Analytical computation of the power spectral density for unimolecular stochastic reaction networks

Ankit Gupta * **Mustafa Khammash** *

* *Department of Biosystems Science and Engineering, ETH-Zürich, 4058-Basel, Switzerland.*

Abstract: Stochastic reaction networks model noisy intracellular processes, like gene-expression, where randomness typically arises due to low copy-numbers of the constituent biomolecular species. The frequency spectrum of each single-cell stochastic trajectory generated by such models contains valuable information about the network architecture and the reaction parameters. In this talk we demonstrate how this frequency spectrum can be analytically computed for any unimolecular reaction network under mass-action kinetics. We provide analytical expressions for the *power spectral density* (PSD) for simple three-node feedforward and feedback topologies in terms of the model parameters. Moreover we establish structural results that highlight the key differences between the PSD generated by these topologies irrespective of the model parameters.

Keywords: Stochastic jump processes, stochastic systems, power spectral density, feedback control, feedforward control.

1. INTRODUCTION

Recent advances in time-lapse microscopy has made it possible to generate single-cell trajectories in a high-throughput fashion. These trajectories typically correspond to abundance levels of fluorescent proteins within each cell and hence they exhibit considerable fluctuations due to the stochastic nature of intracellular dynamics. Stochastic models commonly used to describe noisy intracellular processes represent the dynamics of the underlying reaction network as a *continuous-time Markov chain* (CTMC) whose states are the copy-number vectors of the constituent bio-molecular species. Analysing the dynamics of the fluorescent output trajectories under such stochastic models can reveal useful information about the network structure and its biological function. An important strategy to perform this analysis is to compute the *power spectral density* (PSD) of the output signal and obtain insights about the strengths of various frequency components present in the signal.

In this talk we will present a method to compute the PSD for the stochastic model of any reaction network whose propensity functions are affine functions of the state variables. Our approach generalises the results in Jia et al. (2019) and Song et al. (2019), and we use it to analytically compute the PSD for simple three-node feedforward and feedback topologies (see Fig. 1). We then show that these analytical expressions highlight an important difference between the PSD generated by these two topologies and in particular, only the feedback topology can lead to sustained, discernible oscillations in the single-cell trajectories.

2. PRELIMINARIES

In a CTMC model of a reaction network, the state of the system at any time t is the non-negative integer vector $X(t) = (X_1(t), \ldots, X_d(t))$ denoting the molecular counts of d species $\mathbf{X}_1, \ldots, \mathbf{X}_d$ involved in the network. Each reaction k is characterised by a propensity function $\lambda_k(x)$ of the state-vector $x = (x_1, \ldots, x_d)$ that specifies the rate of firing, and an integer vector ζ_k that specifies the state change upon firing of reaction k. Supposing there are K reactions, when the state is x, the next reaction fires after an exponentially distributed random time with rate $\lambda_0(x) := \sum_{k=1}^{K} \lambda_k(x)$ and this firing reaction is k with probability $\lambda_k(x)/\lambda_0(x)$. Henceforth the output signal for each cell will be the molecular count trajectory $(X_n(t))_{t \geq 0}$ for some species $\mathbf{X}_n$.

We shall assume that the CTMC representing a reaction network is *ergodic*, i.e. the probability distribution of the random state $X(t)$ converges to the stationary distribution π as $t \to \infty$. Methods for checking ergodicity can be found in Gupta et al. (2014) and Gupta and Khammash (2018). For ergodic networks, the long-term behaviour of the CTMC can be studied under the assumption that it is always at stationarity, which implies that the means and covariances of the state components do not depend on time. Let $(\tilde{X}_n(t))_{t \geq 0}$ be the mean-zero output signal

$$\tilde{X}_n(t) = X_n(t) - \bar{x}_n,$$

where $\bar{x}_n$ is the stationary mean of the output. We define the truncated one-sided Fourier Transform of the signal $(\tilde{X}_n(t))_{t \geq 0}$ as

$$\mathcal{F}_T(\omega) = \frac{1}{\sqrt{T}} \int_0^T \tilde{X}_n(t) e^{-i\omega t} dt,$$
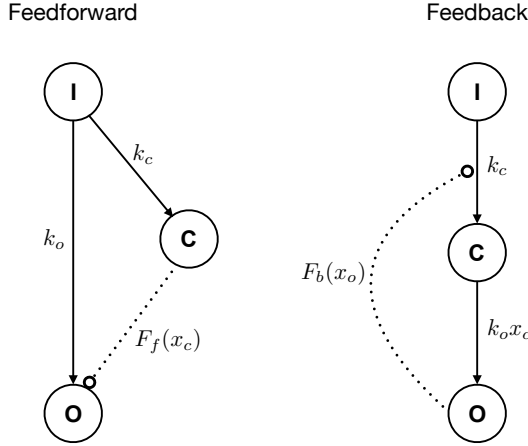
Feedforward   Feedback

Fig. 1. Depicts the feedforward (*left*) and feedback (*right*) topology we consider in our spectral analysis. In the feedforward network, an input **I** directly produces the output **O** at rate k_o and it produces a controller species **C** at rate k_c. The controller species then produces output **O** at rate $F_f(x_c)$ where x_c is the copy-number of **C** and F_f is either a monotonically increasing function (for coherent feedforward) or a monotonically decreasing function (for incoherent feedforward). In the feedback network the output species **O** is only produced by species **C** at rate $k_o x_c$ and the output species **O** produces **C** at rate $F_b(x_o)$ where x_o is the output species copy-number. Here F_b is either a monotonically increasing function (for positive feedback) or a monotonically decreasing function (for negative feedback). In both these networks all arrows denote catalytic production. The input species **I** is assumed to have a constant deterministic copy-number and species **C** and **O** degrade at rates γ_c and γ_o respectively (degradation reactions not shown).

where ω is the frequency and $i = \sqrt{-1}$. The *power spectral density (PSD)* for the process is given by

$$S_{X_n}(\omega) = \lim_{T \to \infty} \mathbb{E}\left(|\mathcal{F}_T(\omega)|^2 \right).$$

Intuitively $S_{X_n}(\omega)$ gives the strength of the component of frequency ω in the signal $(\tilde{X}_n(t))_{t \geq 0}$.
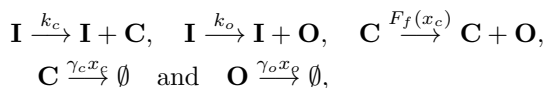
3. MAIN RESULTS

We consider a network where all propensity functions are affine and hence the vector of propensity functions $\lambda(x) = (\lambda_1(x), \dots, \lambda_K(x))$ is described by an affine map

$$\lambda(x) = \Lambda x + b$$

where Λ is some $K \times d$ matrix and b is a $K \times 1$ vector. Under the assumption of ergodicity we provide a method to analytically compute the PSD $S_{X_n}(\omega)$. We now present the results we obtain by applying this method to the two topologies shown in Fig. 1.

Feedforward topology: We can represent the feedforward network shown in Fig. 1 as the following reaction network:

$$\mathbf{I} \xrightarrow{k_c} \mathbf{I} + \mathbf{C}, \quad \mathbf{I} \xrightarrow{k_o} \mathbf{I} + \mathbf{O}, \quad \mathbf{C} \xrightarrow{F_f(x_c)} \mathbf{C} + \mathbf{O},$$

$$\mathbf{C} \xrightarrow{\gamma_c x_c} \emptyset \quad \text{and} \quad \mathbf{O} \xrightarrow{\gamma_o x_o} \emptyset,$$

where the reaction rates are stated above the reaction arrow, and x_c (x_0) denotes the copy-number of species **C** (**O**). The input species **I** does not change its abundance and so its dynamics is not modelled. Here F_f is the feedforward function which we will *linearize* as

$$F_f(x_c) = k + k_{\mathrm{ff}} x_c.$$

The feedforward mechanism can be coherent (if $k_{\mathrm{ff}} > 0$), incoherent (if $k_{\mathrm{ff}} < 0$) or absent (if $k_{\mathrm{ff}} = 0$). Using our method for PSD computation we obtain the following expression for the PSD for this feedforward topology
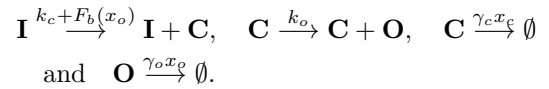
$$S_{X_n}(\omega) = \frac{2}{\gamma_o^2 + \omega^2}\left[k + k_o + \frac{k_{\mathrm{ff}} k_c}{\gamma_c} + \frac{k_{\mathrm{ff}}^2 k_c}{\gamma_c^2 + \omega^2} \right].$$

From this formula it can be shown that for any choice for positive parameters k, k_o, k_c, γ_c, γ_o and any real-valued k_{ff} satisfying

$$k_{\mathrm{ff}} > -\frac{\gamma_c(k_o + k)}{k_c},$$

the mapping $\omega \mapsto S_{X_n}(\omega)$ is monotonically decreasing over the interval $[0, \infty)$. This restriction on k_{ff} is needed to ensure network ergodicity.

Feedback topology: The feedback topology in Fig. 1 as the following reaction network:

$$\mathbf{I} \xrightarrow{k_c + F_b(x_o)} \mathbf{I} + \mathbf{C}, \quad \mathbf{C} \xrightarrow{k_o} \mathbf{C} + \mathbf{O}, \quad \mathbf{C} \xrightarrow{\gamma_c x_c} \emptyset$$

$$\text{and} \quad \mathbf{O} \xrightarrow{\gamma_o x_o} \emptyset.$$

Here F_b is the feedback function which we will *linearize* as

$$F_b(x_o) = k + k_{\mathrm{fb}} x_o.$$

We can have either positive feedback (if $k_{\mathrm{fb}} > 0$), or negative feedback ($k_{\mathrm{fb}} < 0$) or no feedback (if $k_{\mathrm{fb}} = 0$). To ensure ergodicity we must have

$$k_{\mathrm{fb}} < \frac{\gamma_c \gamma_o}{k_o}. \tag{1}$$

Using our method we obtain the following expression for the PSD for this feedback topology

$$S_{X_n}(\omega)$$
$$= \frac{2\gamma_o k_o(k_c + k)(\gamma_c^2 + k_o \gamma_c + \omega^2)}{[\omega^2(\gamma_c + \gamma_o)^2 + (\omega^2 - \gamma_c\gamma_o + k_{\mathrm{fb}}k_o)^2](\gamma_c\gamma_o - k_{\mathrm{fb}}k_o)}.$$

From this formula it can be shown that if the negative feedback is strong enough (i.e. $k_{\mathrm{fb}} < -\mathcal{T}$ for some threshold $\mathcal{T} > 0$) then the mapping $\omega \mapsto S_{X_n}(\omega)$ becomes non-monotonic with a unique maxima at some peak frequency $\omega_{\max} > 0$. We can obtain analytical expressions for $\mathcal{T}$ and $\omega_{\max}$ in terms of parameters k, k_o, k_c, γ_c, γ_o and k_{fb}.

To demonstrate the accuracy of our analytical PSD formulas we simulate a long trajectory for both these topologies, using the *stochastic simulation algorithm* (SSA) in Gillespie (1977), with the parameters $k_o = 1$, $k_c = 0.2$, $k = 50$, $\gamma_c = 2$, $\gamma_o = 0.5$, and $k_{\mathrm{ff}} = k_{\mathrm{fb}} = -3$. We compare the analytically computed PSD with the numerically estimated PSD in Fig. 2 for both the topologies. As one can see, the numerically estimated PSDs are very noisy but they agree well with the analytical PSDs.

Note that both incoherent feedforward and negative feedback topologies can give rise to perfect adaptation (see Ma et al. (2009)), but our results show that in the stochastic setting, irrespective of the parameter values, the incoherent feedforward topology cannot lead to a non-monotonic PSD while the negative feedback topology can, provided
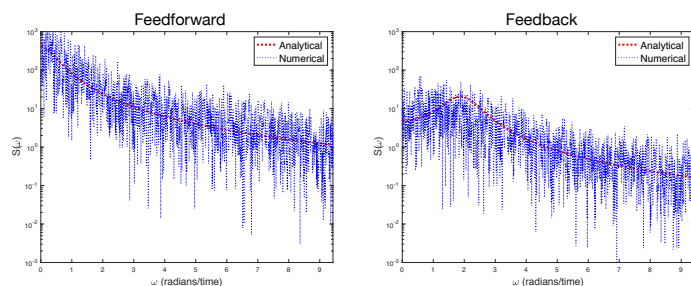
Fig. 2. Comparison of the PSD estimated with a simulated trajectory to the PSD computed with the analytical formula. Note that the PSD for the feedforward topology is monotonic while the PSD for the feedback topology is non-monotonic with a frequency peak $\omega_{\max}$ around 2 radians/time.

that the negative feedback strength is sufficiently high. This key difference may enable differentiating adapting topologies without the need of oscillatory inputs as considered in Rahi et al. (2017).

ACKNOWLEDGEMENTS

REFERENCES

Gillespie, D.T. (1977). Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81(25), 2340–2361.

Gupta, A., Briat, C., and Khammash, M. (2014). A scalable computational framework for establishing long-term behavior of stochastic reaction networks. *PLoS computational biology*, 10(6), e1003669.

Gupta, A. and Khammash, M. (2018). Computational identification of irreducible state-spaces for stochastic reaction networks. *SIAM Journal on Applied Dynamical Systems*, 17(2), 1213–1266.

Jia, C., Zhang, M.Q., and Qian, H. (2019). Analytic theory of stochastic oscillations in single-cell gene expression. *arXiv preprint arXiv:1909.09769*.

Ma, W., Trusina, A., El-Samad, H., Lim, W.A., and Tang, C. (2009). Defining network topologies that can achieve biochemical adaptation. *Cell*, 138(4), 760–773.

Rahi, S.J., Larsch, J., Pecani, K., Katsov, A.Y., Mansouri, N., Tsaneva-Atanasova, K., Sontag, E.D., and Cross, F.R. (2017). Oscillatory stimuli differentiate adapting circuit topologies. *Nature methods*, 14(10), 1010.

Song, S., Yang, G.S., Park, S.J., Hong, S., Kim, J.H., and Sung, J. (2019). Frequency spectrum of chemical fluctuation: A probe of reaction mechanism and dynamics. *PLoS computational biology*, 15(9), e1007356.