Describing function-based approximations of biomolecular systems

Abhishek Dey[1], Shaunak Sen[1] ✉

[1]*Department of Electrical Engineering, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110 016, India*
✉ *E-mail: shaunak.sen@ee.iitd.ac.in*

Abstract: Mathematical methods provide useful framework for the analysis and design of complex systems. In newer contexts such as biology, however, there is a need to both adapt existing methods as well as to develop new ones. Using a combination of analytical and computational approaches, the authors adapt and develop the method of describing functions to represent the input–output responses of biomolecular signalling systems. They approximate representative systems exhibiting various saturating and hysteretic dynamics in a way that is better than the standard linearisation. Furthermore, they develop analytical upper bounds for the computational error estimates. Finally, they use these error estimates to augment the limit cycle analysis with a simple and quick way to bound the predicted oscillation amplitude. These results provide system approximations that can add more insight into the local behaviour of these systems than standard linearisation, compute responses to other periodic inputs and to analyse limit cycles.

1 Introduction

Design of biomolecular systems can enable applications in agriculture, medicine and manufacturing [1]. Complementarily, analysing how naturally occurring biomolecular interactions determine cellular behaviour is a fundamental problem in biology [2]. Mathematical frameworks are useful for both these objectives. These provide system representations to test and compare different design choices as a guide to the actual implementation. These also help to develop useful insight into how system interactions can combine to generate the overall behaviour. Mathematical models used in these cases are typically complex, both due to their large scale as well as the inherent nonlinearity, making their analysis challenging. Therefore, there is a need to adapt existing mathematical methods as well as develop these and new ones for the study of such problems.

Biomolecular systems are frequently represented as ordinary differential equations, formulated based on the principles of mass action. The variables in these equations are the concentrations of various biomolecular species that evolve in time depending on their interactions with each other. One approach to study these mathematical models is exhaustive numerical computations, which can catalogue all possible system behaviours, complemented with simpler calculations to understand the key underlying principles [3, 4]. Another approach is theoretical, exploiting the inherent structure to infer system behaviour. An example is the theory of monotone systems [5]. Intermediary approaches may also exist using various approximations to understand system behaviour. Indeed, methods of linear systems theory have often been used to characterise biomolecular system behaviour such as impulse responses in bacterial chemotaxis [6] as well as frequency responses in the osmolarity pathway of the yeast *Saccharomyces cerevisiae* mediated by the mitogen-activated protein (MAP) kinase cascade [7–9] and in the galactose metabolic pathway, also in the yeast *S. cerevisiae* [10]. Another example of such an intermediary approach is the describing function technique [11, 12], where the frequency response of a nonlinear system to a sinusoidal input of a particular amplitude is approximated by the first harmonic of the resulting output response. This is widely used in classical control engineering to estimate limit cycle behaviour as well as to replace nonlinear input–output responses with corresponding linear approximation. These linear approximations can then be analysed using the well-developed tools of linear

systems theory. In fact, this technique has been applied to analyse biomolecular oscillations [13] and to approximate input–output maps in biomedical contexts [14]. These results present important early work in using this technique for biological systems.

There are at least three striking aspects related to using a describing function-based linearisation to approximate the input–output response of a biomolecular signalling system. One, describing functions naturally allow the analysis of finite amplitude inputs, as opposed to infinitesimal amplitude inputs in the standard linearisation. These may be more relevant to actual biomolecular contexts such as in the experimental studies of frequency response mentioned above [7–10] and provide additional insight into the system behaviour. Two, nonlinearities typically analysed using describing functions in the classical contexts are static nonlinearities such as saturation or hysteresis. Contrastingly, the same nonlinearities in biomolecular contexts can have a dynamic character because of the underlying nonlinear biomolecular interactions are embedded in the overall system dynamics. Three, there may be error involved in the approximation that may depend both on input frequency as well as inherent system parameters, which may be important to quantify. Given these, the describing function-based approximation of these input–output responses including the dependence on system parameters and the nature of approximation error is generally unclear.

Here, we aim to approximate these systems and estimate the resultant error. For this, we used the technique of describing functions, analytically, where possible, as well as computationally. We computed approximations for representative systems with input–output responses exhibiting saturation dynamics with different slopes as well as hysteretic dynamics (formed the basis of preliminary investigation [15]). Next, we computed the approximation error and developed a theoretical error bound for these kinds of systems. Finally, we used these error estimates to augment the classical describing function-based limit cycle analysis by providing a simple way to estimate the range of oscillation amplitudes. These results adapt the existing method of describing functions for the study of biomolecular systems and should be useful both in analysis and design.
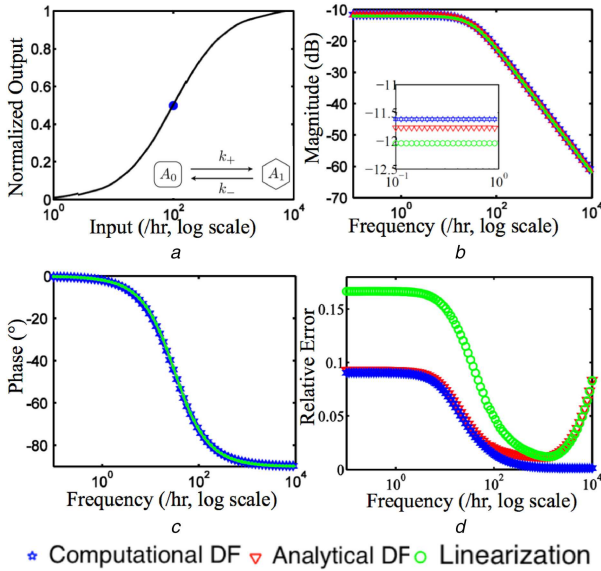
Fig. 1 *Approximation of the simple biomolecular system and error estimate*
(a) Black line is the steady-state input–output response for parameters $A_T = 100\,\text{nM}$, $k_- = 100/\text{h}$, $k_{+0} = k_-$ and $b = k_{+0}/2$. Blue dot is the operating point used for computation of linearisation. Schematic representation of biomolecular system is shown in the inset, *(b)* Red triangle shaped, blue pentagon shaped and green circle shaped markers are the magnitude plots of analytical, computational describing function-based approximations and direct linearisation, respectively, for same parameter set, frequency is varied logarithmically in the range $0.1–10^4/\text{h}$, *(c)* Phase plots corresponding to b, *(d)* Corresponding error plots

2 Computation of describing function approximation

2.1 Describing function method

Consider a nonlinear system governed by equations

$$\dot{x} = f(x, u)$$
$$y = Cx, \tag{1}$$

where x is a vector of species concentrations, u is input which can be a system parameter, C is a row vector and y is the output and typically f is a nonlinear map. We aim to approximate the input–output system from u to y using describing function [11]. To compute the approximation, we set the input to be $u = u_0 + b\sin(\omega t)$, where b is the forcing amplitude, ω is the forcing frequency and u_0 is the input bias. The describing function approximation $G(j\omega, b, u_0)$ can be obtained by taking the first harmonic of the output in response to sinusoidal input. This is defined as

$$\text{Re}\{G(j\omega, b, u_0)\} = \frac{\omega}{\pi b}\int_0^{2\pi} y(t)\sin(\omega t)\,\mathrm{d}t$$

$$\text{Im}\{G(j\omega, b, u_0)\} = \frac{\omega}{\pi b}\int_0^{2\pi} y(t)\cos(\omega t)\,\mathrm{d}t\,. \tag{2}$$

The magnitude and phase of the approximation can be calculated as

$$|G| = \sqrt{\text{Re}\{G\}^2 + \text{Im}\{G\}^2}$$

$$\angle G = \tan^{-1}\frac{\text{Im}\{G\}}{\text{Re}\{G\}} \tag{3}$$

respectively.

Next, we use this technique to approximate canonical biomolecular signalling systems. Here, (1) is an ordinary differential equation model of the system under consideration obtained using the law of mass action. As, biomolecular signals

cannot be negative, we use the constraint $b < u_0$, so that the input is always positive.

2.2 Example 1: biomolecular system with saturating input–output map

One of the simplest signal transduction mechanisms that is representative of diverse signalling contexts is of a biomolecular species that can exist in two states [3, 16, 17]. These can interconvert among each other (Fig. 1a inset) depending on the input level. Typically, one of these states has biological activity and can serve as the output. Consider a simple model of this with a biomolecular species (A) that can interconvert between two forms (A_0 and A_1) at certain forward and reverse rates (k_+ and k_-). In this two-state model, the input such as temperature or pheromone levels can be modelled as modulating the rate k_+ and the output as the concentration of A_1, the active form of the protein A. A mathematical model in the context of biomolecular signal transduction can be obtained using mass action kinetics

$$\frac{\mathrm{d}A_1}{\mathrm{d}t} = k_+(A_T - A_1) - k_- A_1\,. \tag{4}$$

Here, the total concentration of the protein is A_T ($= A_0 + A_1$), which is constant in the simplest scenario considered here. Owing to the presence of the term $k_+ A_1$, where the input term and the output multiply each other, this is a nonlinear equation. The input–output response at steady state can be obtained by setting $\frac{\mathrm{d}A_1}{\mathrm{d}t} = 0 \Rightarrow A_1 = A_T\frac{k_+}{k_+ + k_-}$ (Fig. 1a).

To compute the approximation of the overall response, we set the input to be $k_+ = k_{+0} + b\sin(\omega t)$ and describing function approximation is computed using the output obtained numerically from MATLAB ode23s solver in response to the above input from (2). The results of this computation are shown in Figs. 1b and c.

To obtain an analytical approximation, we set $A_1 = A_{10} + A_{1b}\sin(\omega t + \theta)$ [11] in (4) and collect the like terms to obtain expressions for A_{10}, A_{1b} and θ (the detailed steps are in supplementary material S1)

$$A_{10} = A_T\frac{k_{+0} - \alpha}{k_- + k_{+0} - \alpha} \quad \alpha = \frac{1}{2}b^2\frac{k_{+0} + k_-}{\omega^2 + (k_{+0} + k_-)^2}$$

$$A_{1b} = (A_T - A_{10})\frac{b}{\sqrt{\omega^2 + (k_{+0} + k_-)^2}} \tag{5}$$

$$\theta = -\tan^{-1}\left(\frac{\omega}{k_{+0} + k_-}\right).$$

The describing function approximation is (Figs. 1b and c)

$$G(j\omega, b, k_{+0}) = \frac{A_{1b}}{b}e^{j(\omega t + \theta)}.$$

The analytical and computational results match well converging to the linearised frequency response (Figs. 1b and c) as $b \to 0$

$$\lim_{b \to 0} G(j\omega, b, k_{+0}) = \frac{A_T k_-/(k_- + k_{+0})}{\sqrt{\omega^2 + (k_{+0} + k_-)^2}}e^{j(\omega t + \theta)},$$

with the operating point ($k_+ = k_{+0}$, $A_1 = A_T k_+/(k_+ + k_-)$).

Parametric dependence of approximation: We note that the phase of the approximation is independent of the forcing amplitude b and coincides with that of linearisation. At frequencies lower than crossover frequency (ω_0), the output is in phase, whereas at higher frequencies the output lags the input by $\pi/2$. The magnitude, on the other hand, depends on forcing amplitude b, in an increasing manner (5). This is a counterintuitive result as steady-state solution of (4)

$$A_1 = A_T\frac{k_+}{k_+ + k_-},$$
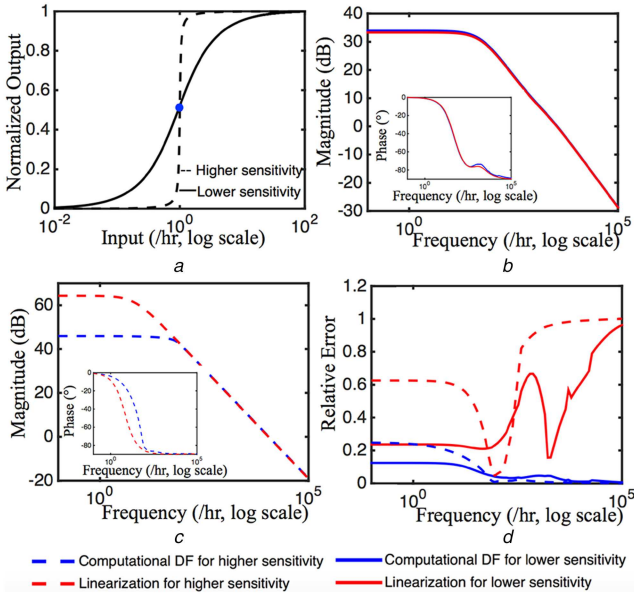
Fig. 2 *Approximation of the biomolecular covalent modification system and error estimates*

(a) Solid-black line represents the steady-state normalised input–output map for parameters $A_T = 200$ nM, $E_{1T} = E_{2T} = 20$ nM, $d_1 = d_2 = 1/\text{h}$, $k_2 = 1/\text{h}$, $k_{10} = k_2$, $b = k_{10}/2$ and $a_1 = a_2 = 10^{-2}/\text{h}$, corresponding to a less steep transition. Dashed-black line represents the steady-state normalised input–output map which corresponds to a steeper transition, with $a_1 = a_2 = 1/\text{h}$ and other parameters are kept constant. Blue dot represents operating point used for computation of linearisation, *(b)* Blue-solid line and red-solid line represent the magnitude plots of computational describing function-based approximation and linearisation approximation of less steep transition, respectively, frequency is varied logarithmically in the range 0.1–$10^4/\text{h}$. Inset shows corresponding phase plots, *(c)* Blue- and red-dashed lines represent magnitude and phase plots for both approximations in steeper transition case, *(d)* Corresponding error involved in both approximations for both sensitivity regimes

is increasing and then saturating function of k_+. However, A_{1b} itself decreases as b is increased. This is because A_{10} decreases as the forcing amplitude b increases (5). Increasing either k_{+0} or k_- increases the crossover frequency ($\omega_0 = k_{+0} + k_-$), making the output phase similar to input phase. In the following two limits:

$$\omega_0 \ll \omega \Rightarrow \frac{A_{1b}}{b} = A_T \frac{k_-}{\omega_0 \omega},$$

$$\omega_0 \gg \omega \Rightarrow \frac{A_{1b}}{b} = A_T \frac{k_-}{\omega_0^2 - b^2/2},$$

magnitude decreases with k_{+0}. However, in the high-frequency limit, magnitude increases with k_-, whereas in the low-frequency limit, the magnitude first increases and then decreases as k_- is increased.

Finally, A_T, the total protein concentration scales the magnitude but does not change the phase.

This presents an overview of how the describing function approximation depends on system parameters. We note that these relations are more accurate than when obtained by standard linearisation and cannot be achieved through numerical simulation.

Approximation error: It is important to quantify the error in any approximation as a measure of its accuracy and to compare different approximation methods. We computed the relative mean-square error

$$e = \frac{1}{A_{1b}} \sqrt{\frac{1}{T} \int_0^T (y(t) - \tilde{y}(t))^2 \, dt}, \tag{6}$$

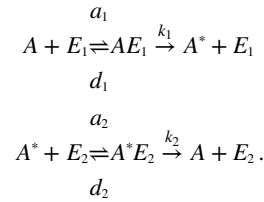where $y(t)$ is the numerical solution of (4) with sinusoidal forcing and

$$\tilde{y}(t) = A_{10} + A_{1b} \sin(\omega t + \theta),$$

which can be found analytically from (5) or can be computed using (2) through computer simulation. A comparison of relative error for different approximations (Fig. 1*d*) shows that the computational describing function-based approximation matches reasonably well with the analytical approximation especially at lower frequencies and is better than the approximation obtained by the standard linearisation.

These results provide a describing function-based approximation of a simple system as well as a quantification of the error involved.

2.3 Example 2: biomolecular covalent modification system with switch-like input–output map

As a second example of a biomolecular signalling system, we investigated a covalent modification scheme that is present in multiple cellular pathways [18]. This is similar to the example considered above, with enzyme-catalysed interconversion reactions. Depending on parameter regime of operation, the steady-state input–output response may have different sensitivities. The reaction scheme for this system is

$$A + E_1 \underset{d_1}{\overset{a_1}{\rightleftharpoons}} AE_1 \overset{k_1}{\rightarrow} A^* + E_1$$

$$A^* + E_2 \underset{d_2}{\overset{a_2}{\rightleftharpoons}} A^*E_2 \overset{k_2}{\rightarrow} A + E_2.$$

The forward and backward conversions between the two forms A and A^* are catalysed by enzymes E_1 and E_2, respectively. As such, these reactions have intermediates $[AE_1]$ and $[A^*E_2]$, giving overall conservation relations as

$$A_T = [A] + [A^*] + [AE_1] + [A^*E_2]$$
$$E_{1T} = [E_1] + [AE_1]$$
$$E_{2T} = [E_2] + [A^*E_2],$$

where A_T is the total substrate concentration and E_{1T} and E_{2T} are the total enzyme concentrations. Using these the mathematical model can be obtained as

$$\frac{d[A]}{dt} = -a_1[A](E_{1T} - [AE_1]) + d_1[AE_1]$$
$$\qquad + k_2(A_T - [A] - [A^*] - [AE_1])$$

$$\frac{d[A^*]}{dt} = -a_2[A^*](E_{2T} - A_T + [A] + [A^*] + [AE_1]) \tag{7}$$
$$\qquad + d_2(A_T - [A] - [A^*] - [AE_1]) + k_1[AE_1]$$

$$\frac{d[AE_1]}{dt} = a_1[A](E_{1T} - [AE_1]) - (d_1 + k_1)[AE_1].$$

As before, we model k_1 as the input and the concentration of A^* as the output. The input–output response at steady state can have different sensitivities depending on the parameter regime of operation. Two such choices with a low sensitivity and a high sensitivity are shown in Figs. 2*b* and *c*. The nonlinearities are distributed through the system and arise due to the principles of mass action. We aim to approximate the overall input–output response with a describing function-based linearisation. Therefore, while this is similar to a saturation nonlinearity [12], there are inherent dynamics.

Calculation of approximation: The describing function approximation is obtained computationally (Fig. 2*b* – blue-solid line and Fig. 2*c* – blue dashed line), as described previously. We also computed, for comparison, the standard linearisation (Fig. 2*b* – red-solid line and Fig. 2*c* – red-dashed line).
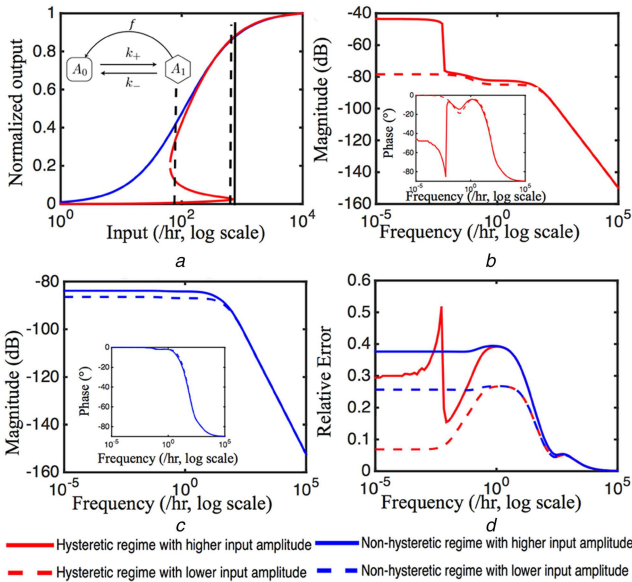
Fig. 3 *Describing function approximation and error estimate for hysteretic system*

(a) Solid-red line represents the steady-state normalised input–output map for parameters $\alpha_0 = 1/15$, $\alpha = 5\,\text{nM/h}$, $n = 2$, $\gamma = 1/\text{h}$, $K = 1\,\text{nM}$, $k_- = 100/\text{h}$, corresponding to a hysteretic response. Solid-blue line represents the steady-state normalised input–output map when $\alpha = 1$ with other parameters same as above. Solid-black vertical line corresponds to the upper limit of the area where describing function approximation is done for higher-input amplitude $k_{+0} = b = 375$. Dashed-black vertical lines correspond to the upper and lower limits of the area where describing function approximation is done for lower-input amplitude, $k_{+0} = 375$, $b = 300$. Schematic representation of the system is shown in inset, *(b)* Red-solid line and dashed line represent the magnitude plot of computational describing function-based approximation for hysteretic response with higher and lower-input amplitudes, respectively, and inset shows phase plot, frequency is varied logarithmically in the range 10^{-5}–10^5/h, *(c)* Blue-solid and dashed line show the magnitude and phase plots for non-hysteretic regime with higher and lower-input amplitudes, respectively, *(d)* Corresponding error associated with the approximation in both regimes for both higher and lower-input amplitudes

Parameter dependence: The describing function-based approximations in both regimes have a similar low-pass nature. The DC gain for the higher sensitivity regime (Fig. 2c) is higher than that of the lower sensitivity regime (Fig. 2b), consistent with the slope of the operating point. For the low sensitivity regime, the describing function-based approximation is similar to that of the standard linearisation. Interestingly, however, the DC gain in the higher sensitivity regime is lower than that of the corresponding standard linearisation, while the bandwidth is larger. This shows that, for finite inputs, the gain may not be as high as the standard linearisation predicts. Similarly, the bandwidth may not decrease to a large extent. Therefore, analysis of the describing function-based approximation provides insight into how the system may behave in case of finite inputs. We emphasise that this realisation is not possible through numerical simulations and the describing function technique provides a better approximation than standard linearisation.

Approximation error: On the basis of the computation (Fig. 2d), it can be observed that the error in the describing function-based approximation is lower than that obtained from the standard linearisation in both parameter regimes considered.

2.4 Example 3: biomolecular system with hysteretic input–output map

Another example of a biomolecular input–output response, often encountered in developmental contexts, is hysteresis. One of the simplest ways through which this can be achieved is, through the addition of transcriptional positive feedback in the circuit showed in Example 1. The resulting all-or-none behaviour has been experimentally observed during maturation of *Xenopus* oocytes

[19, 20]. Here, we consider a simple system to illustrate the hysteretic input–output map (Fig. 3a inset)

$$\frac{d[A_1]}{dt} = k_+(A_T - A_1) - \gamma A_1 - k_- A_1$$

$$\frac{d[A_T]}{dt} = f(A_1) - \gamma A_T, \tag{8}$$

where the feedback is $f(A_1) = \alpha_0 + ((\alpha(A_1)^n)/(K^n + (A_1)^n))$ and there is the conservation law $A_T = A_0 + A_1$. The addition of the nonlinear ultrasensitive ($n > 1$) positive feedback can generate a hysteretic response (Fig. 3a). In the hysteretic regime, there are three steady states, two of which are stable and one is unstable. As input parameter k_+ is varied, a pair of stable–unstable steady states coalesce causing a monostable high or low steady state. Our aim is to develop an approximation of the entire input–output map from k_+ to A_1. While the hysteretic map looks like the classical ones analysed, the difference here is that there are inherent dynamics in the hysteresis nonlinearity.

Calculation of approximation: The describing function approximation is computed by setting the input, k_+, as a biased sinusoid and taking the first harmonic approximation of the output A_1. We analysed parameter regimes where hysteretic response is present as well as the one in which it is not present (Figs. 3b and c). In fact, due to the presence of multiple steady states in the hysteretic regime, it is unclear which point to take for the standard linearisation. For the describing function-based approximation, however, there is no such ambiguity.

Parameter dependence: The frequency response shows low-pass filter nature in both regimes. We note that the frequency response in the hysteretic regime can have a low-frequency part, where the magnitude is constant and the phase is decreasing (Fig. 3b – solid-red line). This is exactly the characteristic associated with a pure delay (transfer function $H(\text{j}\omega) = \text{e}^{-\text{j}\omega T} \Rightarrow |H(\text{j}\omega)| = 1$, $\angle H(\text{j}\omega) = -\omega T$). Therefore, this method of approximating a positive feedback loop adds insight to system analysis by providing a direct map to a pure delay as well as an estimate of the quantitative parameters associated with it. This is consistent with other studies relating positive feedback with a delayed action, for example [21]. With lower-input amplitude, we still observe two humps in the magnitude and phase plot (Fig. 3b), possibly owing to the presence of two time scales – that of protein production and covalent modification. An attenuated version of this trend is also visible in the case when hysteretic response is absent (Fig. 3c). Although, the higher- and lower-input amplitude responses do not have much difference when there hysteresis is not expected (Fig. 3c).

Approximation error: We calculated the error involved (Fig. 3d) in approximation for the hysteretic input–output map for different regimes and different input amplitudes chosen in the input–output plot of Fig. 3a. The error is higher for higher-input amplitude in the hysteretic regime owing to the presence of strong nonlinearity.

3 Analytical error bound

As with any approximation method, it is important to estimate the associated error. As the describing function technique essentially approximates a dynamical system by the first harmonic of its output, when a sinusoidal input is applied, we develop an error bound using methods associated with Fourier series. The Fourier series of a periodic signal is

$$f(t) = a_0 + \sum_{n=1}^{\infty} (a_k \cos n\omega_0 t + b_k \sin n\omega_0 t).$$

This can also be represented in exponential form

$$f(t) = \sum_{n=-\infty}^{\infty} C_n \text{e}^{\text{j}n\omega_0 t} \quad \omega_0 = \frac{2\pi}{T},$$

where
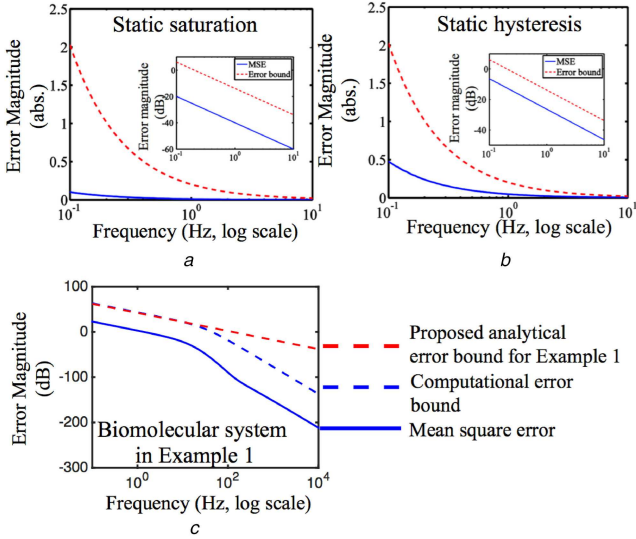
Fig. 4 *Error in describing function approximation and corresponding error bounds*

(a). Blue-solid line and red dotted line represents mean-squared error in describing function approximation and corresponding error bound, respectively, for static saturation, when ω is varied. The same is shown in decibels scale in the inset, *(b)* Similar plots for static hysteresis case, *(c)* For the input–output map in Example 1, solid-blue line represents the mean-square error in the describing function approximation. Red-dashed line represents the theoretical error bound when total variation is taken as $A_T/2$. Dashed-blue line represents the bound when total variation is calculated computationally

$$C_n = \frac{1}{T} \int_{-T/2}^{T/2} f(t) e^{-jn\omega_0 t} \, dt \, . \tag{9}$$

From the convergence of Fourier series, we know that signals that are continuous or have a finite and bounded discontinuity over a period satisfy [22]

$$\lim_{m \to \infty} \| f - f_m \| = 0,$$

where $f_m(t) = \sum_{n=-m}^{m} C_n e^{jn\omega_0 t}$ is the Fourier series taken up to mth harmonic. For the describing function case, we need to estimate

$$\epsilon_1 = \| f - f_1 \| = \int_{-\pi}^{\pi} |f - f_1|^2 \, dt,$$

which is basically a scaled version of the error defined in (6). For Fourier series, an upper bound of this error can be obtained using the principle of bounded variation [23].

Definition 1: [24]: Let $f : [a, b] \to \mathbb{R}$ be a function and $M = \{x_0, x_1, \ldots, x_n\}$ be a partition of $[a, b]$. We define the total variation of f over the $[a, b]$ as

$$V_a^b(f) = \sup_M \left\{ \sum_{k=0}^{n-1} |f(x_{k+1}) - f(x_k)| \right\},$$

where the supremum is taken over all partitions of f. The function is said to have bounded variation if $V_a^b(f)$ is finite in $[a, b]$ and we write $f \in V[a, b]$.

The following properties are useful in calculating the total variation [24]:

i. If $f : [a, b] \to \mathbb{R}$ is monotone in $[a, b]$, then $f \in V[a, b]$ and $V_a^b(f) = |f(b) - f(a)|$.

ii. Let $f : [a, b] \to \mathbb{R}$ be of bounded variation in $[a, b]$. Then $cf \in V[a, b]$ for any $c \in \mathbb{R}$ and $V_a^b(cf) = |c| V_a^b(f)$.

iii. Let $f : [a, b] \to \mathbb{R}$ and c is an arbitrary point in $[a, b]$. Then $f \in V[a, b]$ if and only if $f \in V[a, c]$ and $f \in V[c, b]$. Furthermore, $V_a^b(f) = V_a^c(f) + V_c^b(f)$.

Theorem 1: [23]: If $f(t)$ is periodic with frequency ω_0 and the total variation over one period is bounded by V, then the mean-square error in approximating up to mth harmonic is

$$\epsilon_m^2 \leq \frac{V^2}{\pi \omega_0 m} \, .$$

We use this result to develop an error bound for the describing function approximation. We apply this to classically known static nonlinearities and then illustrate how this can be applied to the dynamic nonlinearities such as the biomolecular systems discussed in Section 2.

3.1 Example 4: static saturation nonlinearity

Consider a saturation nonlinearity with a and k denoting the range and slope of saturation. To obtain the describing function approximation, we use a sinusoidal input $A \sin \omega t$ to the nonlinearity and calculate the first harmonic [12] of the output. The output of the nonlinearity is described as

$$f(t) = \begin{cases} kA \sin \omega t & \text{if } 0 \leq \omega t \leq \delta \\ ka & \text{if } \delta < \omega t \leq \frac{\pi}{2} \end{cases} \text{ where } \delta = \sin^{-1}\left(\frac{a}{A}\right)$$

Using the properties of bounded variation

$$V_0^\delta(f) = f(\delta) - f(0) = kA \sin \delta - 0 = ka,$$

$$V_\delta^{\pi-\delta} = 0,$$

$$V_{\pi-\delta}^\pi = ka \, . \Rightarrow V_0^{2\pi} = 4ka \, .$$

In fact, for any real periodic signal, the total variation over a period is bounded and is four times of the amplitude of the signal. Therefore

$$\epsilon_1^2 = \| f - f_1 \|_2^2 \leq \frac{(4ka)^2}{\pi\omega(1)},$$

for approximation up to first harmonic. We find that the calculated mean-squared error in the first harmonic approximation is bounded by the theoretical error bound for static saturation nonlinearity as frequency is varied (Fig. 4a).

3.2 Example 5: static hysteresis nonlinearity

A relay exhibiting hysteresis is another commonly encountered static nonlinearity [12]. Parameters δ and D denote the hysteretic angle and the range of saturation, respectively. It works like a simple relay with a phase shift of δ. The error bound in this case is $\epsilon_1^2 = \| f - f_1 \|_2^2 \leq \frac{(4D)^2}{\pi\omega(1)}$. Fig. 4b shows the mean-square error between actual output and the first harmonic approximation, which is bounded by theoretical error bound for all ω. It can be noted that the error bound for both cases are similar if the maximum magnitude of the output is set to be same in both cases, i.e. $ka = D$. However, it is observed that the maximum mean-squared error for the relay with hysteresis is more than that of static saturation nonlinearity. The maximum mean-squared error increases for saturation case if the slope of saturation increases and tends to the error of relay with hysteresis when k approaches ∞.
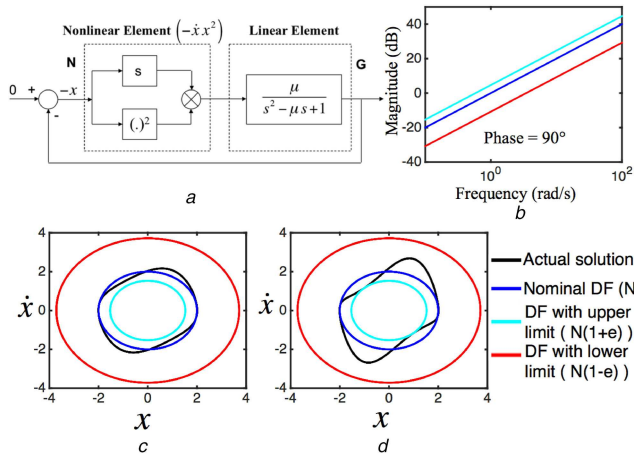
Fig. 5 *Limit cycle analysis of Van der Pol oscillator using describing function technique*

(a) Block diagram of Van der Pol oscillator. *(b)* Solid-blue line represents nominal describing function approximation, whereas upper and lower limits for the approximation is represented by cyan and red lines, respectively, *(c)* Black line represents the actual solution for Van der Pol oscillator. Blue, red and cyan line correspond to the solution considering nominal, upper and lower limits of describing function approximation for $\mu = 0.4$, *(d)* Same is repeated for $\mu = 1$

3.3 Example 1: continued

The nonlinearities encountered in biomolecular systems often have a dynamic character, as analysed in the previous section. To find an error bound for these kinds of nonlinearities, we first need to determine the total variation. We illustrate this for the first example, where we bound the maximum variation for the output A_1 to be A_T, the total concentration. Furthermore, the steady-state input–output response for static saturation case does not take negative values. Comparing these, we find that the total variation of periodic output in this case is $V = A_T/2$ and the error bound is $V^2/\pi\omega$. Computationally, we find the total variation from the time trajectory of simulated output. The theoretical as well as computational error bounds for describing function approximation are shown in Fig. 4c. Both of these bound the mean-square error of describing function approximation.

These results provide an error bound for describing function approximation using the concept of total variation and a general way, using the maximum allowable concentration of the output species, obtain these error bounds for biomolecular systems.

4 Error estimates in limit cycle analysis

Finally, we investigate the use of the error analysis developed above in the analysis of limit cycles, a classically important application of the describing function technique. Although this technique is widely used in control engineering for limit cycle prediction, a treatment of errors involved is relatively less common [25]. Here, we aim to incorporate the simple error estimate developed above to augment the standard analysis.

4.1 Example 6: Van der Pol oscillator

We begin with the Van der Pol oscillation, a benchmark nonlinear oscillator and commonly used to illustrate the describing function-based limit cycle analysis [12]. This oscillator was first analysed in the context of vacuum tube triode circuits and termed as 'relaxation oscillation' by Van der Pol [26]. This class of oscillators has multiple applications in physical as well as biological sciences such as in the Fitzhugh–Nagumo oscillator [27, 28]. The dynamical equations of Van der Pol oscillator are

$$\ddot{x} + \mu(x^2 - 1)\dot{x} + x = 0 \tag{10}$$

This can be separated in a linear and a nonlinear part as shown in Fig. 5a.

Describing function approximation with error estimate: We computed the describing function of the nonlinear part (Fig. 5b). This is denoted as N. We also computed the error (e) in this approximation to get an uncertainty model $N(1 \pm e)$, where the $+$ and $-$ signs give, respectively, the upper and lower limits. Next, we used the condition for sustained oscillation, $1 + NG = 0$, where G is the transfer function of the linear part of the loop. We did this for all three cases – nominal (N), upper limit [$N = N(1 + e)$], lower limit [$N = N(1 - e)$] – to get the corresponding values of the oscillation amplitude (A). The Nyquist plot corresponding to the nominal describing function approximation touches the $-1 + j0$ point at $\omega = 1$ rad/s for $A = 2$. This is consistent with the classical describing function analysis that predicts an oscillation frequency of $\omega = 1$ rad/s and an oscillation amplitude of 2. Incorporating the error information [$N(1 \pm e)$] in the condition of oscillation, we obtain two other values of oscillation amplitude, $A = 1.529$ and 3.718. These suggest that the actual oscillations might lie in this uncertainty band.

To check this, we computed the actual solution of the Van der Pol oscillator along with the solution obtained from the nominal as well as the error analysis for $\mu = 0.4$ and $\mu = 1$ (Figs., 5c and d). At lower values of μ, the actual solution matches the nominal describing function prediction, though the extent of deviation increases as μ increases. In fact, the upper and lower amplitudes provide bounds for the variation of the actual amplitude over a period for these parameter choices. Therefore, these error estimates can be used in predicting an uncertainty band of oscillation amplitudes, providing a larger range of parameters where the describing function approximation can help in the limit cycle analysis.

We note that [25] has also performed a general error calculation of limit cycle, which in this case yields error bounds on amplitude and frequency for a larger range of parameter values. Their approach is based on analysing approximation error and its propagation in the entire closed loop. Here, we focus on the bounds simply obtained due to the uncertainty representation of the nonlinearity.

4.2 Example 7: biomolecular ring oscillator

The repressilator is a biomolecular oscillator consisting of three proteins that repress each other in a ring [29]. It is a benchmark oscillator in the field of synthetic biology – design using biomolecular substrates – for its demonstration of engineering rich dynamic behaviour inside cells. The dynamical equations for the repressilator are

$$\frac{dm_1}{dt} = -\gamma_m m_1 + \alpha_0 + \frac{\alpha}{1 + (p_3/K)^n}$$

$$\frac{dp_1}{dt} = -\gamma_p p_1 + \beta m_1$$

$$\frac{dm_2}{dt} = -\gamma_m m_2 + \alpha_0 + \frac{\alpha}{1 + (p_1/K)^n}$$

$$\frac{dp_2}{dt} = -\gamma_p p_2 + \beta m_2 \tag{11}$$

$$\frac{dm_3}{dt} = -\gamma_m m_3 + \alpha_0 + \frac{\alpha}{1 + (p_2/K)^n}$$

$$\frac{dp_3}{dt} = -\gamma_p p_3 + \beta m_3.$$

The form of these equations and the nominal parameters are taken from [29]. Briefly, m_i and p_i ($i = 1, 2, 3$) represent the concentrations of mRNA and protein, respectively. For simplicity, the production and degradation parameters are assumed to be same for all three mRNA–protein pairs. γ_p and γ_m represent the degradation rate constants for protein and mRNA, respectively. β is the translation rate constant. The repression function for the ith mRNA ($\alpha_0 + (\alpha/(1 + p_j/K)^n))$, models the repression of m_i as p_j increases ($j = 2$ for $i = 1$, $j = 3$ for $i = 2$ and $j = 1$ for $i = 3$). This is a
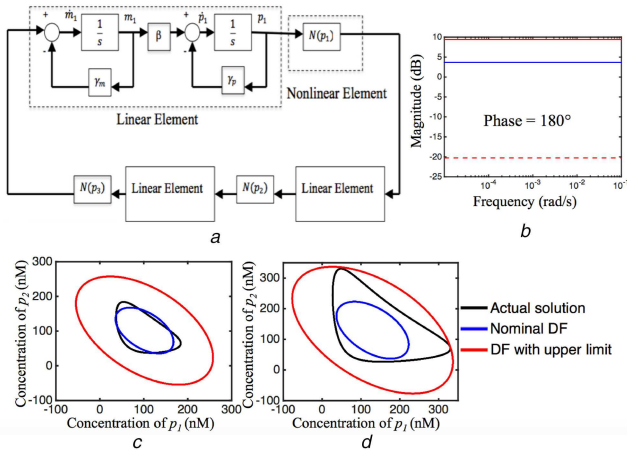
Fig. 6 *Limit cycle analysis of repressilator using describing function technique*

(a) Block diagram of repressilator. *(b)* Solid-blue line represents nominal describing function approximation, whereas upper and lower limits for the approximation is represented by solid-red and dashed-red line, respectively, for parameter values $\alpha = 0.4995 \text{ nM/s}$, $\alpha_0 = 5 \times 10^{-4} \text{ nM/s}$, $\gamma_m = \log 2/120 \text{ s}^{-1}$, $\gamma_p = \log 2/600 \text{ s}^{-1}$, $\beta = \log 2/6 \text{ s}^{-1}$, $K = 40 \text{ nM}$ and $n = 2$, *(c)* Black, blue and red lines represent the actual solution, nominal describing function solution and describing function solution considering the upper error limit for repressilator when $\gamma_p = 20 \log 2/600$ and other parameters are same as b, *(d)* Same is repeated for $\gamma_p = 14 \log 2/600$

special case of a general cyclic gene regulatory network which may give rise to oscillations and the existence criteria and general oscillation profiles are given in [30] using harmonic balance analysis. To apply the describing function technique and associated error analysis, we separate the system to linear and nonlinear parts (Fig. 6*a*). There are three transcriptional modules, each of which has one linear and one nonlinear part which is $F(p) = \alpha_0 + \{\alpha/[(1 + p_j/K)^n]\}$ $(j = 1, 2, 3)$. The gain of the describing function approximation for a single nonlinear element is denoted as N. The input to each nonlinear part is concentration of a protein and thus it is set to a biased sinusoidal input, $p_i = a + b \sin(\omega t)$. Unlike the previous case, the input has a bias part. This is because the concentrations in biomolecular systems concentrations cannot be negative ($b \le a$).

4.2.1 Calculation of frequency and describing function:
From the condition of sustained oscillation

$$1 + N^3 G^3(s) = 0, \tag{12}$$

where $G(s) = (1/((1 + T_{\gamma_m}s)(1 + T_{\gamma_p}s)))$ is the linear element in one-transcriptional module. From this we get one magnitude and one angle condition

$$|(1 + T_{\gamma_m}s)(1 + T_{\gamma_p}s)| = N$$
$$3\angle[(1 + T_{\gamma_m}s)(1 + T_{\gamma_p}s)] = \pi, \tag{13}$$

where $T_{\gamma_m} = (1/\gamma_m)$ and $T_{\gamma_p} = (1/\gamma_p)$. This yields

$$\omega = \frac{-\left(T_{\gamma_m} + T_{\gamma_p}\right) \pm \sqrt{\left(T_{\gamma_m} + T_{\gamma_p}\right)^2 + 4T_{\gamma_m}T_{\gamma_p}\tan^2(\pi/3)}}{2T_{\gamma_m}T_{\gamma_p}\tan(\pi/3)}, \tag{14}$$

and

$$N = \sqrt{\left(1 - T_{\gamma_m}T_{\gamma_p}\omega^2\right)^2 + \omega^2\left(T_{\gamma_m} + T_{\gamma_p}\right)^2} \tag{15}$$

4.2.2 Calculation of the input bias:
The simulation suggests that there is no unique input bias and amplitude pair (a, b) which gives rise to sustained oscillation. So, we need to fix the input bias. Considering the first transcription module in (11), the nonlinear part can be approximated to, $F(p_3) = N_0 - N p_3$. N_0 is the bias term

in the approximation and the negative sign suggests 180° phase for describing function approximation (Fig. 6*b*). N is calculated from (15). Now, putting $p_3 = a + b \sin(\omega t + \delta)$ and $m_1 = a_m + b_m \sin(\omega t + \phi)$, we get

$$\omega b_m \cos(\omega t + \phi) = N_0 - N(a + b \sin(\omega t + \delta))$$
$$- \gamma_m(a_m + b_m \sin(\omega t + \delta)),$$

and

$$\omega b \cos \omega t = \beta(a_m + b_m \sin(\omega t + \phi)) - \gamma_p(a + b \sin \omega t)$$

Equating the constant terms we get

$$N_0 = \left(N + \frac{\gamma_m \gamma_p}{\beta}\right)a, \tag{16}$$

which gives an equation of straight line, where every point corresponds to the condition of sustained oscillation. Now, from simulation also we can compute the value of N_0 when input bias a is varied which produces a family of curves with varying dependence of input bias a on input amplitude b. The intersection points correspond to the value of input bias. The simulation results are closest to the intersection points found by this method for $a = b$, when $n = 2$.
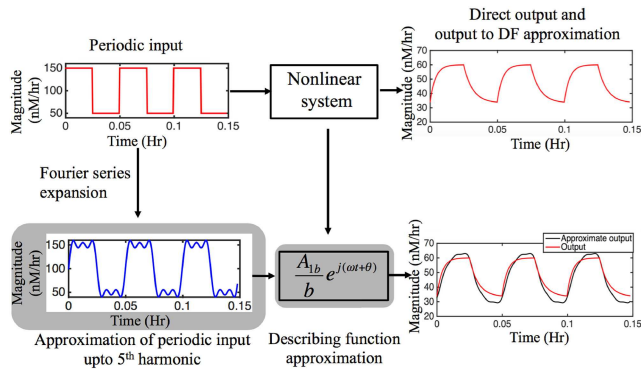
4.2.3 Describing function approximation with error estimate:
The magnitude of the describing function approximation with the upper and lower error limits is shown in Fig. 6*b*. We find the frequency of oscillation from (14) and this matches with simulation result. The describing function gain is calculated from (15) and corresponding input bias is calculated from (16). From the sustained oscillation condition, we get the nominal amplitude of oscillation. The upper limit of the error gives an upper limit of amplitude. The lower limit of error does not give a clear result as the error magnitude is larger than the nominal value. As we have taken same production and degradation constants for all three transcription module, we assume that the phase difference between each protein concentration profile is $2\pi/3$. The actual simulation of the repressilator system is compared with the oscillation profiles obtained from this analysis in Figs. 6*c* and *d* for different parameter values. We conclude that such as in the case of the Van der Pol oscillator, the use of the describing function technique with the simple error model adds to the limit cycle analysis by providing an estimate of the variation in the oscillation amplitude.

5 Discussion

Describing functions can be used to approximate a nonlinear input–output map with its linearisation. Here, we have adapted this method for investigations of biomolecular systems and presented the following three results. First, we used this technique to approximate representative input–output responses, both saturating and hysteretic, and mapped the dependence of the approximation on system parameters. Second, we estimated the approximation error, which was smaller than that of a standard linearisation, and theoretically developed a way to obtain an upper bound of this error. Third, we used the computed error estimate to augment the standard limit cycle prediction by providing a simple way to estimate range of oscillation amplitudes possible. These results should help to augment a framework for approximating biomolecular signalling systems with linearised versions.

It is interesting to note the additional insight that can be obtained by contrasting the describing function-based linearisation with the standard linearisation. For example, in the high sensitivity regime of the enzymatic signalling system (Fig. 2*c*), standard linearisation at the operating point would point to a dynamic response which is very slow and has a high gain. However, an analysis using a finite amplitude input, as in the describing function approximation here, adds insight to the fact that the slow response is present only if the amplitude is infinitesimal. For more realistic inputs, where the amplitude may be finite, the response is not

Fig. 7 *Application of describing function technique in estimating output to square wave input for the system in Fig. 1a for ω/2π = 20.0923 cycles/h. Square wave input is approximated up to the fifth harmonic and each harmonic is treated as a sinusoidal input for the use of describing function. Output to each harmonic summed together to get the approximate output and the results are compared*

significantly slower than in the low sensitivity regime (Fig. 2*b*), though the gain is still higher. This provides a holistic understanding of this signalling system. In case of positive feedback systems, describing function approximation gives an account of the delay (Fig. 3*b*) associated with it in the hysteretic regime. In the presence of multiple steady states, when standard linearisation is not well-defined, describing function approximation gives a quantitative measure of this delay and can further aid in the design of biomolecular oscillators based on this.

An important task for the future is to obtain describing function approximations for different systems, both for building blocks such as analysed over here, as well as for analysis of larger systems obtained by combining such smaller systems. Furthermore, it may be useful to compare these representations with relatively recent experimental data on the response of such systems to periodic input [7–10]. For example, we have computed describing function approximation for the input–output map of the galactose metabolism pathway in *S. cerevisiae* using a model described in [10] consisting of 18 differential equations. This is reported in supplementary material S2. We find that for this case also describing function serves as a better approximation than standard linearisation and can aid the analysis of frequency response data obtained experimentally. It may also be useful to analyse such complex networks by dividing them into multiple stages and replacing each with their linear approximations. Such a cascade network where each stage is a two component signalling system is analysed in supplementary material S3. We have shown how each stage can be approximated with a describing function-based approximation and the overall approximation has lower error than standard linearisation. Similar approach can be adopted for larger biomolecular systems with the aim of enhancing ease of analysis. Finally, the use of random inputs to obtain such describing function approximations may help in treating biomolecular noise that can be potentially important [31].

A mathematical framework can help to understand the behaviour of large complex systems and as a tool for design. Additionally, an appreciation of the limitations of the framework undeniably aids these aims. Here, we have adapted the method of describing functions for approximating biomolecular systems and estimated the corresponding error. These can be also used, for example, in estimating the output response to other periodic inputs such as square waves (Fig. 7). To illustrate, we compare the output with a square wave to the one obtained by decomposing the square wave into three (fundamental, third and fifth harmonics) Fourier components and then adding the output response of these components as predicted by the describing function. The output achieved through these two approaches is reasonably similar. Finally, it should aid in developing a framework for analysis and design of larger, more complex biomolecular systems through systematic interconnections of smaller components.

7 References

[1] Punrick, P.E.M., Weiss, R.: 'The second wave of synthetic biology: from modules to systems', *Nat. Rev. Mol. Cell Biol.*, 2009, **10**, (6), pp. 410–422
[2] Hartwell, L.H., Hopfield, J.J., Leibler, S.*, et al.*: 'From molecular to modular cell biology', *Nature*, 1999, **402**, (6761 Suppl), pp. C47–C52
[3] Huang, C.Y., Ferrell, J.E.: 'Ultrasensitivity in the mitogen-activated protein kinase cascade', *Proc. Natl. Acad. Sci. USA*, 1996, **93**, (19), pp. 10078–10083
[4] Barkai, N., Leibler, S.: 'Robustness in simple biochemical networks', *Nature*, 1997, **387**, (6636), pp. 913–917
[5] Sontag, E.D.: 'Monotone and near-monotone biochemical networks', *Syst. Synth. Biol.*, 2007, **1**, (2), pp. 59–87
[6] Block, S.M., Segall, J.E., Berg, H.C.: 'Impulse responses in bacterial chemotaxis', *Cell*, 1982, **31**, (1), pp. 215–226
[7] Hersen, P., McClean, M.N., Mahadevan, L.*, et al.*: 'Signal processing by the hog map kinase pathway', *Proc. Natl. Acad. Sci. USA*, 2008, **105**, (20), pp. 7165–7170
[8] Mettetal, J.T., Muzzey, D., Gómez-Uribe, C.*, et al.*: 'The frequency dependence of osmo-adaptation in *Saccharomyces cerevisiae*', *Science*, 2008, **319**, pp. 482–484
[9] Muzzey, D., Oudenaarden, A.V.: 'Quantitative time-lapse fluorescence microscopy in single cells', *Annu. Rev. Cell Dev. Biol.*, 2009, **25**, pp. 301–327
[10] Bennett, M.R., Pang, W.L., Ostroff, N.A.*, et al.*: 'Metabolic gene regulation in a dynamically changing environment', *Nature*, 2008, **454**, (7208), pp. 1119–1122
[11] Gelb, A., Vander Velde, W.E.: '*Multiple-input describing functions and nonlinear system design*' (McGraw-Hill Electronic Sciences Series McGraw-Hill, New York, 1968)
[12] Slotine, J.J.E., Li, W.: '*Applied nonlinear control*' (Prentice-Hall, Englewood Cliffs, NJ, 1991)
[13] Wang, Y., Hori, Y., Hara, S.*, et al.*: 'Collective oscillation period of intercoupled biological negative cyclic feedback oscillators', *IEEE Trans. Autom. Control*, 2015, **60**, (5), pp. 1392–1397
[14] Kinnane, O., Ringwood, J., Kelly, D.*, et al.*: 'Describing function approximation for biomedical engineering applications'. Irish Signals and Systems Conf., 2004. pp. 107–112
[15] Dey, A., Sen, S.: 'Describing function-based approximations of biomolecular signalling systems'. Linz, Austria: 14th European Control Conf., 2015, pp. 2292–2297
[16] Detwiler, P.B., Ramanathan, S., Sengupta, A.*, et al.*: 'Engineering aspects of enzymatic signal transduction: photoreceptors in the retina', *Biophys. J.*, 2000, **79**, pp. 2801–2817
[17] Stock, A.M., Robinson, V.L., Goudreau, P.N.: 'Two-component signal transduction', *Annu. Rev. Biochem.*, 2000, **69**, pp. 183–215
[18] Goldbeter, A., Koshland, J.D.E: 'An amplified sensitivity arising from covalent modification in biological systems', *Proc. Natl. Acad. Sci. USA*, 1981, **78**, (11), pp. 6840–6844
[19] Ferrell, J.E., Xiong, W.: 'Bistability in cell signalling: how to make continuous processes discontinuous, and reversible processes irreversible', *Chaos*, 2001, **11**, (1), pp. 227–236
[20] Xiong, W., Ferrell, J.E.: 'A positive-feedback-based bistable 'memory module' that governs a cell fate decision', *Nature*, 2003, **426**, pp. 460–465
[21] Trotta, L., Sepulchre, R., Bullinger, E.: 'Delayed decision-making in bistable models'. IEEE 49th Annual Conf. Decision and Control, 2010, pp. 816–821
[22] Oppenheim, A.V., Willsky, A.S., Young, I.T.: '*Signals and systems*' (Prentice-Hall Signal Processing Series Prentice-Hall, Englewood Cliffs, New Jersey, 1983)
[23] Giardina, C.R., Chirlian, P.M.: 'Bounds on the truncation error of periodic signals', *IEEE Trans. Circuit Theory*, 1972, **19**, (2), pp. 206–207
[24] Lind, M.: 'Functions of bounded variation', Mathematics C-level thesis. Karlstads University. Sweden, 2006
[25] Bergen, A.R., Chua, L.O., Mees, A.I.*, et al.*: 'Error bounds for general describing function problems', *IEEE Trans. Circuits Syst.*, 1982, **29**, (6), pp. 345–354
[26] Van der Pol, B.: 'On relaxation-oscillations', *Lond. Edinb. Dublin Phil. Mag. J. Sci.*, 1926, **2**, (11), pp. 978–992
[27] Fitzhugh, R.: 'Impulses and physiological states in theoretical models of nerve membranes', *Biophys. J.*, 1961, **1**, (6), pp. 445–466
[28] Nagumo, J., Arimoto, S., Yoshizawa, S.: 'An active pulse transmission line simulating nerve axon', *Proc. IRE*, 1962, **50**, (10), pp. 2061–2070
[29] Elowitz, M.B., Leibler, S.: 'A synthetic oscillatory network of transcriptional regulators', *Nature*, 2000, **403**, (6767), pp. 335–338
[30] Hori, Y., Takada, M., Hara, S.: 'Biochemical oscillations in delayed negative cyclic feedback: existence and profiles', *Automatica*, 2013, **49**, (9), pp. 2581–2590
[31] Thattai, M., Oudenaarden, A.V.: 'Intrinsic noise in gene regulatory networks', *Proc. Natl. Acad. Sci. USA*, 2001, **98**, (15), pp. 8614–8619

100

IET Syst. Biol., 2018, Vol. 12 Iss. 3, pp. 93-100