processing can be achieved using the tools of information theory formulated by Shannon Shannon (1948); Shannon and Weaver (1963). Shannon’s formalism provides the most used measure, mutual information (MI) \( I(i; j) \) defined between two random variables \( i \) and \( j \) Cover and Thomas (2006)

\[
I(i; j) = \sum_{i,j} p(i,j) \log_2 \left[ \frac{p(i,j)}{p(i)p(j)} \right].
\]

(1)

Here, \( p(i) \) and \( p(j) \) are the marginal and \( p(i,j) \) is the joint probability function. MI is measured in bits due to base 2 in the logarithmic function. The above relation suggests that a probabilistic interpretation is necessary for variables of interest. This essentially leads to the study of cellular dynamics in a fluctuating environment. To calculate MI, one thus needs a stochastic description of the signal transduction mechanism that mimics the dynamics in a single cell environment.

In general, MI is a measure of the correlation between two random variables whose entropy or fluctuation spaces overlap. From Eq. (1), it can be easily inferred that \( I(i; j) \geq 0 \), the lower bound being realized whenever \( p(i,j) = p(i)p(j) \) or in other words, correlation is lost between \( i \) and \( j \), i.e., these random variables become independent of each other. MI is a special case of what is generally known to mathematicians; relative entropy or the Kullback-Leibler (KL) distance which quantifies the distance between two probability spaces but does not meet the requirement set by the triangle inequality. But unlike the KL distance, MI by construction is symmetric with respect to its argument variables, i.e., \( I(i;j) = I(j;i) \) Cover and Thomas (2006). Due to this feature, MI on its own cannot provide insights into the temporal structure of processed information.
by a set of random variables. From the entropy perspective, the following relations hold between MI and corresponding marginal \( I(H(i), H(j)) \), joint \( I(H(i), j) \), and the conditional \( I(H[ij], H[j]) \) entropy definitions are given by Cover and Thomas (2006).

\[
I(i; j) = H(i) - H(i|j), \tag{2a}
\]
\[
= H(j) - H(j|i), \tag{2b}
\]
\[
= H(i) + H(j) - H(i, j). \tag{2c}
\]

Where, the entropy definitions are

\[
H(i) = - \sum_i p(i) \log_2 p(i), \tag{2d}
\]
\[
H(i, j) = - \sum_{i,j} p(i, j) \log_2 p(i, j), \tag{2e}
\]
\[
H(i|j) = - \sum_{i,j} p(i, j) \log_2 p(i|j). \tag{2f}
\]

Here, \( p(i|j) \) is the conditional probability function. Equations (2a–c), in essence, demonstrate the fact that MI quantifies the average reduction in entropy of one variable when knowledge of the other one is given. The extension of MI into the multivariate realm is straightforward and is utilized in the present paper. Categorizing the associated random variables into sources and targets of information takes us to ask whether assembling more than one source variable simultaneously can benefit us in extracting more information than what is possible otherwise just by adding their contributions. For a three-variable case, the difference between the three-variable total MI and sum of the two constituent two-variable MIs is often used to quantify the information independence between the source variables Schneidman et al. (2003). Since this metric characterizes the effective synergistic aspect between the information sources and can be shown to be equal to the difference between synergistic and redundant information according to the formalism of partial information decomposition (PID) Williams and Beer (2010), it is often dubbed as the net synergy Barrett (2015).

For a system comprised of random variables \( i, j \) and \( k \), the three variable mutual information \( I(i; j, k) \) is written as

\[
I(i; j, k) = \sum_i \sum_{j,k} p(i, j, k) \log_2 \left( \frac{p(i, j, k)}{p(i)p(j, k)} \right), \tag{3}
\]

whose entropy representation is given by Cover and Thomas (2006), Maity et al. (2017)

\[
I(i; j, k) = H(i) + H(j, k) - H(i, j, k). \tag{4}
\]

Now considering \( i \) as a target variable and \( j \) and \( k \) as source variables, one can rewrite Eq. (3) as Barrett (2015)

\[
I(i; j, k) = U(i; j|k) + U(i; k|j) + R(i; j, k), \tag{5a}
\]

where \( U(i; j|k) \) is the unique information between the source variable \( j \) and the target variable \( i \) given the other source variable is \( k \). On a similar note, \( U(k; j|i) \) is the unique information between the source variable \( k \) and the target variable \( i \), out of two source variables \( j \) and \( k \). The synergistic information \( S(i; j, k) \) is the information of the target variable \( i \) when the knowledge of both source variables \( j \) and \( k \) is known. When knowledge of either of the source variables \( j \) and \( k \) is sufficient to predict the nature of the target variable \( i \), one can quantify the redundant information \( R(i; j, k) \). Similarly, two variable mutual information \( I(i; j) \) and \( I(i; k) \) can be written in terms of unique and redundant information

\[
I(i; j) = U(i; j|k) + R(i; j, k), \tag{5b}
\]
\[
I(i; k) = U(i; k|j) + R(i; j, k). \tag{5c}
\]

Following the formalism of PID, one can define the measure \( \Delta I(i; j, k) \) using the above expressions of three variable and two variable mutual information Williams and Beer (2010), Barrett (2015)

\[
\Delta I(i; j, k) = I(i; j, k) - I(i; j) - I(i; k), \tag{5d}
\]
\[
= S(i; j, k) - R(i; j, k). \tag{5e}
\]

The above definition of net synergy suggests that for a multivariate system, the bivariate or the trivariate (multivariate) mutual information alone does not provide useful information about the system. However, an algebraic combination of these measures gives insightful information regarding the process. Depending on the nature of interaction among the nodes of a network \( \Delta I \) can take three different values, i.e., \( > 0, < 0 \) and \( = 0 \) Schneidman et al. (2003). For \( \Delta I > 0 \), the synergistic effect overpowers the redundant effect. In other words, common knowledge of both source variables is more effective than the knowledge of a single source variable. Similarly, \( \Delta I < 0 \) focuses on redundant information sharing between target and source variables. The scenario \( \Delta I = 0 \) leads to the borderline situation where both common information sharing balances redundant information sharing between source and target variables leading to information independence. We note that classification of net synergy \( \Delta I \) has been developed in the context of population coding in neuroscience Schneidman et al. (2003); however, its application in GTRN is limited.

The present paper focuses on this principle information-theoretic metric in the perspective of model GRNs, and the next section shows how this can be computed analytically.
in terms of variances and covariances using a well-known approximation technique. The bacterial GRNs are often found to be enriched with small modular structures named as network motifs Alon (2006). A two-step cascade (TSC) may be regarded as one prime example of this category. In a TSC, one transcription factor (TF) regulates another TF which finally regulates the output gene-expression, which may result in the production of key enzymes and proteins with a possibility for the latter species acting as regulators in other motifs. Regarding genes/gene-products as nodes and transcriptional regulations (activation/repression) as edges have simplified the gene-regulation phenomenon into a network analysis problem. To this end, we have considered four types of TSC depending on the possible interaction topology between the nodes, namely activation-activation (AA), activation-repression (AR), repression-activation (RA) and repression-repression (RR) (see Fig. 1).

In one of our earlier reports Biswas and Banik (2016), we have investigated the responses of net synergy and the signal-to-noise ratio (SNR) as a function of input degradation rate parameter in an AA type of TSC with fixed population levels. In the current manuscript, we intend to extend the analysis to showcase those metrics’ responses in all of the possible topologies of the generic TSC network with tunable signaling strength. In the following, we have studied information processing in four different architectures available to a generic TSC while paying attention to crucial metrics like MI, net synergy, and SNR. Our observations demonstrate the categoric influence played out by transcriptional regulation on fluctuation propagation in TSC pathways. The present work also aims to highlight the dependence of signal fidelity on the interaction topology of a TSC. Moreover, the present construct verifies the role of the net synergy as an indicator of the signaling efficiency in midst of the stochasticity due to the interaction topologies of the TSC.

Model and methods

The Langevin equations governing the dynamics of a generic TSC are Biswas and Banik (2016).

\[
\frac{ds}{dt} = f_s(s) - \mu_s s + \xi_s(t),
\]

\[
\frac{dx}{dt} = f_x(s, x) - \mu_x x + \xi_x(t),
\]

\[
\frac{dy}{dt} = f_y(s, x, y) - \mu_y y + \xi_y(t).
\]

In Eqs. (6–8), s, x and y are the copy numbers of species S, X and Y, respectively, expressed in molecules/V. Here, V is the unit effective cellular volume. Here, copy number stands for population of a species. \(f_s(s)\), \(f_x(s, x)\) and \(f_y(s, x, y)\) are the production term associated with s, x and y, respectively.

We note that the production term of each component, in general, is nonlinear in nature Bintu et al. (2005); Ziv et al. (2007); Tkačik et al. (2008, 2008); de Ronde et al. (2012). The degradation rate parameters associated with \(s, x\) and \(y\) are written as \(\mu_s, \mu_x,\) and \(\mu_y\), respectively. The constants \(\mu_s, \mu_x,\) and \(\mu_y\) correspond to the inverse lifetime of the gene product of the respective component.

The noise processes \(\xi_i(t)\) considered here are independent and Gaussian distributed with properties \(\langle \xi_i(t) \rangle = 0\) and \(\langle \xi_i(t) \xi_j(t') \rangle = \langle |\xi_i|^2 \rangle \delta_{ij} \delta(t - t')\) with \(\langle |\xi|^2 \rangle = \langle f_i \rangle + \mu_i (i = s, x, y)\) at steady state Elf and Ehrenberg (2003); Swain (2004); Paulsson (2004); Tănase-Nicola et al. (2006); Warren et al. (2006); van Kampen (2007); Mehta et al. (2008); de Ronde et al. (2010). The nature of noise correlation also suggests that the noise processes are uncorrelated with each other Swain (2016). Here, \(\langle \cdots \rangle\) represents steady state ensemble average. The noise processes we consider here are intrinsic in nature. As a result, variability induced in the dynamics is controlled by the intrinsic noise processes although it has been observed that fluctuations in gene expression in a genetically identical population are mostly controlled by extrinsic noise processes Elowitz et al. (2002); Raser and O’Shea (2004). Depending on the nature of interactions, activation or repression shown in Fig. 1, we consider specific form of the functions \(f_s(s)\), \(f_x(s, x)\) and \(f_y(s, x, y)\) (see Table 1).

We now employ linear noise approximation (LNA) Elf and Ehrenberg (2003); van Kampen (2007) to calculate the second moments associated with \(s, x\) and, \(y\). To this end, we use perturbation of linear order \(\delta u(t) = u(t) - \langle u \rangle\) where \(\langle u \rangle\) is the average population of \(u\) at steady state and recast Eqs. (6–8) in the following form

\[
\frac{d\delta W}{dt} = J_{W \to W} \delta W(t) + \Xi(t).
\]

In the above equation, \(\delta W(t)\) denotes the fluctuation matrix containing the linear order perturbation terms. The noise matrix \(\Xi(t)\) along with \(\delta W(t)\) can be written as

| Table 1 List of functions used in four different TSC motifs |
|-----------------------------------------------------------|
| Motif       | \(f_s(s)\)     | \(f_x(s, x)\)     | \(f_y(s, x, y)\)     |
| AA          | \(k_s\)        | \(k_s \frac{e^{-K_n s}}{K_n s + e^{-K_n s}}\) | \(k_s \frac{e^{-K_n s}}{K_n s + e^{-K_n s}}\) |
| (S → X → Y) |               |                   |                   |
| AR          | \(k_s\)        | \(k_s \frac{e^{-K_n s}}{K_n s + e^{-K_n s}}\) | \(k_s \frac{e^{-K_n s}}{K_n s + e^{-K_n s}}\) |
| (S → X → Y) |               |                   |                   |
| RA          | \(k_s\)        | \(k_s \frac{e^{-K_n s}}{K_n s + e^{-K_n s}}\) | \(k_s \frac{e^{-K_n s}}{K_n s + e^{-K_n s}}\) |
| (S → X → Y) |               |                   |                   |
| RR          | \(k_s\)        | \(k_s \frac{e^{-K_n s}}{K_n s + e^{-K_n s}}\) | \(k_s \frac{e^{-K_n s}}{K_n s + e^{-K_n s}}\) |
| (S → X → Y) |               |                   |                   |
The diagram, $\rightarrow$ and $\leftarrow$ stand for activation and repression, respectively.

$$\delta W(t) = \begin{pmatrix} \delta s(t) \\ \delta x(t) \\ \delta y(t) \end{pmatrix}, \Xi(t) = \begin{pmatrix} \xi_s(t) \\ \xi_x(t) \\ \xi_y(t) \end{pmatrix}. $$

$J$ represents the Jacobian matrix at steady state,

$$J = \begin{bmatrix} f_{s,s}'(s) - \mu_s & 0 & 0 \\ f_{x,s}'(s, x) & f_{x,x}'(s, x) - \mu_x & 0 \\ f_{y,s}'(s, y) & f_{y,x}'(s, x, y) & f_{y,y}'(s, x, y) - \mu_y \end{bmatrix}. $$

Here, $f_{x, s}'$ stands for differentiation of $f_x$ with respect to $s$ and evaluated at $s = \langle s \rangle$, and so on. The variance and covariance are evaluated using the Lyapunov equation at steady state Keizer (1987); Elf and Ehrenberg (2003); Paulsson (2004, 2005); van Kampen (2007).

$$J\Sigma + \Sigma J^T + D = 0. $$

In Lyapunov equation, $\Sigma$ and $D$ stand for covariance matrix and diffusion matrix, respectively. The diffusion matrix $D$ incorporates different noise strength using the relation $D = \langle \Xi \Xi' \rangle$. As the noise processes $\xi_s$, $\xi_x$ and $\xi_y$ are uncorrelated, the off-diagonal elements of the diffusion matrix $D$ are zero. Here, $T$ represents transpose of a matrix. Analytical solution of Eq. (10) yields variance and covariance associated with different TSC motifs (see Table 2).

The variance and covariance associated with different TSC motifs given in Table 2 can be used to calculate the two and three variable MIs using Gaussian framework Shannon (1948); Cover and Thomas (2006); Barrett (2015).

### Table 2

| Motif | $\Sigma(s)$ | $\Sigma(x, x)$ | $\Sigma(x, y)$ | $\Sigma(x)$ | $\Sigma(x, y)$ | $\Sigma(y)$ |
|-------|-------------|----------------|----------------|-------------|----------------|-------------|
| AA    | $\langle s \rangle$ | $\frac{\partial K(s)}{(\partial s)^2} + (\frac{\partial \Sigma(s)}{(\partial s)^2})$ | $\frac{\partial K(s)}{(\partial s)^2} + (\frac{\partial \Sigma(s)}{(\partial s)^2})$ | $\frac{\partial K(s)}{(\partial s)^2} + (\frac{\partial \Sigma(s)}{(\partial s)^2})$ | $\frac{\partial K(s)}{(\partial s)^2} + (\frac{\partial \Sigma(s)}{(\partial s)^2})$ | $\frac{\partial K(s)}{(\partial s)^2} + (\frac{\partial \Sigma(s)}{(\partial s)^2})$ |
| AR    | $\langle s \rangle$ | $\frac{\partial K(s)}{(\partial s)^2} + (\frac{\partial \Sigma(s)}{(\partial s)^2})$ | $\frac{\partial K(s)}{(\partial s)^2} + (\frac{\partial \Sigma(s)}{(\partial s)^2})$ | $\frac{\partial K(s)}{(\partial s)^2} + (\frac{\partial \Sigma(s)}{(\partial s)^2})$ | $\frac{\partial K(s)}{(\partial s)^2} + (\frac{\partial \Sigma(s)}{(\partial s)^2})$ | $\frac{\partial K(s)}{(\partial s)^2} + (\frac{\partial \Sigma(s)}{(\partial s)^2})$ |
| RA    | $\langle s \rangle$ | $\frac{\partial K(s)}{(\partial s)^2} + (\frac{\partial \Sigma(s)}{(\partial s)^2})$ | $\frac{\partial K(s)}{(\partial s)^2} + (\frac{\partial \Sigma(s)}{(\partial s)^2})$ | $\frac{\partial K(s)}{(\partial s)^2} + (\frac{\partial \Sigma(s)}{(\partial s)^2})$ | $\frac{\partial K(s)}{(\partial s)^2} + (\frac{\partial \Sigma(s)}{(\partial s)^2})$ | $\frac{\partial K(s)}{(\partial s)^2} + (\frac{\partial \Sigma(s)}{(\partial s)^2})$ |
| RR    | $\langle s \rangle$ | $\frac{\partial K(s)}{(\partial s)^2} + (\frac{\partial \Sigma(s)}{(\partial s)^2})$ | $\frac{\partial K(s)}{(\partial s)^2} + (\frac{\partial \Sigma(s)}{(\partial s)^2})$ | $\frac{\partial K(s)}{(\partial s)^2} + (\frac{\partial \Sigma(s)}{(\partial s)^2})$ | $\frac{\partial K(s)}{(\partial s)^2} + (\frac{\partial \Sigma(s)}{(\partial s)^2})$ | $\frac{\partial K(s)}{(\partial s)^2} + (\frac{\partial \Sigma(s)}{(\partial s)^2})$ |

$\Sigma(s)$, $\Sigma(x)$ and $\Sigma(y)$ are the variance of $s$, $x$ and $y$, respectively. The covariance between $s$ and $x$, $s$ and $y$, and $x$ and $y$ are written as $\Sigma(s, x)$, $\Sigma(s, y)$ and $\Sigma(x, y)$, respectively.
I(s, x) = \frac{1}{2} \log_2 \left[ \frac{\Sigma(s) \Sigma(x)}{\Sigma(s) \Sigma(x) - \Sigma^2(x, x)} \right], \quad (11)

I(s, y) = \frac{1}{2} \log_2 \left[ \frac{\Sigma(s) \Sigma(y)}{\Sigma(s) \Sigma(y) - \Sigma^2(s, y)} \right], \quad (12)

I(s, x, y) = \frac{1}{2} \log_2 \left[ \frac{\Sigma(s) (\Sigma(x) \Sigma(y) - \Sigma^2(x, y))}{|\Delta|} \right], \quad (13)

where

\[ \Delta = \begin{pmatrix} \Sigma(s) & \Sigma(s, x) & \Sigma(s, y) \\ \Sigma(s, x) & \Sigma(x) & \Sigma(x, y) \\ \Sigma(s, y) & \Sigma(x, y) & \Sigma(y) \end{pmatrix} \]

In the above expressions, \( \Sigma(s), \Sigma(x) \) and \( \Sigma(y) \) are the variance of \( s, x \) and \( y \), respectively. The covariance between \( s \) and \( x \), \( s \) and \( y \), and \( x \) and \( y \) are written as \( \Sigma(s, x), \Sigma(s, y) \) and \( \Sigma(x, y) \), respectively.

We note here that calculations presented up to Eq. (10) are applicable for Hill coefficient \( n \geq 1 \) within the purview of LNA. In the next section, we have presented results for \( n = 1 \) and \( n = 2 \). For \( n > 1 \), say \( n = 2 \), etc., will bring in higher levels of nonlinearity in the dynamics. We have, however, excluded the phenomenon of gene switching, autoregulation (positive/negative), and cross-regulation to avoid bimodal distribution of system variables. Our model also does not show the generation of noise-induced bistability for an otherwise deterministic monostable system To and Maheshri (2010).

The kinetics of signal transmission in all the TSC motifs are simulated using stochastic simulation algorithm Gillespie (1976, 1977). The associated kinetic schemes and propensities used for numerical simulation are given in Table 3. Table 3 suggests that even for \( n = 1 \) the propensities used are of Hill type. The propensities used in the present work are not derived from the first principle and hence may lead to misleading results even under the purview of LNA Thomas et al. (2012). The limitation of using intrinsic noise, Hill coefficient \( n = 1 \), and phenomenological propensities within the framework of LNA could be overcome by the recent extension of LNA Thomas et al. (2014); Keizer et al. (2019). In our calculation, we have used \( \mu_1 < \mu_2 < \mu_3 \) for maximum information flow due to separation of time scale Maity et al. (2015). To this end, we used \( \mu_1 = 0.1 \text{ sec}^{-1}, \mu_2 = 1 \text{ sec}^{-1}, \mu_3 = 10 \text{ sec}^{-1} \). We set \( K_1 = 50 \text{ (molecules/V)} \) and \( K_2 = 100 \text{ (molecules/V)} \). The mean copy numbers of \( s, x \) and \( y \) are 50, 100 and 100, respectively, with unit molecules/V. The average population of signal \( s \) is tuned to compute different statistical measures. Statistical averaging is done using the steady state output of \( 10^6 \) independent trajectories.

For biochemical kinetics under the influence of Gaussian noise processes, one can make use of different variance and covariance to define two variable and three variable MI. These MI are then used to calculate net synergy for the TSC motif Schneidman et al. (2003), Barrett (2015)

\[ \Delta I(s; x, y) = I(s; x, y) - I(s; x) - I(s; y). \]

Earlier analysis Biswas and Banik (2016) of a TSC system reveals that \( I(s; x, y) \approx I(s; x) \) and \( I(s; x) \geq I(s; y) \) due to data processing inequality (DPI) and Markov chain property of

| AA | AR |
|---|---|
| Kinetics | Rate Constant | Propensity | Kinetics | Rate Constant | Propensity |
| \( \phi \to S \) | \( k_s \) | \( k_s \) | \( \phi \to S \) | \( k_s \) | \( k_s \) |
| \( S \to \phi \) | \( \mu_s \) | \( \mu_s \) | \( S \to \phi \) | \( \mu_s \) | \( \mu_s \) |
| \( S \to S + X \) | \( k_x \) | \( k_x \frac{K_1}{K_1 + s} \) | \( S \to S + X \) | \( k_x \) | \( k_x \frac{K_1}{K_1 + s} \) |
| \( X \to \phi \) | \( \mu_x \) | \( \mu_x \) | \( X \to \phi \) | \( \mu_x \) | \( \mu_x \) |
| \( X \to X + Y \) | \( k_y \) | \( k_y \frac{K_2}{K_2 + y} \) | \( X \to X + Y \) | \( k_y \) | \( k_y \frac{K_2}{K_2 + y} \) |
| \( Y \to \phi \) | \( \mu_y \) | \( \mu_y \) | \( Y \to \phi \) | \( \mu_y \) | \( \mu_y \) |
| RA | RR | RA | RR |
| \( \phi \to S \) | \( k_s \) | \( k_s \) | \( \phi \to S \) | \( k_s \) | \( k_s \) |
| \( S \to \phi \) | \( \mu_s \) | \( \mu_s \) | \( S \to \phi \) | \( \mu_s \) | \( \mu_s \) |
| \( S \to S + X \) | \( k_x \) | \( k_x \frac{K_1}{K_1 + s} \) | \( S \to S + X \) | \( k_x \) | \( k_x \frac{K_1}{K_1 + s} \) |
| \( X \to \phi \) | \( \mu_x \) | \( \mu_x \) | \( X \to \phi \) | \( \mu_x \) | \( \mu_x \) |
| \( X \to X + Y \) | \( k_y \) | \( k_y \frac{K_2}{K_2 + y} \) | \( X \to X + Y \) | \( k_y \) | \( k_y \frac{K_2}{K_2 + y} \) |
| \( Y \to \phi \) | \( \mu_y \) | \( \mu_y \) | \( Y \to \phi \) | \( \mu_y \) | \( \mu_y \) |
the linear architecture Cover and Thomas (2006). To be specific for a TSC \( I(s;x) > I(s;y) \). Considering this Eq. (14) can be written as

\[
\Delta I(s; x, y) = -\frac{1}{2} \log_2 \left[ 1 + \frac{\Sigma^2(s, y) - \Sigma^2(s, y)}{\Sigma(s)\Sigma(y)} \right],
\]

where the different MI terms are expressed in the unit of bits due to the base 2 in the logarithmic function. We note here that the quantity \( \Sigma^2(s, y) [\Sigma(s)\Sigma(y) - \Sigma^2(s, y)] \) is defined as signal-to-noise ratio (SNR). SNR provides a measure of the redundancy of the signaling cascade Cheong et al. (2011); Bowsher et al. (2013). A high value of SNR between the input and the output reveals high signal processing capacity of a biochemical network.

**Results and discussion**

In the present article, we present a comparative study among the four TSC motifs with different architecture. Our goal here is to find out how architecture dependent information transmission is correlated with the fidelity of the networks. To this end, we tune the average input signal population \( \langle s \rangle \) within the physiological range to study the proposed metrics of TSC motifs. To do so, we vary \( k_s \), the synthesis rate constant of S, according to the relation \( \langle s \rangle = k_s/\mu_s \), keeping the degradation rate constant \( \mu_s \) fixed at a particular value. The mentioned relation can be obtained from Eq. (6) at the steady state.

We first quantify the two variable and three variable MIs - \( I(s; x), I(s; y) \), and \( I(s; x, y) \) using Shannon’s definition for \( n = 1 \) and \( n = 2 \). These three MIs are plotted against the mean copy number of \( S, \langle s \rangle \) in Fig. 2 for all four motifs. We observe a decreasing trend of MIs for AA and AR motifs and an increasing trend for RA and RR motifs. Validity of DPI and Markov chain property, i.e., \( I(s;x, y) \approx I(s;x) \) and \( I(s;x) > I(s;y) \) of the TSC motifs can be verified from Fig. 2. We note that the profiles for AA and AR are the same irrespective of the architectural difference between the two motifs. RA and RR motifs also show similarity in their MI profiles. This observation suggests that the information transduction along the cascades depends only on the regulation \( S \rightarrow X \) but not on \( X \rightarrow Y \). We also quantify net synergy and SNR for all motifs as a function of \( \langle s \rangle \) (Fig. 3).

It is important to mention that for \( n = 2 \), we observe disagreement between theoretical profiles and simulation results at low \( \langle s \rangle \) (see Figs. 2e–h and 3b, d). As mentioned in the previous section, \( n > 1 \) incorporates higher order nonlinearity in the system which also becomes prominent at low \( \langle s \rangle \). The higher value of the Hill coefficient \( (n > 1) \) takes into account the cooperative binding of TF to its binding site. This in turn brings in higher order nonlinearity in the dynamics due to which the network species experiences greater fluctuations. These fluctuations are more prominent when the mean population of signal \( \langle s \rangle \) is low. This can be understood from Fig. 3b, d which shows that at low \( \langle s \rangle \) simulation results deviate from the theoretical profiles. For AA and AR motifs, such deviation is large. The deviations may even get larger in the case of \( n \geq 2 \) due to the breakdown of LNA.

Figure 3a, b shows the values of net synergy are on the negative region, indicating the fact that this metric quantifies the net redundant information expressed by the output \( Y \). This is further confirmed from Eq. (15), which shows that the net synergy value is equal to the negative magnitude of \( I(s; y) \). The net synergy profiles for AA and AR fall on each other, and so are RA and RR. This behavior is similar to the MI profiles shown in Fig. 2. The profiles for AA and AR are increasing while RA and RR are decreasing, and both become equal at high \( \langle s \rangle \). To account for all these behavior, we focus on the analytical expression of SNR as the expression of net synergy has a logarithmic function (see Eq. 15). Using the expressions of various moments listed in Table II, we have expressions of SNR for four different motifs (for \( n = 1 \))

\[
\text{SNR}_{\text{AA}} = \frac{B^2_1}{a_1 + (\langle y \rangle + F_1)\langle s \rangle (K_1 + \langle s \rangle)^2},
\]

\[
\text{SNR}_{\text{AR}} = \frac{B^2_2}{a_2 + (\langle y \rangle + F_2)\langle s \rangle (K_1 + \langle s \rangle)^2},
\]

\[
\text{SNR}_{\text{RA}} = \frac{B^2_3\langle s \rangle}{a_3\langle s \rangle + (\langle y \rangle + F_3)\langle s \rangle (K_1 + \langle s \rangle)^2},
\]

\[
\text{SNR}_{\text{RR}} = \frac{B^2_4\langle s \rangle}{a_4\langle s \rangle + (\langle y \rangle + F_4)\langle s \rangle (K_1 + \langle s \rangle)^2}.
\]

We refer to Appendix for detailed derivation of the above expressions of SNR along with the explicit forms of \( a_i, b_i, f_i \) and \( f_i \) \( (i = 1, 2, 3, 4) \). These expressions have been plotted in Fig. 3c, d, which shows similar collating phenomena of AA & AR and RA & RR, as observed for net synergy. The profiles shown in Fig. 3c, d are obtained in the limit \( K_1 \gg \langle s \rangle \). We further consider \( K_1 + \langle s \rangle \approx K_1 \) and approximate Eqs. (16–19) leading to

\[
\text{SNR}_{\text{AA}} \approx \frac{B^2_1}{a_1 + K_1^2(\langle y \rangle + F_1)\langle s \rangle},
\]

\[
\text{SNR}_{\text{AR}} \approx \frac{B^2_2}{a_2 + K_1^2(\langle y \rangle + F_2)\langle s \rangle},
\]
\[
\text{SNR}_{\text{RA}} \approx \frac{B_2^2 \langle s \rangle}{\alpha_3 \langle s \rangle + K_1^2 \langle y \rangle + F_3},
\]
(22)

\[
\text{SNR}_{\text{RR}} \approx \frac{B_4^2 \langle s \rangle}{\alpha_4 \langle s \rangle + K_1^2 \langle y \rangle + F_4}.
\]
(23)

From Appendix, it is clear that \(B_1 = B_2\), \(\alpha_1 = \alpha_2\), and \(F_1 = F_2\), since in generating the results we have used, \(\langle x \rangle = K_2\) (see all figure captions). This makes \(\text{SNR}_{\text{AA}} = \text{SNR}_{\text{AR}}\). Similarly, it can also be established that \(\text{SNR}_{\text{RA}} = \text{SNR}_{\text{RR}}\). This explains the similarity between AA and AR and between RA and RR. As mentioned earlier, the information transmission depends only on the regulation of \(S\) to \(X\) and not on the regulation of \(X\) to \(Y\). This fact can also
be understood from Eqs. (20–23), which contain only \( K_1 \) as a signature of S to X regulation. However, the effect of X to Y regulation is encapsulated in the parameter \( K_2 \), which we surpass by making it equal with \( \langle s \rangle \).

It is clear from Eqs. (20–21) that SNR for AA and AR is a decreasing function of \( \langle s \rangle \), while Eqs. (22-23) indicate that SNR for RA and RR is, indeed, an increasing function of \( \langle s \rangle \). Such dependency on \( \langle s \rangle \) is consistent with the simulation result shown in Fig. 3c, d. These variations are significant only at low \( \langle s \rangle \). At high \( \langle s \rangle \), SNR of all motifs becomes equal. Since net synergy is the negative logarithm of SNR, the profiles of SNR are replicated in the profiles of net synergy with an opposite trend due to the negative sign.

In an earlier study, it has been shown that the intermediate component X in a TSC motif possesses some hindrance ability. As a result, information flow from S to Y gets reduced as the signal strength increases Nandi et al. (2019). Such information restriction is prominent in the case of AA and AR motifs. In both the motifs, the signal S activates the intermediate X, whose restriction ability increases as a function of signal strength. Hence, the information content \( I(s; y) \) decreases (see Fig. 2) and consequently net synergy increases (see Fig. 3a, b) as a function of \( \langle s \rangle \). This enhanced restriction character also causes a reduction in signalling fidelity of AA and AR motifs (see Fig. 3c,d). On the other hand, in the case of RA and RR motifs, S represses the synthesis of X. Hence, increasing signal strength here may decrease the restriction ability of X, which indeed results in elevated information transmission from S to Y, i.e., \( I(s; y) \) increases with \( \langle s \rangle \) (see Fig. 2). Reduction in restriction ability of X, again, in turn, decreases the net synergy and increases the fidelity of RA and RR motifs (see Fig. 3).

The above analysis indicates that the information transmission capacity along the four TSC motifs not only depends on the signal strength but also depends on the architecture of the motifs. In particular, the activation and repression phenomena also play an essential role. However, at a high signal population, such interplay of activation and repression does not remain valid, and all the motifs behave similarly. We have mentioned earlier that the results are presented in the limit \( K_1 \geq \langle s \rangle \), which indicates that at \( \langle s \rangle = 50 \) half-maximum point is reached (since for each plot we set \( K_1 = 50 \)). We set the parameter \( K_1 \) in such a way that the extreme right points in each plot correspond to the half-maximum; whatever be the value of \( \langle s \rangle \), up to which the proposed metrics are calculated. The convergence of net synergy and SNR continues even for \( \langle s \rangle > 50 \) by setting \( K_1 = \max \langle s \rangle \). We note that if one crosses the half-maximum (by setting \( K_1 = 50 \) and by tuning \( \langle s \rangle \) beyond 50), such convergence will break down and a crossover will occur.

Here, it is imperative to understand the biological relevance of increasing \( \langle s \rangle \) on the proposed metrics. We note here that we focus only on the activation-repression processes for S to X regulation. This is, perhaps, because the information transduction along the cascades depends only on S to X regulation. The signal S activates gene X by binding at the promoter region upstream of RNA polymerase (RNAP) binding site (see Fig. 4a). Consequently, the TF (or the signal) S together with RNAP stimulates the transcription processes so that protein X gets produced. On the other
hand, S binds to the site downstream of RNAP binding site to inhibit the process of transcription (see Fig. 4b). As a result, S represses the production of protein X.

When the signal population is high, S has a greater probability to get attached to their binding sites to either activate or inhibit the process of transcription. Such increased probability may lead to the incorporation of less fluctuations into the system of AA and AR. In consequence, information content $I(s; y)$ becomes less since information transmission along a network deciphers the propagation of the fluctuations. We note here that these fluctuations capture the knowledge of the concentration changes at the signal level Nandi et al. (2018). Hence, net synergy increases for both AA and AR. On the other hand, increased binding probability of S leads to suppression of transcription rates in RA and RR. In other words, a reduced transcription rate may imply more randomness in the output copy number. Therefore, $I(s; y)$ increases, and net synergy decreases with $\langle s \rangle$. All these ups and downs in information transmission, in turn, affect the fidelity of the networks, as discussed previously.

**Conclusion**

In the present work, we studied information transmission in four different architecture of TSC motifs. Using the formalism of LNA, we have calculated various two variable and three variable MIs, net synergy, and SNR in terms of variance and covariance terms associated with the networks variables. We approximate these variables to follow Gaussian distribution. We also validated the proposed metrics by numerical simulation using Gillespie algorithm.

Quantification of MIs shows validation of DPI and Markov chain properties for four linear architectures having different interaction topology. We further validate our finding of redundancy in information transmission using the formalism of PID. Signal fidelity is quantified using the measure SNR. The variations in SNR can well explain the behavior of net synergy, as the latter is a logarithmic function of the former. From the approximate expressions of SNR, the clustering phenomena are understood.

The decreasing (of AA and AR) and increasing (of RA and RR) nature of different metrics with signal population are consistent with the approximate analytical expressions of SNR. We also postulate that the information transmission in these motifs depends only on the regulation of S to X. In other words, when the regulation is activating in nature, information transmission capacity goes down with increasing signal population. On the other hand, for repressive regulation, information transmission capacity increases with the signal population. Our analysis thus suggests that in a generic TCS motif, X to Y regulation has an insignificant role in information transmission. It has also been observed that the profiles of net synergy and SNR for AA and AR, and RA and RR differ significantly at low signal population. However, when the signal population is high, all the architectures have almost equal magnitudes of net synergy as well as SNR. At a high signal population, the interplay between activation and repression is no more a determining factor in the information dynamics of TSC motifs.

The four TSC motifs we considered in this work have simple architecture and can be constructed using genetic engineering and recombinant technology. To the best of our knowledge, only the RR motif has been constructed synthetically to study response delays in a transcription network of *E. coli* Rosenfeld.
and Alon (2003). Using the similar principle of synthetic biology, one can construct the rest of the motifs (AA, AR, and RA) and experimentally verify signal fidelity in linear architectures with different interaction topologies. Our results demonstrate that a greater amount of redundant information in comparison with its synergistic counterpart ensures a significant proportion of the SNR or signal fidelity in the TSC irrespective of its topological specifics as well as the signaling strength.

**Appendix: Calculation of SNR**

For AA, we have from Table II

\[
\Sigma(s,x) = \frac{k_y K_1 \langle s \rangle}{(K_2 + \langle x \rangle)^2(\mu_s + \mu_x)},
\]

\[
= \frac{\mu_s(\langle x \rangle)(K_1 + \langle s \rangle)}{\langle s \rangle \times (K_1 + \langle s \rangle)^2(\mu_s + \mu_x)},
\]

\[
= \frac{K_1 \mu_s \langle x \rangle}{\mu_s + \mu_x} \times \frac{1}{K_1 + \langle s \rangle},
\]

\[
= \frac{A_1}{K_1 + \langle s \rangle},
\]

where we have used \( k_y = \mu_s(\langle x \rangle)(K_1 + \langle s \rangle)/\langle s \rangle \) and \( A_1 = K_1 \mu_s(\langle x \rangle)/\mu_s + \mu_x \) for. Again using the expression of \( \Sigma(s,x) \) and \( k_y = \mu_s(\langle x \rangle)(K_2 + \langle s \rangle)/\langle s \rangle \), we rewrite the expression of \( \Sigma(s,y) \),

\[
\Sigma(s,y) = \frac{k_y K_2 \Sigma(s,x)}{(K_2 + \langle x \rangle)^2(\mu_s + \mu_x)},
\]

\[
= \frac{\mu_s(\langle y \rangle)(K_2 + \langle x \rangle)}{\langle x \rangle \times (K_2 + \langle x \rangle)^2(\mu_s + \mu_y)},
\]

\[
= \frac{K_1 \mu_s \langle y \rangle}{\mu_s + \mu_x} \times \frac{1}{K_1 + \langle s \rangle},
\]

\[
= \frac{B_1}{K_1 + \langle s \rangle},
\]

where \( B_1 = K_1 K_2 \mu_s(\langle y \rangle)/(K_2 + \langle x \rangle)(\mu_s + \mu_x)(\mu_s + \mu_y) \). It is important to note here that the expressions of \( k_y \) and \( k_y \) are obtained by considering the Langevin equations corresponding to the motif at steady state. The expressions of other second moments are obtained by similar fashion and are given by,

\[
\Sigma(x) = \langle x \rangle + \frac{C_1}{\langle s \rangle(K_1 + \langle s \rangle)^2},
\]

\[
\Sigma(x,y) = D_1 + \frac{E_1}{\langle s \rangle(K_1 + \langle s \rangle)^2},
\]

\[
\Sigma(y) = \langle y \rangle + F_1 + \frac{G_1}{\langle s \rangle(K_1 + \langle s \rangle)^2},
\]

with

\[
C_1 = \frac{K_1^2 \mu_s(\langle x \rangle)^2}{\mu_s + \mu_x},
\]

\[
D_1 = \frac{K_2 \mu_s(\langle y \rangle)}{(K_2 + \langle x \rangle)(\mu_s + \mu_y)},
\]

\[
E_1 = \frac{K_1^2 K_2 \mu_s(\langle x \rangle)(\mu_s + \mu_y)}{(K_2 + \langle x \rangle)(\mu_s + \mu_x)(\mu_s + \mu_y)(\mu_s + \mu_y)},
\]

\[
F_1 = \frac{K_2^2 \mu_s(\langle y \rangle)^2}{(K_2 + \langle x \rangle)^2(\mu_s + \mu_y)},
\]

\[
G_1 = \frac{K_1^2 K_2 \mu_s(\langle y \rangle)^2(\mu_s + \mu_x)}{(K_2 + \langle x \rangle)^2(\mu_s + \mu_x)(\mu_s + \mu_y)(\mu_s + \mu_y)}.
\]

Now utilizing the expressions of different moments one can express SNR associated with the motif AA in terms of model parameters and \( \langle s \rangle \) as

\[
\text{SNR}_{AA} = \frac{B_1^2}{\alpha_1 + (\langle y \rangle + F_1)(\langle s \rangle(K_1 + \langle s \rangle))^2},
\]

with \( \alpha_1 = G_1 - B_1^2 \). At low population of the signal S, we approximate that \( K_1 \gg \langle s \rangle \) and hence \( K_1 + \langle s \rangle \approx K_1 \). This makes the SNR to be a hyperbolic function of \( \langle s \rangle \),

\[
\text{SNR}_{AA} \approx \frac{B_1^2}{\alpha_1 + K_1^2(\langle y \rangle + F_1)(\langle s \rangle)},
\]

For AR, we evaluate the expressions of different moments as a function of \( \langle s \rangle \) following the same strategy

\[
\Sigma(s,x) = \frac{A_2}{K_1 + \langle s \rangle},
\]

\[
\Sigma(s,y) = - \frac{B_2}{K_1 + \langle s \rangle},
\]

\[
\Sigma(x) = \langle x \rangle + \frac{C_2}{\langle s \rangle(K_1 + \langle s \rangle)^2},
\]

\[
\Sigma(x,y) = -D_2 - \frac{E_2}{\langle s \rangle(K_1 + \langle s \rangle)^2},
\]

\[
\Sigma(y) = \langle y \rangle + F_2 + \frac{G_2}{\langle s \rangle(K_1 + \langle s \rangle)^2},
\]

where \( A_2 = A_1 \), \( B_2 = (\langle x \rangle/K_2)B_1 \), \( C_2 = C_1 \), \( D_2 = (\langle x \rangle/K_2)D_1 \), \( E_2 = (\langle x \rangle/K_2)E_1 \), \( F_2 = (\langle x \rangle/K_2)^2F_1 \) and \( G_2 = (\langle x \rangle/K_2)^2G_1 \). The expression of SNR becomes

\[
\text{SNR}_{AR} = \frac{B_2^2}{\alpha_2 + (\langle y \rangle + F_2)(\langle s \rangle(K_1 + \langle s \rangle))^2}.
\]

Here, \( \alpha_2 = G_2 - B_2^2 = (\langle x \rangle/K_2)^2 \alpha_1 \). At low \( \langle s \rangle \) value, the SNR follows the same trend as was observed in case of AA motif.
For RA, we write the expressions of second moments as,

\[
\Sigma(s, x) = - A_3 \frac{\langle s \rangle}{K_1 + \langle s \rangle}, \\
\Sigma(s, y) = - B_3 \frac{\langle s \rangle}{K_1 + \langle s \rangle}, \\
\Sigma(x) = \langle x \rangle + C_3 \frac{\langle s \rangle}{(K_1 + \langle s \rangle)^2}, \\
\Sigma(x, y) = D_3 + E_3 \frac{\langle s \rangle}{(K_1 + \langle s \rangle)^2}, \\
\Sigma(y) = (\langle y \rangle + F_3) + G_3 \frac{\langle s \rangle}{(K_1 + \langle s \rangle)^2},
\]

where

\[
A_3 = \frac{\mu_s(x)}{\mu_s + \mu_e}, \\
B_3 = \frac{K_2 \mu_s \mu_e \langle y \rangle}{(K_2 + \langle x \rangle)(\mu_s + \mu_e)}, \\
C_3 = \frac{\mu_s \langle x \rangle^2}{\mu_s + \mu_e}, \\
D_3 = \frac{\mu_s \mu_e \langle y \rangle}{(K_2 + \langle x \rangle)(\mu_s + \mu_e)}, \\
E_3 = \frac{K_2 \mu_s \mu_e \langle x \rangle \langle y \rangle (\mu_s + \mu_e + \mu_i)}{(K_2 + \langle x \rangle)(\mu_s + \mu_e)(\mu_s + \mu_i)(\mu_s + \mu_e)}, \\
F_3 = \frac{K_2 \mu_s \mu_e \langle y \rangle^2}{(K_2 + \langle x \rangle)(\mu_s + \mu_e)(\mu_s + \mu_i)(\mu_s + \mu_e)}, \\
G_3 = \frac{K_2 \mu_s \mu_e \langle y \rangle^2 (\mu_s + \mu_e + \mu_i)}{(K_2 + \langle x \rangle)(\mu_s + \mu_e)(\mu_s + \mu_i)(\mu_s + \mu_e)}.
\]

Now, the expression of SNR is,

SNR_{RA} = \frac{B_3^2 \langle s \rangle}{\alpha_3 \langle s \rangle + (\langle y \rangle + F_3)(K_1 + \langle s \rangle)^2},

with \( \alpha_3 = G_3 - B_3^2 \). At low \( \langle s \rangle \), SNR will become,

SNR_{RA} \approx \frac{B_3^2 \langle s \rangle}{\alpha_3 \langle s \rangle + \beta_3^2 (\langle y \rangle + F_3)},

which is an increasing function of \( \langle s \rangle \).

On a similar note, for RR, the expressions of second moments are,

\[
\Sigma(s, x) = - A_4 \frac{\langle s \rangle}{K_1 + \langle s \rangle}, \\
\Sigma(s, y) = B_4 \frac{\langle s \rangle}{K_1 + \langle s \rangle}, \\
\Sigma(x) = \langle x \rangle + C_4 \frac{\langle s \rangle}{(K_1 + \langle s \rangle)^2}, \\
\Sigma(x, y) = D_4 - E_4 \frac{\langle s \rangle}{(K_1 + \langle s \rangle)^2}, \\
\Sigma(y) = (\langle y \rangle + F_4) + G_4 \frac{\langle s \rangle}{(K_1 + \langle s \rangle)^2},
\]

where \( A_4 = A_3 \), \( B_4 = (\langle x \rangle/K_2)B_3 \), \( C_4 = C_3 \), \( D_4 = (\langle x \rangle/K_2)D_3 \), \( E_4 = (\langle x \rangle/K_2)E_3 \), \( F_4 = (\langle x \rangle/K_2)^2F_3 \) and \( G_4 = (\langle x \rangle/K_2)^2G_3 \). The expression of SNR is,

\[
\text{SNR}_{RR} = \frac{B_3^2 \langle s \rangle}{\alpha_4 \langle s \rangle + \langle y \rangle + F_4)(K_1 + \langle s \rangle)^2},
\]

with \( \alpha_4 = G_4 - B_3^2 = (\langle x \rangle/K_2)^2\alpha_3 \). At low \( \langle s \rangle \), SNR shows similar trend as a function of \( \langle s \rangle \) as was observed for RA motif.

Acknowledgements T.S.R. and M.N. thank CSIR, India [09/015(0495)/2016-EMR-I] and UGC, India [22/06/2014(i)EU-V], respectively, for research fellowships.

References

Alon U (2006) An introduction to systems biology: design principles of biological circuits. CRC Press, Boca Raton
Alon U (2007) Network motifs: theory and experimental approaches. Nat Rev Genet 8:450
Barrett AB (2015) Exploration of synergistic and redundant information sharing in static and dynamical Gaussian systems. Phys Rev E 91:052802
Biswas A, Banik SK (2016) Redundancy in information transmission in a two-step cascade. Phys Rev E 93:052422
Bowsher CG, Swain PS (2014) Environmental sensing, information transfer, and cellular decision-making. Curr Opin Biotechnol 28:149
Bowsher CG, Voliotis M, Swain PS (2013) The fidelity of dynamic signaling by noisy biomolecular networks. PLoS Comput Biol 9:1002965
Cheong R, Rhee A, Wang CJ, Nemenman I, Levchenko A (2011) Information transduction capacity of noisy biochemical signaling networks. Science 334(6054):354
Cover TM, Thomas JA (2006) Elements of Information Theory, 2nd edn. Wiley-Interscience, New York
de Ronde WH, Tostevin F, ten Wolde PR (2010) Effect of feedback on the fidelity of information transmission of time-varying signals. Phys Rev E 82:031914
de Ronde WH, Tostevin F, ten Wolde PR (2012) Feed-forward loops and diamond motifs lead to tunable transmission of information in the frequency domain. Phys Rev E 86:021913
Elf J, Ehrenberg M (2003) Fast evaluation of fluctuations in biochemical networks with the linear noise approximation. Genome Res 13:2475
Elowitz MB, Levine AJ, Sigga ED, Swain PS (2002) Stochastic gene expression in a single cell. Science 297:1183
Gillespie DT (1976) A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. J Comp Phys 22:403
Gillespie DT (1977) Exact stochastic simulation of coupled chemical reactions. J Phys Chem 81:2340
Hansen AS, O’Shea EK (2015) Limits on information transduction through amplitude and frequency regulation of transcription factor activity. Elife 4:e06559
Jost J (2020) Biological information. Theory Biosci 139:361
Keizer J (1987) Statistical Thermodynamics of Nonequilibrium Processes. Springer-Verlag, Berlin
Keizer EM, Bastian B, Smith RW, Grima R, Fleck C (2019) Extending the linear-noise approximation to biochemical systems influenced by intrinsic noise and slow lognormally distributed extrinsic noise. Phys Rev E 99:052417
Maity AK, Chaudhury P, Banik SK (2015) Role of relaxation time scale in noisy signal transduction. PLoS One 10:123242
Maity AK, Chaudhury P, Banik SK (2017) Information theoretical study of cross-talk mediated signal transduction in MAPK pathways. Entropy 19:e19090469
Mc Mahon SS, Sim A, Filippi S, Johnson R, Liepe J, Smith D, Stumpf MP (2014) Information theory and signal transduction systems: from molecular information processing to network inference. Semin Cell Dev Biol 35:98
Mehta P, Goyal S, Wingreen NS (2008) A quantitative comparison of sRNA-based and protein-based gene regulation. Mol Syst Biol 4:221
Nandi M, Biswas A, Banik SK, Chaudhury P (2018) Information processing in a simple one-step cascade. Phys Rev E 98:042310
Nandi M, Banik SK, Chaudhury P (2019) Restricted information in a two-step cascade. Phys Rev E 100:032406
Paulsson J (2004) Summing up the noise in gene networks. Nature 427:415
Paulsson J (2005) Models of stochastic gene expression. Phys Life Rev 2:157
Raser JM, O’Shea EK (2004) Control of stochasticity in eukaryotic gene expression. Science 304:1811
Rosenfeld N, Alon U (2003) Response delays and the structure of transcription networks. J Mol Biol 329:645
Schneidman E, Bialek W, Berry MJ II (2003) Synergy, redundancy, and independence in population codes. J Neurosci 23:11539
Selimkhanov J, Taylor B, Yao J, Pilko A, Albeck J, Hoffmann A, Tsirring L, Wollman R (2014) Systems biology. Accurate information transmission through dynamic biochemical signaling networks. Science 346:1370
Shannon CE (1948) The mathematical theory of communication. Bell Syst Tech J 27:379
Shannon CE, Weaver W (1963) The mathematical theory of communication. University of Illinois Press, Urbana
Swain PS (2016) Lecture notes on stochastic models in systems biology. arXiv:1607.07806
Swain PS (2004) Efficient attenuation of stochasticity in gene expression through post-transcriptional control. J Mol Biol 344:965
Tânase-Nicola S, Warren PB, ten Wolde PR (2006) Signal detection, modularity, and the correlation between extrinsic and intrinsic noise in biochemical networks. Phys Rev Lett 97:068102
Thomas P, Straube AV, Grima R (2012) The slow-scale linear noise approximation: an accurate, reduced stochastic description of biochemical networks under timescale separation conditions. BMC Syst Biol 6:39
Thomas P, Popović N, Grima R (2014) Phenotypic switching in gene regulatory networks. Proc Natl Acad Sci USA 111:6994
Tkačik G, Callan CG, Bialek W (2008) Information flow and optimization in transcriptional regulation. Proc Natl Acad Sci USA 105:12265
Tkačik G, Gregor T, Bialek W (2008) The role of input noise in transcriptional regulation. PLoS ONE 3(7):e2774
To TL, Maheshri N (2010) Noise can induce bimodality in positive transcriptional feedback loops without bistability. Science 327:1142
van Kampen NG (2007) Stochastic Processes in Physics and Chemistry, 3rd edn. North-Holland, Amsterdam
Warren PB, Tânase-Nicola S, ten Wolde PR (2006) Exact results for noise power spectra in linear biochemical reaction networks. J Chem Phys 125:144904
Williams PL, Beer RD (2010) Nonnegative decomposition of multivariate information. arXiv:1004.2515
Ziv E, Nemenman I, Wiggins CH (2007) Optimal signal processing in small stochastic biochemical networks. PLoS ONE 2:e1077

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.