Nonparametric Analysis of Nonlinear Distortions for Biomolecular Systems

Agarwal, Atul; Dey, Abhishek; Relan, Rishi; Sen, Shaunak

Published in:
IFAC-PapersOnLine

Link to article, DOI:
10.1016/j.ifacol.2018.05.035

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnline CONFERENCE PAPER ARCHIVE

IFAC PapersOnLine 51-1 (2018) 313–318

Nonparametric Analysis of Nonlinear Distortions for Biomolecular Systems

Atul Agarwal * **Abhishek Dey** ** **Rishi Relan** *** **Shaunak Sen** ****

* *Mediatek Bangalore Private Limited, Bengaluru, Karnataka - 560103, India. (e-mail: atulagarwal893@gmail.com).*
** *Department of Electrical Engineering, Indian Institute of Technology Delhi, Hauz Khas, New Delhi - 110016, India. (e-mail: abhishek.dey@ee.iitd.ac.in)*
*** *DTU Compute, Technical University of Denmark, Asmussens Allé 2800 Kgs. Lyngby, Denmark (e-mail: risre@dtu.dk)*
**** *Department of Electrical Engineering, Indian Institute of Technology Delhi, Hauz Khas, New Delhi - 110016, India. (e-mail: shaunak.sen@ee.iitd.ac.in)*

Abstract: System identification of biomolecular circuits is a challenging problem, including due to the nonlinearities that are often present in them. The extent to which these nonlinearities contribute to the overall behaviour of the biomolecular circuit is unclear. Here, we address this issue for simple biomolecular circuit models by exploiting the properties of broadband random phase multisine excitations. We analysed the classical models of a two-state signaling system, an enzymatic signaling system, and of a transcriptional feedback circuit for the presence of nonlinear distortions at certain parametric settings and studied their dependence on the input parameters. These results should help the modeller in quantifying the effect of nonlinearities and assessing the validity of the linear models at a particular operating condition.

Keywords: System identification, Biomolecular systems, Nonlinear distortion, Frequency spectrum

1. INTRODUCTION

The mathematical modeling of biological processes is usually done by ordinary differential equations which represent the evolution over time of a particular biomolecule. The variables in these equations are often the concentration of species involved in the process. These equations depend on several parameters like generation and decay constants, reaction rates etc. Some of these parameters can be measured experimentally but it is not possible for many of them. Therefore it is required to determine unknown parameters indirectly from the measurement of other quantities. Besides, system identification plays an important role in deriving the dynamical model of the biomolecular systems directly from the input-output measurements.

Many identification techniques like linear and nonlinear least squares fitting (Mendes and Kell, 1998), Bayesian method (Wilkinson, 2007), Kalman and Particle filters (Liu and Niranjan, 2012) have been used to estimate the parameters of biological circuits. All these techniques try to approximate the system either by a linear or a nonlinear model. Since biological systems exhibit inherent stochasticity and are highly complex, therefore, linear models might not always work for these systems at different parameter settings.

For example, due to a change in the operating conditions such as temperature, humidity etc. or input signal amplitude, the output may be disturbed by nonlinear distortions (NL). One solution to this problem is to do nonlinear system identification (Westwick and Kearney, 2003; Billings, 2013; Schetzen, 1980) but nonlinear identification is computationally and resource expensive. So, it is wise to use this method only when there is a sufficient return on the additional resources. If an assessment about the nature and level of nonlinearities present in the system can be done beforehand, then an early decision about the use of a particular identification methodology can be made. Therefore, the goal is to first detect nature and level of nonlinearities present in the biomolecular system and then ascertain, whether it is still safe to use the linear system identification approach, even in the presence of nonlinear distortions. If not, then to approximate the system with a best linear approximation plus a nonlinear noise source (Schoukens et al., 2016).

Here, we concentrate on the first goal of identifying the nature as well as the level of nonlinearities present in the system. In this paper, we introduce a nonparametric technique for the analysis of NL for simple biomolecular systems by exploiting the properties of random phase multisines signals. To make things more clear, first, static nonlinear models are used to illustrate the technique. After that, it is utilised to analyse three simple biomolecular

2405-8963
Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2018.05.035

models. Finally, the effect of certain parameters is studied on the level of nonlinear distortions for these systems.

The paper is organized as follows. The basic principle, details about the excitation signal and the illustration of the method on simple static models is presented in Section 2. Section 3 discusses the nonparametric analysis of simple biomolecular classical models such as two-state signaling, covalent modification system and transcriptional feedback system. Finally, Section 4 presents conclusion based on the results obtained and future works.

2. METHOD

For detection of nonlinearities present in the system at certain operating conditions or parameter settings, a nonparametric analysis of the system's output response is performed as described in Schoukens et al. (2016). For completeness, essential aspects of this method are summarized in the following three sections.

2.1 Basic Idea

The basic idea is straightforward: a linear system cannot transfer power from one frequency to another or, in other words, output of a linear system contains power only at the frequencies which are present in the input spectrum. However, this is not necessarily true for a nonlinear system. Hence, by exciting the system at selected frequencies in a predefined frequency band of interest and utilising the information present at the non-excited harmonics in the output spectrum, a qualitative analysis of the kind and levels of nonlinearities present in the system can be performed. To quantify the behaviour of the biomolecular system regarding the level as well as the kind of nonlinearities, it is necessary to use a broadband excitation signal, so that maximum information about the behaviour in the band of excitation can be extracted (Schoukens et al., 2016).

2.2 Choice of the excitation signal

Gaussian excitations are widely used for this purpose. As compared to Gaussian excitation signals, a multisine excitation signal allows full control over the amplitude spectrum and power spectral density (PSD), while maintaining noise-like properties in the time-domain (Schoukens et al., 1998). A quick look at the output spectrum of system to a multisine excitation indicates whether the system behaves nonlinearly or not at a certain operating point. Further advantages of periodic excitations and steady state measurements compared to random excitations like Gaussian noise, periodic noise etc. can be found in Schoukens et al. (2006).

Here, our goal is to characterise a nonlinear system for the Gaussian excitation signals, using random phase multisine excitations. The amplitude spectrum of the multisine excitation should be such that the equivalence between the random phase multisine and the Gaussian random noise with respect to the nonlinear behavior is always guaranteed (Schoukens et al., 2016). Hence, the equivalence class E_{S_u} is defined which contains all signals that are (asymptotically) Gaussian distributed, and have asymptotically, for $N \to \infty$, where N is the number of excited harmonics, the same power on each finite frequency interval. This is defined precisely in the definition stated below.

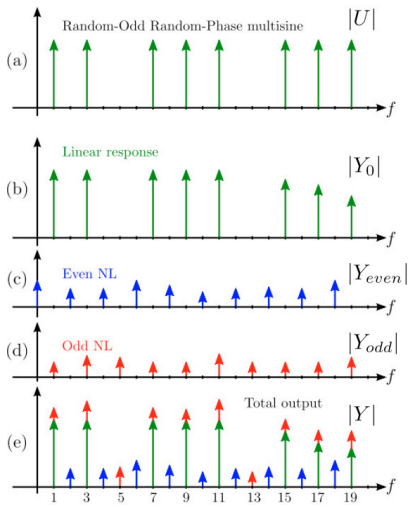

Fig. 1. Response PISPO: Total output is the sum of linear contributions (at excited lines), even NL (at even lines), odd NL (at odd lines) and noise (at all lines, not displayed here). Figure courtesy: (Relan et al., 2017)

Definition 1. Consider a signal u with a power spectrum $S_U(j\omega)$, which is piecewise continuous, with a finite number of discontinuities. A random signal belongs to the Riemann equivalence class of E_{S_u} if it obeys by any of the following statements (Schoukens et al., 2016): (1) It is a Gaussian noise excitation with power spectrum $S_U(j\omega)$. (2) It is a random multisine or random phase multisine (Pintelon and Schoukens, 2012) such that:

$$\frac{1}{N}\sum_{k=k_1}^{k_2}\mathbb{E}\{|U(j\omega_k)|^2\} = \frac{1}{2\pi}\int_{\omega_{k_1}}^{\omega_{k_2}} S_U(\nu)d\nu + \mathcal{O}(N^{-1}), \quad (1)$$

where $\omega_k = k\frac{2\pi f_s}{N}, k \in \mathbb{N}, 0 < \omega_{k_1} < \omega_{k_2} < \pi f_s$ and f_s is the sample frequency. The frequency domain representation of the multisine signal is the sum of the Fourier transforms of the individual sines and is given by:

$$U_{ms}(j\omega) = \frac{1}{\pi\sqrt{N_k}}\sum_{k\in\pm\mathbb{K}_{exc}} A(k)\delta(\omega_k - \omega_{k_e})e^{j\varphi_k}, \quad (2)$$

where $\delta(\bullet)$ is the Dirac delta function, $\mathbb{K}_{exc}$ is the set of positive excited frequencies, N_k the number of excited frequencies and $\varphi_k \sim \mathcal{U}[0, 2\pi[$ are the phases. The amplitudes of the multisine components $A(k) \geq 0$ can be chosen arbitrarily, depending on the application. In addition, some frequencies are not excited i.e. $A(k_{n.exc}) = 0$. The signal at these unexcited frequencies (*detection lines*) in the output spectrum contains valuable information about the presence of nonlinearities of the system.

2.3 Nonparametric data analysis

Assumption 1. The biomolecular system can be modeled as a weakly nonlinear periodic-in-same-period-out (PISPO) system described by a Volterra series (see Pintelon and Schoukens (2012) for more details).

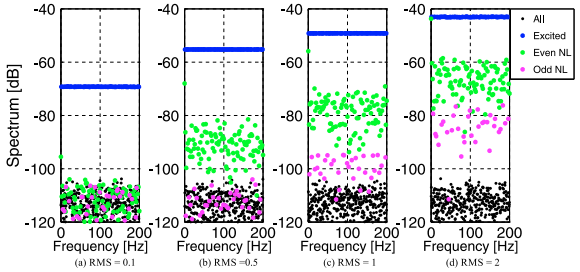

Fig. 2. Frequency Spectrum of output for $y = \frac{x}{k+x}$ with $k = 25$ and varying RMS of input. Black dots - 'All' represents all frequencies present in the output spectrum of the full data, blue dots - 'Excited' are the excited frequencies, green dots - 'Even NL' represents the even harmonics in the frequency band of interest and the magenta dots - 'Odd NL' are the non-excited odd harmonics in the frequency band of interest.

Remark 1. A nonlinear system is called PISPO if the steady state response to a periodic input is also a periodic signal with the same period as the input (with preservation of the period length), (Pintelon and Schoukens, 2012; Schoukens et al., 2016) for a more formal definition.

Since the nonlinear system (here biomolecular system) is operating in open-loop, the output Discrete Fourier Transform (DFT) spectrum of each period p where $p = 1, 2, 3, ..., P$, of the steady state response (with known periodic input) to an odd random phase multisine with random harmonic grid is given by:

$$Y^{[p]}(k) = Y_0(k) + V^{[p]}(k) + Y_S(k). \tag{3}$$

The total response of the system (see Fig. 1) is the sum of linear ($Y_0(k)$), stochastic nonlinear (even & odd) contributions $Y_S(k)$, and $V^{[p]}$ is the noise term (Pintelon and Schoukens, 2012; Schoukens et al., 2016). The information about the effect of NL and the need for differentiating between odd and even frequencies during the nonparametric test is essential in order to get an idea about the contributions of NL to the frequency response function of the biomolecular system under varying conditions.

2.4 Toy Examples: Static Systems

This method is illustrated with two simple static models. Four periods of input-output data were acquired and the first period was neglected to remove the effect of any transients. In the examples discussed below, a small zero mean Gaussian noise is added to the final output which represents the measurement noise. Only odd frequencies are excited in the frequency band of interest and one out of every three consecutive odd frequencies is also skipped randomly to check the effect of odd as well even nonlinearities present in the system. The sampling frequency is 500 Hz and maximum excited frequency is 200 Hz with frequency resolution of 2 Hz. For precise RMS value of excitation, the signal is divided by its RMS value and then multiplied by the desired RMS value.

(1) The system described by (4) is similar to a Hill function with Hill coefficient = 1, which corresponds to simple non-cooperative binding (Del Vecchio and Murray, 2015).

$$y = \frac{x}{k+x}. \tag{4}$$

For a particular value of k, this system is excited with x the odd random phase multisine input and y as output. Fig. 2 shows the output for $k = 25$ and when the RMS value of the multisine excitation is increased from 0.1 to 2. In this system, as RMS value of the input approaches magnitude of parameter k, the system will show more nonlinear behavior because the operating region will be closer to steep region of the input-output map. Fig. 2 supports the previous statement as it can be seen that the level of magenta nonlinear distortions increase with an increase in the RMS value of multisine excitation.

(2) Now, to understand it more, a comparatively more complex model described by (5) is studied

$$y = \frac{x^2}{k+x^2}. \tag{5}$$

This is equivalent to cooperative binding involving dimerization (Del Vecchio and Murray, 2015) though it is a static model. A similar analysis is done for this model also and compared with the previous model. Here, a small DC bias is also added to the input signal. Fig. 3 shows a comparison between the two systems. Since the model in (5) has more nonlinear characteristic (Hill coefficient = 2) than in (4), the level of nonlinear distortions is higher for the same RMS value of excitation signal.

As seen from the above analysis, this simple nonparametric analysis directly gives a qualitative and quantitative insight into the nonlinear distortions based on input-output data without needing any user interaction. This information can be used by the modeler to decide on the use of identification methodology.

3. CASE STUDIES

3.1 Two-State Signaling System

One of the widely used models to represent simple kind of nonlinearities present in biomolecular circuits is a two-state system (Huang and Ferrell, 1996; Stock et al., 2000). It is represented as a reversible reaction of the form

$$B \underset{K_-}{\overset{K_+}{\rightleftharpoons}} A \tag{6}$$

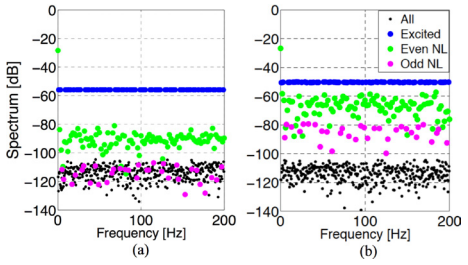

Fig. 3. Frequency Spectrum of output for (a) $y = \frac{x}{k+x}$ and (b) $y = \frac{x^2}{k+x^2}$ with dc bias of input = 1. Black, blue, green and magenta dots represent the noise, excited frequencies, even harmonic and odd harmonic distortion respectively.

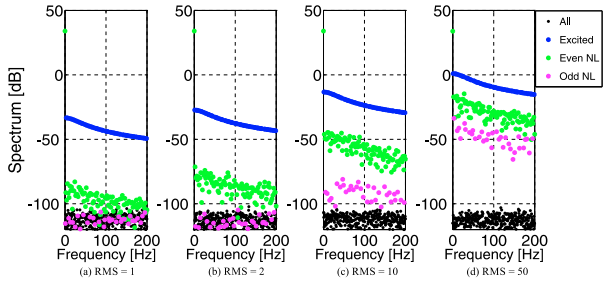

Fig. 4. Frequency spectrum of output for two-state signaling system for parameters $A_T = 100nM$, $K_- = 100/hr$ and different values of RMS of multisine. Black, blue, green and magenta dots represent the noise, excited frequencies, even harmonic and odd harmonic distortion respectively.

in which a protein B converts into protein A with forward rate constant K_+ and A converts into B with backward rate constant K_-. The total number of molecules in the system remain constant ($A_T = (A + B)$). The dynamic model of this system can be described as,

$$\dot{a} = K_+ A_T - a(K_+ + K_-), \tag{7}$$
$$y = a, \tag{8}$$

where a is molecular population of A and y is measured population of A. Here, K_+ is taken as input and y as output, due to which system becomes nonlinear as the state is multiplied with the input. This is similar to saturating input-output maps like (4), but with inherent dynamics. We aim to do nonlinear distortion analysis of this system using random phase multisine input (Appendix A),

$$K_+ = K_{+0} + \frac{1}{\pi\sqrt{N_k}} \sum_{k \in \mathbb{K}_{exc}} A(k)\, cos(\omega_k t + \varphi_k), \tag{9}$$

where $K_{+0} = 100$ is the DC bias added to the random phase multisine, rest all symbols have the same meaning as defined before. The value of the parameter K_- is kept constant, equal to the DC bias K_{+0}. The RMS value of the input excitation is increased from 1 to 50 to see its effect on the nonlinear distortions.

For this system, the NL is less if the RMS value of multisine input is lower and we operate in the linear region of input-output map. However, if the RMS value of the signal is comparable to the backward rate constant K_-, then we encounter the saturating region of the input-output map, so the system will exhibit more nonlinear characteristics and distortions will be higher. Similar behaviour can be observed from Fig. 4, which shows frequency spectrum of output y for different excitation levels. It can be clearly observed that at low excitation levels (see Figs. 4a and 4b), the distortions are ≈ 80 dB below the linear contribution. Therefore, in this case, linear estimation techniques can be used with slight approximation. But in case of higher excitations levels (Figs. 4c and 4d), the level of nonlinear contributions is significantly higher ≈ 30 dB and ≈ 15 dB respectively. Hence, a nonlinear model might be needed in order to capture the effects of nonlinear distortions at this level of input excitation.

3.2 Biomolecular covalent modification system: G-K Switch

Enzyme driven reversible covalent modification system is another example of biomolecular signaling system, analyzed by Goldbeter and Koshland (Goldbeter and Koshland, 1981), which is present in different signaling and metabolism pathways. This system can exhibit two different input-output properties (Figs. 5a and 5b) depending upon the parameter regime. One of them is more similar to the two-state case, whereas the other shows more switch-like response. In this system, a protein can coexist in unmodified form A and in modified form A^*. The conversion between the two forms is catalyzed by two converter enzymes, E_1 and E_2 as shown below,

$$A + E_1 \underset{d_1}{\overset{a_1}{\rightleftharpoons}} AE_1 \xrightarrow{k_1} A^* + E_1 \tag{10}$$

$$A^* + E_2 \underset{d_2}{\overset{a_2}{\rightleftharpoons}} A^*E_2 \xrightarrow{k_2} A + E_2. \tag{11}$$

Clearly, the total number of protein molecules of each type is conserved as there is no production or degradation in the model. Hence, the conservation equations are

$$\begin{aligned} A_T &= [A] + [A^*] + [AE_1] + [AE_2], \\ E_{1T} &= [E_1] + [AE_1], \\ E_{2T} &= [E_2] + [A^*E_2], \end{aligned} \tag{12}$$

where A_T is the total concentration of protein, E_{1T} and E_{2T} are the total concentration of catalyst of reaction 10 and 11 respectively. The kinetic equations described in (Goldbeter and Koshland, 1981) which govern the system model are modified using the above conservation equations and are given as (Dey and Sen, 2015),

$$\begin{aligned} \frac{d[A]}{dt} &= -a_1[A](E_{1T} - [AE_1]) + d_1[AE_1] \\ &\quad + k_2(A_T - [A] - [A^*] - [AE_1]), \\ \frac{d[A^*]}{dt} &= -a_2[A^*](E_{2T} - A_T + [A] + [A^*] + [AE_1]) \\ &\quad + d_2(A_T - [A] - [A^*] - [AE_1]) + k_1[AE_1], \\ \frac{d[AE_1]}{dt} &= a_1[A](E_{1T} - [AE_1]) - (d_1 + k_1)[AE_1]. \end{aligned} \tag{13}$$

The nonparametric analysis to detect the nonlinearities present in this system is performed with A^* as output and k_1 as input. With k_1 as random phase multisine (Eqn. 9), the system of equations (13), is solved with initial condition of states calculated by solving a cubic polynomial as in Goldbeter and Koshland (1981). The DC value of k_1 is 100 and its RMS value is taken 1 and 30. The choice of excited frequencies is same as taken in previous examples.

The effect of change in parameters a_1 and a_2 is shown in Fig. 5. In the regime, where the input-output map is less steep, nonlinear distortions increase with increase in the excitation level (Figs. 5c and 5e) as expected. Interestingly, in the regime where the response is more switch-like, nonlinear distortions are high even for lower excitation (Fig. 5d). Also, the distortions are very high at frequency ≈ 150 Hz which is a very different characteristic from other parameter setting. Therefore, linear identification techniques cannot be used for this case. This happens because even for lower excitation level, the operating

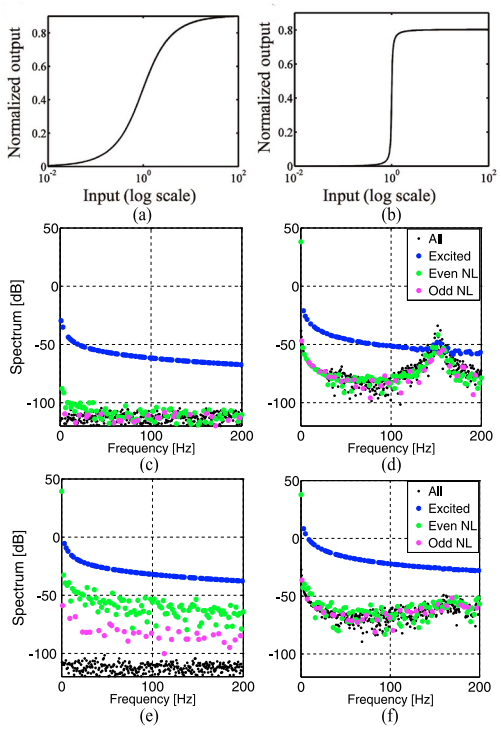

Fig. 5. Nonlinear distortion analysis for G-K switch. Left column is for parameter $a_1 = a_2 = 1/nM - hr$ and right column for $a_1 = a_2 = 100/nM - hr$. (a) and (b) show normalized input and normalized output map. (c), (d) and (e), (f) are frequency spectrum of A^* for RMS = 1 and 30 respectively. The value of other parameters is: $A_T = 200nM$, $E_{1T} = E_{2T} = 20nM$, $d_1 = d_2 = 100/hr$ and $k_2 = 100/hr$. Black, blue, green and magenta dots represent the noise, excited frequencies, even harmonic and odd harmonic distortion respectively.

region is around the steep input-output map (Fig. 5b). It should be noted that all this can be deduced from a simple nonparametric nonlinear analysis.

3.3 Transcriptional feedback system

Another example of biomolecular system which is of particular interest in signaling pathways is a transcriptional feedback system. If a protein activates or represses the rate of production of mRNA from DNA then it acts as a transcription factor. Here, we take a system which is similar to the two-state signaling system but with feedback from the output which activates its own production. Similar dynamics has been observed in *Xenopus* oocytes maturation (Ferrell and Xiong, 2001) leading to an all or none behaviour. The dynamical equations are given as,

$$\begin{aligned} \dot{a} &= K_+(A_T - a) - \gamma a - K_- a, \qquad (14) \\ \dot{A_T} &= f(a) - \gamma A_T, \end{aligned}$$

where,

$$f(a) = \alpha_0 + \frac{\alpha a^n}{K^n + a^n}.$$

This nonlinear feedback from output a may lead to hysteretic response depending upon parameter regime depending on the input amplitudes. We take K_+ as a random phase multisine input (Eqn. 9) and a as the output, with

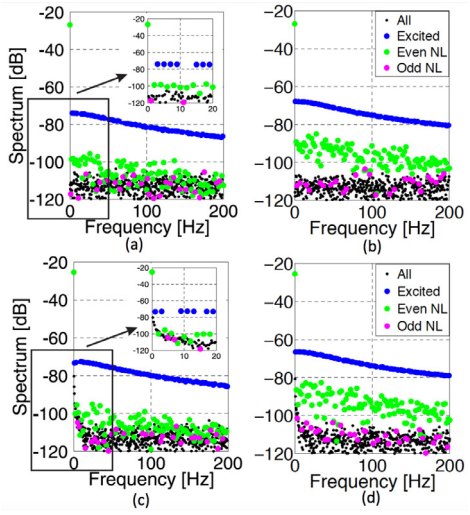

Fig. 6. Frequency spectrum plots of transcriptional feedback system. (a), (b) without hysteretic input-output map ($\alpha = 1nM/hr$) and (c), (d) with hysteretic input-output map ($\alpha = 5nM/hr$), at different excitation levels (RMS = 30 for a,c; RMS = 60 for b,d). Other parameters are : $\alpha_0 = 1/15nM/hr$, $K = 1nM$, $\gamma = 1/hr$ and $n = 2$. Black, blue, green and magenta dots represent the noise, excited frequencies, even harmonic and odd harmonic distortion respectively.

$K_{+0} = 200$ and solve Eqn. 14 computationally. Here, eight periods of input-output data were acquired, dropping the first four periods to eliminate the effect of transients. The RMS value of the signal is varied to check the level of nonlinear distortions.

For lower RMS values (see Figs. 6a,c) the level of nonlinear distortions are lower compared to linear contribution, whereas for higher RMS values (see Figs. 6b,d) the difference becomes less than 30 dB. Interestingly, as the system approaches the parameter values near the hysteretic regime, the level of noise and NL at lower frequencies is higher even for lower RMS value (Fig. 6c), one thing to note here is that the PISPO assumption is still valid at this input amplitude, as the period length is preserved.

4. CONCLUSION

In this paper, we proposed a data-driven frequency domain nonparametric analysis technique to extract the information about the level and nature (odd or even) of nonlinearities present in the representative biomolecular circuits. Here, we have considered three such examples which are important in different cellular context. First, we have chosen a random phase multisine input and computed the discrete Fourier transform of the output. Next, we have studied the effect of change in RMS value of the input as well as effect of change in system parameters. Finally, we have quantified the level of nonlinear distortion in all the cases. In all the examples considered, it is observed that the level of nonlinear distortions increases with an increase in the excitation level. In the case of the Goldbeter-Koshland switch, the operating parameter regime also have significant effect on choice of system identification method to use. Interestingly, in the regime where the response is more switch-like, nonlinear distortions are

found to be high even for lower excitation. Similar results are also observed for transcriptional feedback system. The nonlinearity information does not differ much if different parameters in the models are used as input.

As this technique discussed in the paper only provides a nonparametric representation of the system, the next step is to find a parametric representation of the nonlinear biomolecular systems by computing the best parametric linear approximation and nonlinear noise source (Eqn. 3) because a parametric representation is very useful in stability analysis of the dynamic systems. Another natural task for future is to use this technique for analysing nonlinear distortions in more complex and larger biomolecular systems. The application and design of the random-phase multisine signals has already been proven in many of the real life biological and electrochemical case studies experimentally (Van Ingelgem et al., 2009; Sanchez and Bragos, 2009; Olarte et al., 2014), hence designing these signal in practice is not cumbersome and difficult at all with all the modern as well as cheap instrumentation and data acquisition equipment available.

It has been demonstrated that this technique can be used to study directly from the input-output data, effect of any parameter change and its effect on the system dynamics. The proposed technique gives the modeller a quantitative and qualitative insight of the nonlinear distortions present in the biological systems. Hence, it gives a possibility to decide at an early stage, on the feasibility of different identification methodologies as well as on the suitability of different class of models for biological systems, which eventually can save time and resources.

REFERENCES

Billings, S.A. (2013). *Nonlinear system identification: NARMAX methods in the time, frequency, and spatio-temporal domains*. John Wiley & Sons.

Del Vecchio, D. and Murray, R.M. (2015). *Biomolecular feedback systems*. Princeton University Press.

Dey, A. and Sen, S. (2015). Describing function-based approximations of biomolecular signalling systems. In *14th European Control Conference (ECC)*, 2292–2297. Linz, Austria.

Ferrell, Jr., J.E. and Xiong, W. (2001). Bistability in cell signalling: How to make continuous processes discontinuous, and reversible proceses irreversible. *Chaos*, 11(1), 227–236.

Goldbeter, A. and Koshland, D.E. (1981). An amplified sensitivity arising from covalent modification in biological systems. *Proceedings of the National Academy of Sciences*, 78(11), 6840–6844.

Huang, C.Y. and Ferrell, J.E. (1996). Ultrasensitivity in the mitogen-activated protein kinase cascade. *Proceedings of the National Academy of Sciences*, 93(19), 10078–10083.

Liu, X. and Niranjan, M. (2012). State and parameter estimation of the heat shock response system using kalman and particle filters. *Bioinformatics*, 28(11), 1501–1507.

Mendes, P. and Kell, D. (1998). Non-linear optimization of biochemical pathways: applications to metabolic engineering and parameter estimation. *Bioinformatics*, 14(10), 869–883.

Olarte, O., Barbé, K., Van Moer, W., Van Ingelgem, Y., and Hubin, A. (2014). Measurement and characterization of glucose in nacl aqueous solutions by electrochemical impedance spectroscopy. *Biomedical Signal Processing and Control*, 14, 9 – 18.

Pintelon, R. and Schoukens, J. (2012). *System identification: a frequency domain approach*. John Wiley & Sons.

Relan, R., Firouz, Y., Timmermans, J.M., and Schoukens, J. (2017). Data-driven nonlinear identification of li-ion battery based on a frequency domain nonparametric analysis. *IEEE Transactions on Control Systems Technology*, 25(5), 1825–1832.

Sanchez, B. and Bragos, R. (2009). *Fast Electrical Impedance Spectroscopy for Moving Tissue Characterization Using Bilateral QuasiLogarithmic Multisine Bursts Signals*, 1084–1087. Springer Berlin Heidelberg, Berlin, Heidelberg.

Schetzen, M. (1980). *The Volterra and Wiener theories of nonlinear systems*. Wiley.

Schoukens, J., Rolain, Y., and Pintelon, R. (2006). Analysis of windowing/leakage effects in frequency response function measurements. *Automatica*, 42(1), 27 – 38.

Schoukens, J., Dobrowiecki, T., and Pintelon, R. (1998). Parametric and nonparametric identification of linear systems in the presence of nonlinear distortions-a frequency domain approach. *IEEE Transactions on Automatic Control*, 43(2), 176–190.

Schoukens, J., Vaes, M., and Pintelon, R. (2016). Linear system identification in a nonlinear setting: Nonparametric analysis of the nonlinear distortions and their impact on the best linear approximation. *IEEE Control Systems Magazine*, 36(3), 38–69.

Stock, A.M., Robinson, V.L., and Goudreau, P.N. (2000). Two-component signal transduction. *Annual review of biochemistry*, 69(1), 183–215.

Van Ingelgem, Y., Tourwé, E., Blajiev, O., Pintelon, R., and Hubin, A. (2009). Advantages of odd random phase multisine electrochemical impedance measurements. *Electroanalysis*, 21(6), 730–739.

Westwick, D.T. and Kearney, R.E. (2003). *Identification of nonlinear physiological systems*, volume 7. John Wiley & Sons.

Wilkinson, D.J. (2007). Bayesian methods in bioinformatics and computational systems biology. *Briefings in bioinformatics*, 8(2), 109–116.

Appendix A. ALGORITHM

Algorithm 1 Algorithm for NL analysis in biomolecular systems

1: Choose sampling frequency f_s and frequency band of interest and number of periods.
2: Choose random phase multisine input (as in Eqn. 9) for biomolecular system with suitable bias, phase uniformly distributed in $[0,2\pi[$, exciting only the odd frequency lines and leaving one odd frequency line randomly in each three consecutive blocks of odd frequency lines.
3: Compute output DFT ignoring first few periods for steady-state analysis.
4: NL is the non-excited frequency components in output DFT.