
Proceedings Paper:
Foo, M, Bates, DG and Kim, J orcid.org/0000-0002-3456-6614 (2017) System identification of gene regulatory networks for perturbation mitigation via feedback control. In: 2017 IEEE 14th International Conference on Networking, Sensing and Control (ICNSC). 2017 IEEE 14th International Conference on Networking, Sensing and Control (ICNSC), 16-18 May 2017, Calabria, Italy. IEEE .

https://doi.org/10.1109/ICNSC.2017.8000094

System identification of gene regulatory networks for perturbation mitigation via feedback control

Mathias Foo, Jongrae Kim and Declan G. Bates

Abstract— **In Synthetic Biology, the idea of using feedback control for the mitigation of perturbations to gene regulatory networks due to disease and environmental disturbances is gaining popularity. To facilitate the design of such synthetic control circuits, a suitable model that captures the relevant dynamics of the gene regulatory network is essential. Traditionally, Michaelis-Menten models with Hill-type nonlinearities have often been used to model gene regulatory networks. Here, we show that such models are not suitable for the purposes of controller design, and propose an alternative formalism. Using tools from system identification, we show how to build so-called S-System models that capture the key dynamics of the gene regulatory network and are suitable for controller design. Using the identified S-System model, we design a genetic feedback controller for an example gene regulatory network with the objective of rejecting an external perturbation. Using a sine sweeping method, we show how the S-System model can be approximated by a second order linear transfer function and, based on this transfer function, we design our controller. Simulation results using the full nonlinear S-System model of the network show that the designed controller is able to mitigate the effect of external perturbations. Our findings highlight the usefulness of the S-System modelling formalism for the design of synthetic control circuits for gene regulatory networks.**

I. INTRODUCTION

In any complex networks such as the traffic systems, power grids, irrigation networks, etc, the presence of external disturbances can have adverse effects on the overall system. These unwarranted effects include gridlock in the movement of transportation, major power outages in residential and industrial areas and poor water supply to farming areas. In view of this, network control particularly in the presence of disturbances has been subjected to intensive studies, resulting in the development of many useful tools for the control of complex networks.

Due to advances in this area, synthetic biologists have recently began to investigate the application of the aforementioned tools to the control of biological networks and systems. Some notable examples can be found in [1], [2], [3], [4], [5], where strategies based on feedback control theory have been used to analyse the controllability, observability and stability of biological networks such that appropriate sets of control design rules can be developed.

In this paper, we focus our attention on the *modelling* and *control design* of gene regulatory networks. The ability to 'control' the dynamics of gene regulatory networks, especially in the presence of disturbance, has many useful applications in the field of synthetic biology, where synthetic circuits can be developed to implement the proposed controllers and hence curb the effect of external disturbances due to disease or environmental changes. Here, we use system identification techniques to build models of gene regulatory networks that are suitable for control system design. From the identified models, we design a feedback controller that can be implemented genetically in order to reject external disturbances that enter the network.

This paper is organised as follows. In Section II, we present the example gene regulatory network that is used to build our model for control design. In Section III, we evaluate different types of models used to describe gene regulatory networks from the perspective of control system design, and we propose a system identification approach for model building. The control design procedure is described and closed-loop simulation results are provided in Section IV. Conclusions are given in Section V.

II. DREAM GENE REGULATORY NETWORK

The DREAM *in silico* gene regulatory network challenge is established to serve as a benchmark to assess different proposed approaches to infer gene regulatory networks from given experimental data [6], [7], [8]. Often in the DREAM challenge, the time-series data for each gene (or node) in the network are provided and the aim is for the participants to deduce the underlying network to attain insights such as the interconnecting edges, the direction of the information flow, etc. The provided gene regulatory networks are typically subsets of actual transcriptional networks in model organisms such as *E. coli* and *S. cerevisiae*, and hence, they are representative of real biological systems.

In this paper, we choose the DREAM3 Size 10 data set (hereafter we use the term DREAM3 to denote this network), which consists of mRNA temporal data on a network composed of 10 interconnecting genes that is a subset of a *S. cerevisiae* gene regulatory network. As the dataset does not include separate protein data, in the following, we make the following two assumptions: (i) the temporal evolution of the protein is similar to the mRNA and (ii) the protein is linearly translated from mRNA. Following these two assumptions, we can lump the protein dynamics into the transcription rate of the mRNA at steady state, and this results in a complete network that can be described solely using mRNA levels. In this DREAM3 data set, information regarding the direction of the interconnectivity between each genes is provided and the depiction of these interactions is shown in Fig. 1(A).

M. Foo, and D.G. Bates are with Warwick Integrative Synthetic Biology Centre, School of Engineering, University of Warwick, Coventry, CV4 7AL, UK. J. Kim is with School of Mechanical Engineering, University of Leeds, Leeds, LS2 9JT, UK. M.Foo@warwick.ac.uk, menjkim@leeds.ac.uk, D.Bates@warwick.ac.uk

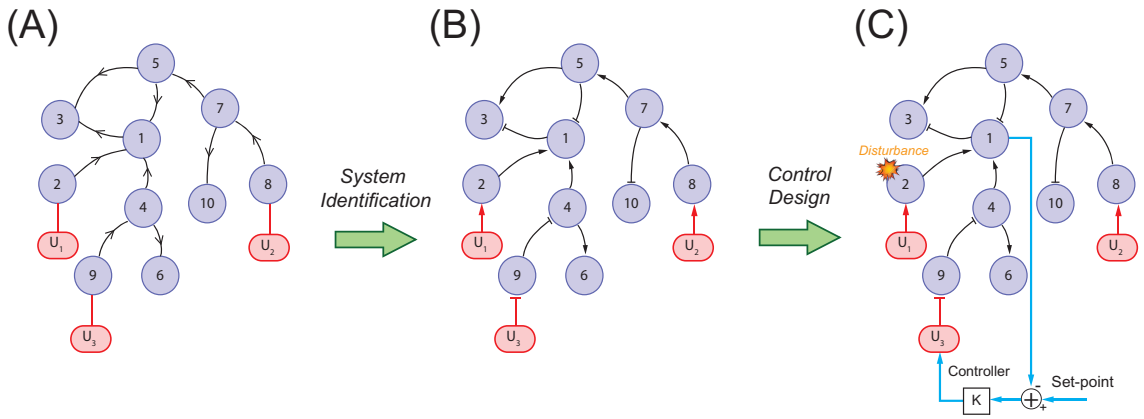

Fig. 1: (A) DREAM3 gene regulatory network. Purple circles represent genes and red rectangles represent external inputs. The direction of regulation is indicated by the triangle arrow. (B) Using system identification, the types of regulation in the network are identified. Arrow head indicates activation and Bar head indicates inhibition. (C) Proposed control design configuration for disturbance rejection.

III. MODEL DESCRIPTION

A. On the Michaelis-Menten and Hill-type nonlinearity model structure

Model structures employing Michaelis-Menten and Hill-type nonlinearities are commonly used to describe the dynamics of gene regulatory networks. If the regulation type and the cooperative binding are known, the modeller can either specify

$$F_a = k_0 N_P^h / (K_M + N_P^h) \tag{1}$$

for an activation type of regulation or

$$F_i = k_0 / (K_M + N_P^h) \tag{2}$$

for an inhibition type of regulation. In both Eqns. (1) and (2), N_P is the transcription factor, k_0 and K_M are associated with the Michaelis-Menten constants and h is the Hill coefficient.

In the context of network inference, this type of model structure can be used only if the type of regulation (activatory or inhibitory) between each gene in the network is known. In the event that the type of regulation is unknown, then this model structure is not suitable as the structure of an activation or an inhibition type of regulation is different and arbitrarily assigning them in the model building stage could thus lead to poor model accuracy. An additional problem in the context of synthetic biology is that models of this type are not suitable for subsequent use in the design of synthetic controllers. To illustrate this, let us consider Eqn. (1) and assume that our control action (i.e. output of the controller) is given by N_P. If $N_P \gg K_M$, then $F_a \approx k_0 (N_P^h / N_P^h) = k_0$, resulting in a saturated control action, which is undesirable from control design perspective. In view of these two limitations, an alternative model structure is thus required. The alternate model structure needs to have a general structure that can accommodate either type of regulation and be useful for controller design.

B. S-System models for gene regulatory networks

Here, we choose the so-called S-System modeling formalism as an alternative approach to describe the dynamics of gene regulatory networks. The S-System modeling framework was originally developed from the field of biochemical system theory (see e.g. [9], [10]), and has been used to describe the dynamics of gene regulation (see e.g. [11], [12]), where it has been shown to be as accurate as Michaelis-Menten with Hill-type nonlinearity models (see [13]). The S-System model has the following form:

$$\frac{dN_i}{dt} = a_i \prod_{j=1}^{M_1} N_j^{p_{i,j}} + b_i \prod_{j=1}^{M_2} N_j^{q_{i,j}} + c_i U \tag{3}$$

where i denotes the number of biochemical component, $a > 0$, $b < 0$ and $c \in (-\infty, +\infty)$ are constants, N represents the biochemical component, M_1 and M_2 are the total number of biochemical components involved in the interaction and U is the external input. The power exponent terms, p and q are associated with the production and degradation terms respectively. For simplicity, we assume $q = 1$ throughout this paper. Additionally, a positive value of p represents activation while a negative value of p represents inhibition.

Note that the S-System model has a general structure that can accommodate either an activation or inhibition type of regulation via the sign of p. Thus, no prior knowledge of the type of regulation is required in the model building exercise. Moreover, the S-System model can be used for the purposes of controller design as it does not suffer from the issues affecting the controller action described in Section III-A.

Remarks: In a metabolic pathway, M_1 and M_2 are required due to the different number of components interacting with the respective production and degradation components. For gene transcription, since only mRNA itself is degrading, $M_2 = 1$ and setting $q = 1$ is also consistent with the standard linear degradation model.

C. System identification of an S-System model

Fig. 1(A) shows the interconnection between the genes in the DREAM3 gene regulatory network. The DREAM3 network provides no information regarding the type of regulation between the interconnecting genes, and therefore we will use system identification techniques (see e.g. [14]) to infer the type of regulation within the network.

System identification techniques have been used to build models of gene regulatory networks in several previous studies, including [15], [16], [17], where linear *black box* network models were considered and the directions and the types of regulation were identified based on available data on gene expression profiles. In this paper, we consider a *grey box* S-System model, given that we have prior knowledge about the network interconnections, and focus our attention on the identification of the type of regulation between the interconnecting genes. As per standard system identification procedures, we use one data set for model estimation and another data set for validation.

Thus, the S-System model for the DREAM3 gene regulatory network following Fig. 1(A) is given by

$$\begin{aligned}
\frac{dN_1}{dt} &= a_1 N_2^{p_{1,2}} N_4^{p_{1,4}} N_5^{p_{1,5}} + b_1 N_1, & \frac{dN_2}{dt} &= b_2 N_2 + c_2 U_1 \\
\frac{dN_3}{dt} &= a_3 N_1^{p_{3,1}} N_5^{p_{3,5}} + b_3 N_3, & \frac{dN_4}{dt} &= a_4 N_9^{p_{4,9}} + b_4 N_4 \\
\frac{dN_5}{dt} &= a_5 N_7^{p_{5,7}} + b_5 N_5, & \frac{dN_6}{dt} &= a_6 N_4^{p_{6,4}} + b_6 N_6 \\
\frac{dN_7}{dt} &= a_7 N_8^{p_{7,8}} + b_7 N_7, & \frac{dN_8}{dt} &= b_8 N_8 + c_8 U_2 \\
\frac{dN_9}{dt} &= b_9 N_9 + c_9 U_3 + d_9, & \frac{dN_{10}}{dt} &= a_{10} N_7^{p_{10,7}} + b_{10} N_{10}
\end{aligned} \tag{4}$$

Note that for dN_9/dt, as mRNA levels are physical quantity, a constant value denoted by d_9 is added to the model to ensure that the overall mRNA level stays positive since U_3 is negatively correlated with N_9 and b_9 is negative due to the degradation term. Note that the inclusion of d_9 does not change the structure of S-System model as the equivalent model structure can be obtained by setting $d_i = a_i$ and $p_{i,j} = 0$.

Let $\theta = \{a_i, b_i, c_i, d_1, p_{i,j}\}$ with i and j represent the appropriate indices in Eqn. (4), the values of θ can be estimated using the prediction error method with a quadratic criterion, i.e.

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \frac{1}{L} \sum_{i=1}^{10} \sum_{t=1}^{L} [N_i(t) - \hat{N}_i(t,\theta)]^2 \tag{5}$$

where $L = 20$ is the length of the data, $\hat{N}$ denotes the simulated data from the S-System model while N denotes the real data and Eqn. (5) is solved using MATLAB function *fminsearch*, which uses the Nelder-Mead simplex algorithm. Table I tabulates the estimated model parameters of the S-System model and Fig. 2 shows the comparison between the S-System model and the real data on the validation data set.

TABLE I: Esimated parameters for the S-System model.

Gene	Values
N_1	a_1 = 0.2757, $p_{1,2}$ = 0.3502, $p_{1,4}$ = 0.0559, $p_{1,5}$ = -0.2789, b_1 = -0.4023
N_2	b_2 = -0.1875, c_2 = 0.0946
N_3	a_3 = 0.1478, $p_{3,1}$ = -0.0021, $p_{3,5}$ = 0.1393, b_3 = -0.1481
N_4	a_4 = 0.0023, $p_{4,9}$ = -5.1622, b_4 = -0.3555
N_5	a_5 = 0.1199, $p_{5,7}$ = 0.0760, b_5 = -0.2057
N_6	a_6 = 0.2567, $p_{6,4}$ = -0.0120, b_6 = -0.3035
N_7	a_7 = 0.0607, $p_{7,8}$ = 0.1104, b_7 = -0.1237
N_8	b_8 = -0.0298, c_8 = 0.0108
N_9	b_9 = -0.1793, c_9 = -0.0268, d_9 = 0.1733
N_{10}	a_{10} = 0.0139, $p_{10,7}$ = -1.5609, b_{10} = -0.0480

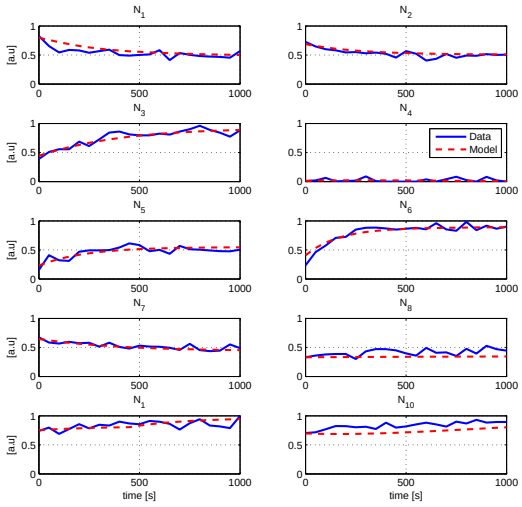

Fig. 2: Comparison between S-System model and DREAM3 data on the validation data set.

From the estimated parameters shown in Table I, we are able to determine the type of regulation in the network, where a positive value of the power term denotes activation while a negative value of the power term denotes inhibition. Reassuringly, all the known degradation terms were identified to have negative values, in accordance with biological reality.

The comparison between the S-System model and the real data on the validation data set shows good agreement, suggesting a good level of accuracy of the model. To quantify this, we calculate the Mean Square Error (MSE) for each gene between the S-System model and the real data. The MSE is computed using,

$$\text{MSE} = \frac{1}{L} \sum_{t=1}^{L} [N_i(t) - \hat{N}_i(t,\theta)]^2 \tag{6}$$

where $i = 1, 2, \ldots, 10$. Table II shows the computed MSE for both the estimation and validation data sets.

The total MSE, MSE_T, is obtained by summing all the individual MSE from each genes. In general, the MSE values are small and similar between the two data sets. With the

TABLE II: MSE for both estimation and validation data sets.

Gene	MSE (Estimation)	MSE (Validation)
N_1	0.0029	0.0054
N_2	0.0013	0.0021
N_3	0.0014	0.0031
N_4	0.0009	0.0010
N_5	0.0010	0.0037
N_6	0.0017	0.0036
N_7	0.0019	0.0016
N_8	0.0012	0.0088
N_9	0.0033	0.0050
N_{10}	0.0017	0.0128
MSE_T	0.0171	0.0470

regulation types in the DREAM3 network as identified, the network interactions are as shown in Fig. 1(B).

IV. CONTROLLER DESIGN

To achieve an implementable controller for a gene regulatory network, a genetic based controller is required, and there are frameworks available for such designs (see e.g. [18], [19]), where combination of several proteases can be utilised to achieve a genetic based lead-lag type of controller. In this paper, we employ a frequency domain control design methodology to control the DREAM3 network, motivated by the design framework proposed in [19]. In order to design controllers in the frequency domain, a linear model is required. As the S-System is a nonlinear model, the standard procedure is to linearise the model. However, linearising the S-System is not trivial due to the presence of the non-integer power exponent terms. Thus, as alternative we approximate the S-System model with a linear transfer function obtained using the sine sweeping method (see e.g. [14], [20]).

A. Sine sweeping method

In the sine sweeping method, sinusoidal input signals within the frequency range of interest are given as the inputs to the system. The output responses within the frequency range are then analysed in terms of their magnitude and phase relative to the input signal. By collecting these magnitude and phase values, the Bode plot of the system can be easily obtained. Here, we summarise the procedure for obtaining the Bode plot using the sine sweeping method and refer readers to [14], [20] for complete details.

Consider a sinusoidal input $u(t) = A\sin(\omega_0 t)$, where A is the amplitude and ω_0 is the frequency. For any linear time invariant system, the output would be also sinusoidal with the same frequency but with scaled amplitude and a phase shift. In practice, the output response is subjected to transient effects, as well as effects of nonlinearities and disturbance $d(t)$, yielding,

$$y(t) = B\sin(\omega_0 t + \phi) + d(t) + \text{transient} + \text{nonlinearities} \quad (7)$$

where $B = A|G(j\omega_0)|$, $\phi = \angle G(j\omega_0) = \tan^{-1}\frac{\text{Im}|G(j\omega_0)|}{\text{Re}|G(j\omega_0)|}$ and $G(j\omega_0)$ is the transfer function relating the input and output. The effect of transient and nonlinearities can be reduced by not considering the initial part of the data and assuming the linear contribution dominates the nonlinearities respectively. To reduce the effect of $d(t)$ on $y(t)$, one can use a correlation method [14], where the idea is to correlate y with a sine and cosine of the same frequency and average it over the length of the data N_L (see Fig. 3).

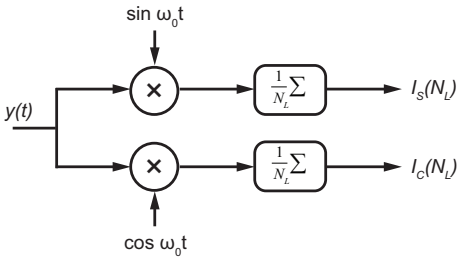

Fig. 3: Correlation method.

From Fig. 3, we obtain,

$$I_S(N_L) = \frac{1}{N_L}\sum_{t=1}^{N_L} y(t)\sin(\omega_0 t) \quad , I_C(N_L) = \frac{1}{N_L}\sum_{t=1}^{N_L} y(t)\cos(\omega_0 t) \quad (8)$$

Substituting Eqn. (7) into (8), and after some algebraic manipulation, we arrive at

$$\begin{aligned} I_S(N_L) &= \frac{A}{2}|G(j\omega_0)|\cos\phi - \frac{A}{2}|G(j\omega_0)|\frac{1}{N_L}\sum_{t=1}^{N_L}\cos(2\omega_0 t \\ &\quad + \phi) + \frac{1}{N_L}\sum_{t=1}^{N_L} d(t)\sin(\omega_0 t) \\ I_C(N_L) &= \frac{A}{2}|G(j\omega_0)|\sin\phi - \frac{A}{2}|G(j\omega_0)|\frac{1}{N_L}\sum_{t=1}^{N_L}\sin(2\omega_0 t \\ &\quad + \phi) + \frac{1}{N_L}\sum_{t=1}^{N_L} d(t)\cos(\omega_0 t) \end{aligned} \quad (9)$$

From Eqn. (9), the second term for both $I_S(N_L)$ and $I_C(N_L)$ will go to zero as $N_L \to \infty$. Assuming $d(t)$ is a stationary stochastic process with zero mean value and covariance function $R_d(l)$ such that $\sum_{l=0}^{\infty} l|R_d(l)| < \infty$, the third term for both $I_S(N_L)$ and $I_C(N_L)$ will be zero as $N_L \to \infty$ as the variance of the third term decays at a rate of $1/N_L$ [14]. From the remaining terms of Eqn. (9), the magnitude, $|G(j\omega_0)|$ and the phase, $\angle G(j\omega_0)$ can be estimated using the following equations, i.e.

$$|G(j\omega_0)| = \frac{2}{A}\sqrt{I_S^2(N_L) + I_C^2(N_L)},\ \angle G(j\omega_0) = \tan^{-1}\frac{I_C(N_L)}{I_S(N_L)} \quad (10)$$

For the DREAM3 network, we assume that the input to the network is through U_3 and the output of interest is the expression of gene N_1. We apply sinusoidal signals in the frequency range from 0.001 rad/s to 1.000 rad/s. Despite using a nonlinear model, we note that the output sinusoidal responses have the same frequency as the input and no subharmonics are apparent, indicating a dominant linearity

of the model. By computing the magnitude and phase values using Eqn. (10), the Bode plot of the DREAM3 network from input U_3 to output N_1 is obtained and shown in Fig. 4.

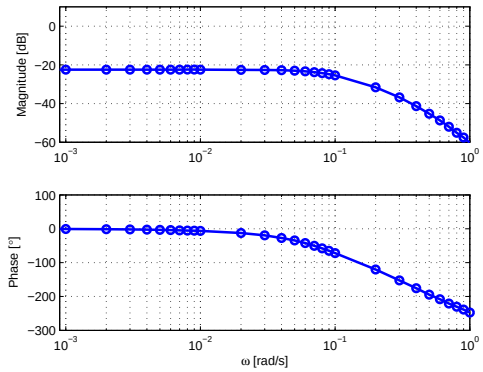

Fig. 4: Bode plot of DREAM3 network from input U_3 to output N_1.

From the Bode plot, we note the following: (i) At low frequency, the magnitude of the system is about -22.5dB. (ii) The corner frequency is 0.11 rad/s. (iii) At the corner frequency, the slope is close to -40dB/dec and the phase is approximately -90°, suggesting a second order system with repeating poles. Thus, the transfer function relating input U_3 to output N_1 can be approximated by

$$\frac{N_1(s)}{U_3(s)} = \frac{0.0750}{(1+\frac{s}{0.11})^2} = \frac{0.0009}{s^2+0.22s+0.012} \qquad (11)$$

With this transfer function identified, we can proceed with the design of the controller using a frequency domain approach.

B. Design of a genetic phase lag controller for disturbance rejection

In this section, we illustrate the design of the genetic phase lag controller. A phase lag controller is chosen, as this type of controller is typically used to improve disturbance rejection and reduce steady state error.

The phase lag controller has the following form:

$$K(s) = \frac{K_1}{s+a_P} + K_2 = \frac{K_2(s+a_P+\frac{K_1}{K_2})}{s+a_P} \qquad (12)$$

where the zero of the controller $z = -(a_P+(K_1/K_2))$ and the pole of the controller $p = -a_P$, with the gain of the controller K_2. As both the gain and phase margins of the system obtained from the Bode plot are infinite, our primary focus is on improving the transient dynamics of the disturbance rejection and reducing the steady state error.

The transfer function given in Eqn. (11) is a type 0 system, and with the use of a phase lag controller, there is no integrator in the open loop gain to eliminate the steady state error. As such, when choosing the pole of the phase lag controller, we try to place the pole, a_P as close as possible to the origin. Likewise, the static error constant, $K_p = 0.0027K_2$ should be chosen as large as possible to reduce the steady state error. The choice of the design parameters are constrained by the achievable biological values and following the range of allowable values given in [19]; the following inequalities should be adhered to: $0.0002 \leq a_P \leq 0.0040$, $K_1 < 2.3$ and $K_2 < 1.8$.

C. Simulation examples

While the design of the controller is carried out using the linear model, for implementation, we carried out our simulation using the S-System model. In most gene regulatory network perturbation mitigation problems, we are interested in maintaining the steady state level of a particular gene of interest in the presence of a perturbation. Biologically, this can be interpreted as maintaining the level of expression of a gene of interest to ensure optimal biological function. Thus, in this simulation example, we are interested in maintaining the steady state level of N_1 at its desired reference value in the presence of a disturbance. Here, we assume that the disturbance enters the network through U_1 and our control action is provided by U_3 as depicted in Fig. 1(C).

In the absence of a disturbance, the steady state level of N_1 is 0.486, thus, our control objective is to maintain the steady state level of N_1 close to 0.486 in the presence of a disturbance. In our simulation, a step disturbance with amplitude of 2 enters the network at time 4000s. As can be seen in Fig. 5(A), without control, the steady state level of N_1 increase to 0.63 and is unable to return to its desired value.

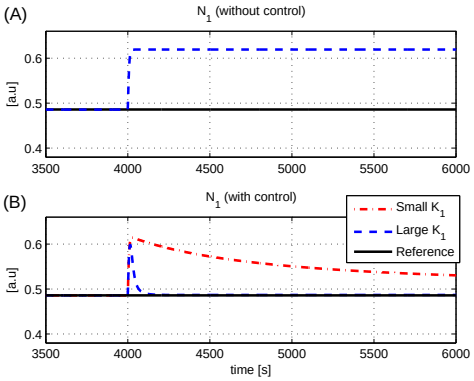

Fig. 5: (A) N_1 set-point regulation (without control). (B) N_1 set-point regulation (with control). Black solid line: Set-point. Red dash-dotted line: N_1 response to small K_1. Blue dashed line: N_1 response to large K_1.

In the design of the phase lag controller, the following values are chosen. To have the pole close to the origin, we choose $a_P = 0.0002$. To have the static error constant as large as possible, we choose $K_2 = 1.7$. For K_1, we consider two cases, i.e. $K_1 = 0.04$ (controller's zero close to origin) and $K_1 = 2$ (controller's zero far from the origin). The simulation results are shown in Fig. 5(B). For a small value of K_1, we see that the performance of the system is slow and at

time 6000s, there is still a noticeable steady state error, i.e. 0.044. On the other hand, for a large value of K_1, we see a significant improvement in the performance, where we get a faster response and an almost zero steady state error, i.e. 0.0008. The Bode plots of the system with and without

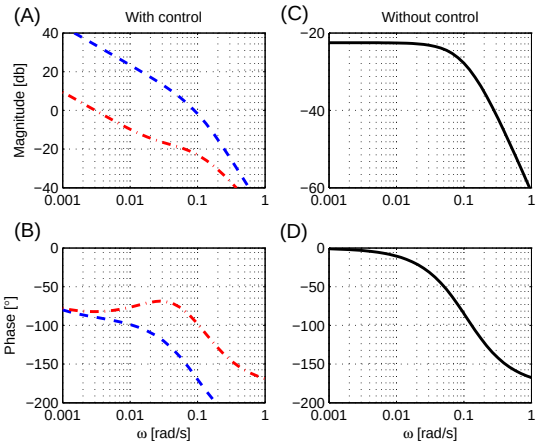

Fig. 6: (A) & (B) Gain and phase plots of with control. Red dash-dotted line: Small K_1, Blue dashed line: Large K_1. (C) & (D) Gain and phase plots without control.

control are shown in Fig. 6. For a small value of K_1, we note that the phase margin of the system is 97°. On the other hand, for a large value of K_1, despite the good performance, we note that the phase margin of the system reduces from 97° to 10°, which is less than typically specified values. Thus, a compromise between the transient performance and overall stability robustness needs to be performed when designing the controller, and this trade-off can be effectively managed through the choice of the controller parameter K_1.

V. CONCLUSIONS

In this paper, we use system identification techniques to build a model of a gene regulatory network that is suitable for the purposes of control system design. We show that standard approaches employing Michaelis-Menten models with Hill-type nonlinearities are not appropriate model structures if the type of regulation between interacting genes in the network is unknown, and are also not suitable for controller design. As an alternative approach, we propose the use of the S-System modeling formalism to model the gene regulatory network. Through system identification, we are able to obtain realistic model parameters, identify the type of regulation between each gene, and derive a model that is suitable for the design of a synthetic genetic feedback controller. Using the sine sweeping method, the S-System model can be approximated by a second order linear transfer function and, based on this transfer function, we design a genetic phase lag feedback controller. Simulation results show the satisfactory performance of the controller in mitigating external network perturbations. Our proposed modelling and control system design approach has great potential for application in diverse application domains in the field of synthetic biology.

VI. ACKNOWLEDGEMENTS

We gratefully acknowledge the financial support EPSRC and BBSRC via research grants BB/M017982/1 and from the School of Engineering of the University of Warwick.

REFERENCES

[1] Y.-Y. Liu, J.J. Slotine, and A.-L. Barabasi, "Controllability of complex networks", *Nature* vol. 473, no. 7346, pp. 167-173, 2011.

[2] Y. Tang, G. Huijin, D. Wei, L. Jianquan, A.V. Vasilakos, and J. Kurths, "Robust multiobjective controllability of complex neuronal networks", *IEEE/ACM Transactions on Computational Biology and Bioinformatics* vol. 13, no. 4, pp. 778-791, 2015.

[3] Y.-Y. Liu, and A.-L. Barabasi, "Control principles of complex systems", *Reviews of Modern Physics* vol. 88, no. 3, pp. 035006, 2016.

[4] C. Nowzari, V.M. Preciado, and G.J. Pappas, "Analysis and control of epidemics: a survey of spreading processes on complex networks", *IEEE Control Systems* vol. 36, no. 1, pp. 26-46, 2016.

[5] A. Vinayagam, T.E. Gibson, H.-J. Lee, B. Yilmazel, C. Roesel, Y. Hu, Y. Kwon, A. Sharma, Y.Y. Liu, N. Perrimon, and A.-L. Barabasi, "Controllability analysis of the directed human protein interaction network identifies disease genes and drug targets", *Proceedings of the National Academy of Sciences, USA* vol. 113, no. 18, pp. 4976-4981, 2016.

[6] D. Marbach, T. Schaffter, T. Mattiussi, and D. Floreano, "Generating realistic *in silico* gene networks for performance assessment of reverse engineering methods", *Journal of Computational Biology* vol. 16, no.2, pp. 229-239, 2009.

[7] G. Stolovitzky, D. Monroe, and A. Califano, "Dialogue on reverse-engineering assessment and methods: the DREAM of high-throughput pathway inference", In G. Stolovitzky, and A. Califano, editor, *Annals of the New York Academy of Science*, vol. 1115, pp. 11-22, 2007.

[8] G. Stolovitzky, R.J. Prill, and A. Califano, "Lessons from the DREAM2 challenges", In G. Stolovitzky, and A. Califano, editor, *Annals of the New York Academy of Science*, vol. 1158, pp. 159-195, 2009.

[9] M. Savageau, *Biochemical systems analysis: a study of function and design in molecular biology*, Addison-Wesley, Reading, MA, 1976.

[10] E.O. Voit, *Canonical nonlinear modeling. S-System approach to understanding complexity*, Van Nostrand Reinhold, NY, 1991.

[11] S. Kikuchi, D. Tominaga, M. Arita, K. Takahashi, and M. Tomita, "Dynamic modeling of genetic networks using genetic algorithm and S-system", *Bioinformatics* vol. 19, no. 5, pp. 643-650, 2003.

[12] S. Kimura, K. Ide, A. Kashihara, M. Kano, M. Hatakeyama, R. Masui, N. Nakagawa, S. Yokoyama, S. Kuramatsu, and A. Konogaya, "Inference S-system models of genetic networks using cooperative coevolutionary algorithm", *Bioinformatics* vol. 21, no. 7, pp. 1154-1163, 2005.

[13] E.O. Voit, and M. Savageau, "Accuracy of alternative representations for integrated biochemical systems", *Biochemistry* vol. 26, no. 21, pp. 6869-6880, 1987.

[14] L. Ljung, *System identification: theory for the user*, 2nd Edition, Prentice-Hall, Upper Saddle River, NJ, 1999.

[15] T.S. Gardner, D. di Bernado, D. Lorenz and J.J. Collins, "Inferring genetic networks and identifying compound mode of action via expression profiling", *Science* vol. 301, no. 5629, pp. 102-105, 2003.

[16] D. di Bernado, M.J. Thomphson, T.S. Gardner, S.E. Chobot, E.L. Eastwood, A.P. Wojtovich, S.J. Elliot, S.E. Schaus, and J.J. Collins, "Chemogenomic profiling on a genome-wide scale using reverse-engineered gene networks", *Nature Biotechnology* vol. 23, no. 3, pp. 377-383, 2005.

[17] M. Bansal, V. Belcastro, A. Ambesi-Impiombato and D. di Bernado, "How to infer gene networks from expression profiles", *Molecular Systems Biology* vol. 78, 2007.

[18] J. Ang, S. Bagh, B.P. Ingalls, and D.R. McMillen, "Considerations for using integral feedback control to construct perfectly adapting synthetic gene network", *Journal of Theoretical Biology* vol. 266, no.4, pp. 723-738, 2010.

[19] A.W.K. Harris, J.A. Dolan, C.L. Kelly, J. Anderson and A. Papachristodoulou, "Designing genetic feedback controllers", *IEEE Transactions on Biomedical Circuits and Systems* vol. 9, no.4, pp. 475-484, 2015.

[20] T. Soderstrom, and P. Stoica *System identification*, Englewood Cliffs, NJ, 1988.