

Biochemical noise enables a single optogenetic input to control identical cells to track asymmetric and asynchronous reference signals

Michael P May¹ and Brian Munsky²

Abstract—Optogenetics is a powerful technology to control synthetic gene circuits using external and computer-programmable light inputs. Like all biological processes, these systems are subject to intrinsic noise that arises from the stochastic process of gene regulation at the single-cell level. Many engineers have sought to mitigate this noise by developing more complex embedded bio-circuits, but recent work has shown that noise-exploiting stochastic controllers could enable new control strategies that take advantage of noise, rather than working against it. These noise-exploiting controllers were initially proposed to solve a single-input-multi-output stationary control problem, where symmetry was broken in a means reminiscent to the concept of Maxwell’s Demon. In this paper, we extend those results and show through computation that transient, asymmetric, and asynchronous stochastic control of the single-input-multi-output (SIMO) control problem is possible to achieve by cycling through different controllers in time. We show that such a method is able control two cells to two different periodic fates with different frequencies and different phases despite the use of only one control input.

I. INTRODUCTION

The combination of synthetic biology, fluorescence labeling, optical microscopy, and computer controlled optogenetics has created the possibility for cyber-organic control mechanisms that watch cellular behaviors in real time and then actuate internal cellular processes through application of external light inputs [1], [2]. Such circuits could enable many potential applications that require precise spatial and temporal control of individual cells, but unique challenges arise from the unique spatiotemporal scales of the biochemical reactions that define cellular process. Most importantly, biological behaviors are governed by fundamentally stochastic processes due to small numbers of DNA, RNA, and proteins that interact in discrete and randomly timed events – these events can create heterogenous responses within a population of isogenic cells [3]. Because the regulation of genes in synthetic biocircuits also exhibit such behavior, stochastic gene regulation has been treated as a fundamental limit to the controllability of biological systems. Many have tried to build biological systems that reduce noise [4], but recent work has shown that noise can create new control opportunities [5], [6].

*Research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health under award number R35GM124747.

¹Michael P May is with the School of Biomedical Engineering, Colorado State University, Fort Collins, Colorado, 87503. michael.may@rams.colostate.edu

²Brian Munsky is with the Department of Chemical and Biological Engineering and the School of Biomedical Engineering, Colorado State University, Fort Collins, Colorado, 87503. munsky@colostate.edu

Although many models for optogenetically controlled synthetic cell dynamics have utilized ordinary differential equations, linear noise approximations with additive noise, or Monte Carlo methods such as Gillespie’s stochastic simulation algorithm [7], improved computational tools are still needed to improve understanding for these dynamics and to assist efforts to mitigate or exploit randomness in stochastic gene regulation. In this paper, we use direct analyses of the full probability distributions of system responses under transient inputs (i.e., solutions to the chemical master equation) to find deeper understanding for the stochastic cellular process. We show that this analysis can support the design of controllers that can improve performance of optogenetic systems including for SIMO control. We show how the combination of optogenetics, fluorescent microscopy, and noise-exploiting feedback control could yield new control techniques which could able the control of two cells to different periodic fates.

The integration of stochasticity and feedback to stretch the limits of physical possibilities extends back to 1867, when James Maxwell proposed a theoretical demon who could violate the 2nd law of thermodynamics to create a temperature gradient without consuming energy. In essence, the demon uses the random transient state of the system, combined with a controllable trap door to direct fast (hot) molecules to one side and slow (cold) molecules to the other, despite the fact that all molecules sample from the same velocity distribution. In previous work, we proposed a similar optogenetic Maxwell’s demon (OMD) that exploits the transient nature of biological noise in order to direct two identical cells to two different arbitrarily chosen phenotypes. The unique challenge of this control is that the demon cannot just “open the door for a single cell”, but rather if the demon changes the light to control one cell, that same light signal will affect and potentially destabilize the other Fig1(A). The demon must play a difficult balancing game to drive each cell to a different phenotype simultaneously without destabilizing the other.

Although this previous work performed well at breaking the symmetry between the two cells, our previous work only considered the stationary behavior about a single static reference point. When the control of a system to a periodic process is desired [8], such a control method may not be effective. In this paper, we extend the analysis of previous work to enable the stochastic control toward dynamically changing reference points Fig1(A).

In the next section, we introduce a nonlinear stochastic model for two-cell SIMO control, and we demonstrate how to

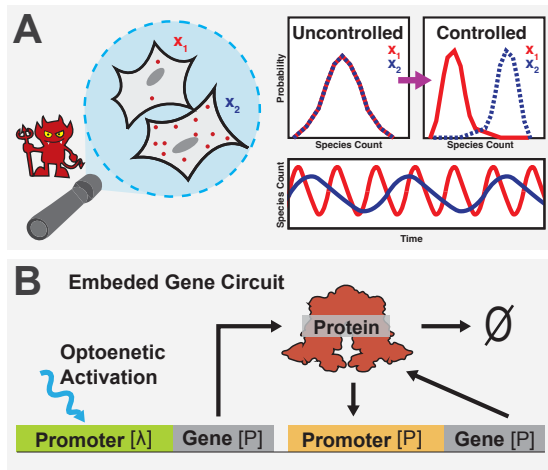


Fig. 1. (A). Shows an OMD which seeks to control a stochastic two cell system to two different reference points despite the fact that they both share the same embedded gene circuit and global light input. (B). Schematic of a non-linear and optogenetically controllable embedded gene circuit for the production of an arbitrary protein.

compute the stationary and transient evolution of probability mass under different control laws. Next, we introduce a metric to evaluate the probabilistic performance error of this controller and devise a gradient based search to optimize the controller to minimize this error. In Section III, we show our results that for the tracking of different dynamic reference signal where two cells are forced to follow different asymmetric and asynchronous trajectories. Finally, in the Discussion Section, we summarize our results and discuss current challenges and future efforts to control multiple different cells using a limited number of control inputs.

II. METHODS

A. Nonlinear Model for Noise-Exploiting Feedback Control

We previously showed that the combination of stochastic single-cell variation, nonlinear gene regulation, and global feedback has the potential to break the symmetry between identical cells and enable the individualized control of multiple cells using a single control input [5]. One possibility for such a system is illustrated in Fig1(B), which depicts a primary actuator promoter that allows for external feedback control and a secondary self-actuated promoter that provides a nonlinear stabilization once the system is activated. Although the primary promoter in the model to be described in this manuscript is based on the optogenetic T7 dimerization system from Baumschlager et al, [1], whose model parameters we fit in previous work [5], other actuator promoters could utilize alternative electronic, optical, or chemical signals to induce gene expression.

To model this system, we assume that the total protein production rate, $H(x)$, due to the two promoters is defined by:

$$H(x) = \kappa \frac{x^\eta}{x^\eta + \beta^\eta} + \alpha + u, \quad (1)$$

where x is the number of protein molecules; κ , η and β are the max-value, Hill coefficient, and half-max activation concentration, respectively, for the self-actuated promoter; u is the strength of the optogenetically controlled promoter;

TABLE I

PARAMETERS OF THE MODEL

Parameter Name	Value	Units
κ	0.406	molecules minute ⁻¹
β	20.0	molecules
η	8.00	
k_0	0.0001	molecules minute ⁻¹
γ	0.0203	minute ⁻¹

and α is the total basal (or leaky) expression level from the two promoters. The protein decay rate is assumed to be a first order reaction with rate:

$$D(x) = \gamma x, \quad (2)$$

where γ represents the degradation rate. Specific values for all parameters are provided in Table I.

For the discrete stochastic description of such a system consisting of two cells, let x_1 denote the integer protein population in the first cell, and let x_2 denote the protein population in the second cell. In this case, there are four possible production events (one for each promoter in each cell) and two degradation events (one for each cell). The stoichiometry for these six possible events can be written in matrix form as:

$$S = [s_1 \ s_2 \ s_3 \ s_4 \ s_5 \ s_6] = \begin{bmatrix} 1 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & -1 \end{bmatrix}, \quad (3)$$

and the corresponding propensity functions (i.e., the stochastic reaction rates) for these reactions are given by:

$$\begin{aligned} \omega^1(\mathbf{x}) &= \kappa \frac{x_1^\eta}{x_1^\eta + \beta^\eta} + \alpha, & \omega^2(\mathbf{x}) &= u(\mathbf{x}; \mathbf{r}), & \omega^3(\mathbf{x}) &= \gamma x_1, \\ \omega^4(\mathbf{x}) &= \kappa \frac{x_2^\eta}{x_2^\eta + \beta^\eta} + \alpha, & \omega^5(\mathbf{x}) &= u(\mathbf{x}; \mathbf{r}), & \omega^6(\mathbf{x}) &= \gamma x_2. \end{aligned} \quad (4)$$

The propensity functions for the optogenetic promoter, ω^2 and ω^5 , are identical due to the fact that a single light-modulated control signal is applied to both cells at the same time. The control signal, $u(\mathbf{x}; \mathbf{r})$, can be any positive valued function of the current state $\mathbf{x} = [x_1, x_2] \in \mathbb{N}_{\geq 0}^2$ and some reference signal $\mathbf{r} = [r_1, r_2] \in \mathbb{N}_{\geq 0}^2$. In other words, u maps $\mathbb{N}_{\geq 0}^4 \mapsto \mathbb{R}_{\geq 0}$. We stress that the control signal $u(\mathbf{x}; \mathbf{r})$ is equal for both cells due to the fact that the global light input that shines on both cells at once and gives rise to the SIMO control system as shown in Fig1(A). The goal of the stationary OMD from [5] was to find an optimal $u(\mathbf{x}; \mathbf{r})$ to direct the discrete stochastic process $\bar{\mathbf{x}}(t)$ to one specific reference point \mathbf{r} . In this manuscript, we explore the possibility to allow \mathbf{r} to vary as a function of time.

B. Chemical Master Equation description for the joint dynamics of two cells under a single control input

The state space for the two-cell process introduced above is the set of all non-negative integer vectors $\mathbf{x} \in \mathbb{N}_{\geq 0}^2$. Using the stoichiometry vectors $\{s_1, \dots, s_6\}$ and propensity functions $\{\omega^1(\mathbf{x}), \dots, \omega^6(\mathbf{x})\}$ defined above, we can state the chemical master equation (CME [9]) to quantify the time rate of change for the probability of each state \mathbf{x} as:

$$\frac{dP(\mathbf{x})}{dt} = - \sum_{k=1}^6 \omega^k(\mathbf{x}) P(\mathbf{x}) + \sum_{k=1}^6 \omega^k(\mathbf{x} - s_k) P(\mathbf{x} - s_k), \quad (5)$$

where the first sum describes flow of probability away from state \mathbf{x} and the second sum describes flow of probability into state \mathbf{x} from its neighbors. We can enumerate the countable state space $\mathbf{x} \in \mathbb{N}_{\geq 0}$ according to: $\mathbf{x}_i = [x_{i_1}, x_{i_2}]$, where the $i = i(i_1, i_2)$ is an arbitrary enumeration scheme to be defined below. Under any such enumeration scheme, we can define the probability mass vector: $\mathbf{P} = [P(\mathbf{x}_1(t), P(\mathbf{x}_2(t), \dots)]^T$, and write the CME in linear matrix form as:

$$\frac{d}{dt}\mathbf{P} = \mathbf{A}_r\mathbf{P}, \quad (6)$$

where the infinitesimal generator, \mathbf{A}_r , for a specific reference point \mathbf{r} is given by:

$$[\mathbf{A}_r]_{ij} = \begin{cases} -\sum_{n=1}^6 w^n(\mathbf{x}_j), & \text{for } i = j, \\ w^n(\mathbf{x}_j), & \text{for } \mathbf{x}_i = \mathbf{x}_j + \mathbf{s}_n, \\ 0, & \text{otherwise.} \end{cases} \quad (7)$$

Although there is an infinite number of possible states, $\mathbf{x} \in \mathbb{R}_{\geq 0}$, one can approximate the CME solution using a truncation such as the Finite State Projection approach [10] for finite times or by adding a reflecting boundary condition for use with steady state analyses. Specifically, for the purposes of this manuscript, we truncate the state space such that $x_1 < N$ and $x_2 < N$, which allows us to use the relatively simple enumeration scheme:

$$i = n \cdot i_2 + i_1 + 1. \quad (8)$$

For all analysis presented here, $n = 50$, such that the number of possible states (and therefore the length of the vector \mathbf{P}) is $(n+1)^2 = 2500$.

To approximate steady state of the CME, we assume that the propensity functions ω^1 and ω^2 are both zero when $x_1 = n-1$ and that ω^4 and ω^5 are both zero when $x_2 = n-1$, such that the number of proteins is not allowed to exceed n molecules in either cell. Under this approximation, the restricted steady state probability distribution can be estimated first by computing the null space of the infinitesimal generator,

$$\mathbf{v}_r = \text{null}(\mathbf{A}_r), \quad (9)$$

and then normalizing this result to have a sum of unity:

$$\mathbf{P}_r^{ss} = \mathbf{v}_r / \sum \mathbf{v}_r. \quad (10)$$

In the above formulation, \mathbf{P}_r^{ss} should be understood as the joint steady state probability distribution over the enumerated state space $\{\mathbf{x}_i \in \mathbb{N}_{\geq 0}^2\}$ under a fixed control law $u(\mathbf{x}; \mathbf{r})$. We will later use $\mathbf{P}_{r(t)}(t)$ to denote the solution of the master equation under a potentially time-varying set of control laws, $u(\mathbf{x}; \mathbf{r}(t))$. In either case, the corresponding marginal distributions $P(x_1)$ or $P(x_2)$ can be calculated by using a linear operations:

$$\mathbf{P}(x_1) = \mathbf{M}_1\mathbf{P}, \text{ or } \mathbf{P}(x_2) = \mathbf{M}_2\mathbf{P}, \quad (11)$$

where \mathbf{P} can be replaced with \mathbf{P}_r^{ss} or $\mathbf{P}_{r(t)}(t)$ and for the appropriate linear marginalization matrix \mathbf{M}_1 or \mathbf{M}_2 .

C. Metric to Quantify Control Performance Error

To rank outcomes of different control laws, we use the expected the mean **squared** error between the cell state and the reference signal,

$$J_r = \mathbf{E}\{|\mathbf{x} - \mathbf{r}|_2^2\}. \quad (12)$$

This score is calculated by summing over all possible states:

$$J_r = \sum_{i_1, i_2} ((x_1 - r_1)^2 + (x_2 - r_2)^2) P(x_1 = i_1, x_2 = i_2), \quad (13)$$

which can be written more concisely in terms of the enumerated state space:

$$J_r = \sum_i (|\mathbf{x}_i - \mathbf{r}|_2^2) P(\mathbf{x}_i) = \mathbf{C}_r\mathbf{P}, \quad (14)$$

where the i^{th} entry of the reference-dependent output vector is given by $[\mathbf{C}_r]_i = (|\mathbf{x}_i - \mathbf{r}|_2^2)$.

D. Optimizing the Control Law to Minimize Error

Using the definitions above, we can now optimize to find a control law $u(\mathbf{x}; \mathbf{r})$ that minimizes the mean squared error for any specific reference point \mathbf{r} (i.e., to minimize J_r). We assume that all control signals must be constrained such that:

$$0 \leq u(\mathbf{x}; \mathbf{r}) \leq 2, \text{ for all } \mathbf{x} \text{ and } \mathbf{r}. \quad (15)$$

For any given fixed \mathbf{r} , this optimization problem requires finding a vector of control values $\mathbf{u}_r = [u(\mathbf{x}_1; \mathbf{r}), u(\mathbf{x}_2; \mathbf{r}), \dots]$ with one entry for every \mathbf{x}_i (i.e., for 2,500 different states in our case). The optimization is conducted using a gradient search, where the effect of the control on the error is:

$$\nabla_{\mathbf{u}} J_r = \nabla_{\mathbf{u}} (\mathbf{C}_r \mathbf{P}^{ss}) = \mathbf{C}_r \nabla_{\mathbf{u}} \mathbf{P}^{ss}. \quad (16)$$

From Eq. 6 at steady state, we can write:

$$\nabla_{\mathbf{u}} (\mathbf{A}_r \mathbf{P}^{ss}) = \nabla_{\mathbf{u}} (0), \quad (17)$$

$$\nabla_{\mathbf{u}} \mathbf{A}_r \mathbf{P}^{ss} + \mathbf{A}_r \nabla_{\mathbf{u}} \mathbf{P}^{ss} = 0, \quad (18)$$

from which we calculate $\nabla_{\mathbf{u}} \mathbf{P}^{ss}$ using the generalized minimal residual method [11].

To estimate an optimized solution, we begin with some initial control law \mathbf{u}_0 and follow the gradient descent procedure according to:

$$\mathbf{u}_{k+1} = \mathbf{u}_k - \delta \mathbf{C}_r \nabla_{\mathbf{u}} \mathbf{P}^{ss}, \quad (19)$$

where δ is a small learning rate. The search defined by Eq. 19 is looped over many iterations $k = 1, 2, \dots$, where on each iteration \mathbf{A}_r and $\nabla_{\mathbf{u}} \mathbf{A}_r$ are updated according to Eq. 7; then \mathbf{P}^{ss} is found using Eqs. 9 and 10; and $\nabla_{\mathbf{u}} \mathbf{P}^{ss}$ is determined from Eq. 18. Finally, after the optimization converges, the control inputs are bounded such that any \mathbf{u} greater than \mathbf{u}_{max} is set to \mathbf{u}_{max} .

E. Control to follow a path of changing reference points

For a single fixed reference point \mathbf{r} , the optimization above performs well to find a locally optimal controller that minimizes a static $J_{\mathbf{r}}$ for that specific reference state. For systems where a changing reference point may be desired, a single reference point solution is no longer sufficient. To extend these results for a dynamic reference signal, we compute the locally optimal steady state controllers at several discrete points along the desired pathway.

For our examples below, we restrict the reference signals to periodic functions with a period T , although non-periodic functions could be considered using the same approach. The piecewise constant definition of the reference signal is given as

$$\mathbf{r}(t) = \mathbf{r}_{\theta(t)} \in [\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N], \quad (20)$$

where the index, $\theta(t)$ is defined as:

$$\theta(t) = \lceil N \frac{\text{mod}(t, T)}{T} \rceil \in [1, 2, \dots, N], \quad (21)$$

and the notation $\lceil \cdot \rceil$ rounds its argument up to the nearest integer. The corresponding piecewise constant controller $\mathbf{u}(t)$ is then defined by selecting the appropriate reference value at the appropriate time:

$$\mathbf{u}(t) = \mathbf{u}_{\theta(t)} \in [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_N]. \quad (22)$$

Finally, to solve for resulting time-inhomogeneous CME, we simply integrate:

$$\frac{d}{dt} \mathbf{P} = \mathbf{A}(\mathbf{u}(t)) \mathbf{P}, \quad (23)$$

but where the infinitesimal generator matrix (Eq. 7) is now a piecewise constant function in time.

As a result of the dynamic changes to the reference point, the resulting probability distribution $\mathbf{P}(t) = \mathbf{P}(\mathbf{x}, t)$ changes continually in time and no longer reaches a fixed steady state value. To visualize the error of the control performance, we define the random error $\tilde{\mathbf{y}}(t) = (\tilde{\mathbf{x}}(t) - \mathbf{r}(t)) \in \mathbb{N}^2$ and translate the computed probability distribution to calculate the joint distribution of the error at any specific instance in time as:

$$\mathbf{P}(\mathbf{y}, t) = \mathbf{P}(\mathbf{x} + \mathbf{r}(t), t). \quad (24)$$

The best possible controller is one where most of the probability mass concentrates at $\mathbf{y} = [0, 0]$, and where very little probability mass accumulates at large magnitudes of \mathbf{y} . In order to visualize the average error (direction and magnitude) over a periodic trajectory, we integrate to average over time:

$$\mathbf{P}_{\text{Ave}}(\mathbf{y}) = \frac{\int_0^{t_{\text{final}}} \mathbf{P}(\mathbf{y}, t) dt}{t_{\text{final}}}. \quad (25)$$

III. RESULTS

Optimized control inputs were found on a square grid of static reference points which minimized the score $J_{\mathbf{r}_k}$ at each point in a coarse grid. Each element of the grid $G \in [r_1, r_2]$ is created by stepping r_1 and r_2 from zero to forty in steps of five for a total of $9^2 = 81$ grid points. Optimal scores $J_{\mathbf{r}_k}$ were found for each static point in the grid and are plotted on the domain of grid points as a heat map Fig 2(A). The heat map shows that some static reference points in the domain are more easily controlled than others, with lower scores (lighter colors) indicating more controllable static reference values and higher scores (darker colors) indicating worse control. Higher scores are seen in the top-left, bottom-right, and the middle region of the heat map Fig 2(A). The edge regions likely have poor control due to the geometry of the system, while the middle point $\mathbf{x} = [20, 20]$ has poor control due to instability of the rate equations for that particular value of \mathbf{x} . The dashed black line in Fig 2(A) shows a line of symmetry in the score values. Transposing the heat map along this line yields a perfectly symmetric result because transposing the control matrix is the same as swapping $u(\mathbf{x}, [r_1, r_2])x \rightarrow u(\mathbf{x}, [r_2, r_1])$.

Figures 2(A, B1, and C1) shows the optimized control laws ($u(\mathbf{x}, \mathbf{r})$) for two examples of static reference points: $\mathbf{r} = [20, 25]$ (green highlighted star) and $\mathbf{r} = [25, 10]$ (blue highlighted star). Figures 2(B2 and C2) shows the stationary joint probability distributions under those fixed-reference controllers. These figures show that controllers optimized to different fixed reference points required very different control laws and result in qualitatively different joint probability distributions. The corresponding marginal distributions depicted in Fig 2(B3, B4, C3, C4) show how symmetry between the two cells has been broken despite using only a single global light input signal and for when both cells share identical gene circuits. Joint and marginal distributions of the responses reveal that tighter control of the probability distribution leads to lower score for the more asymmetric reference point ($J_{[25, 10]} = 68$ in Fig 2.C2) compared to the more symmetric reference point ($J_{[20, 25]} = 110$ in Fig 2.B2).

Using different combinations of the fixed-reference control laws, we specified three controllers to track periodic reference signals:

- 1) a synchronized trajectory given by

$$\mathbf{r}(t) = 10 \left[\sin \left(\frac{2\pi t}{T} \right) + 2, \sin \left(\frac{2\pi t}{T} \right) + 2 \right] \quad (26)$$

where the desired reference points are the same;

- 2) a phase lagged trajectory given by

$$\mathbf{r}(t) = 10 \left[\sin \left(\frac{2\pi t}{T} \right) + 2, \cos \left(\frac{2\pi t}{T} \right) + 2 \right] \quad (27)$$

where the desired two reference points contains a $\pi/2$ phase shift between each other; and

- 3) a phase lagged trajectory with a two different frequencies

$$\mathbf{r}(t) = 10 \left[\sin \left(\frac{2\pi t}{T} \right) + 2, -\cos \left(\frac{6\pi t}{T} \right) + 2 \right] \quad (28)$$

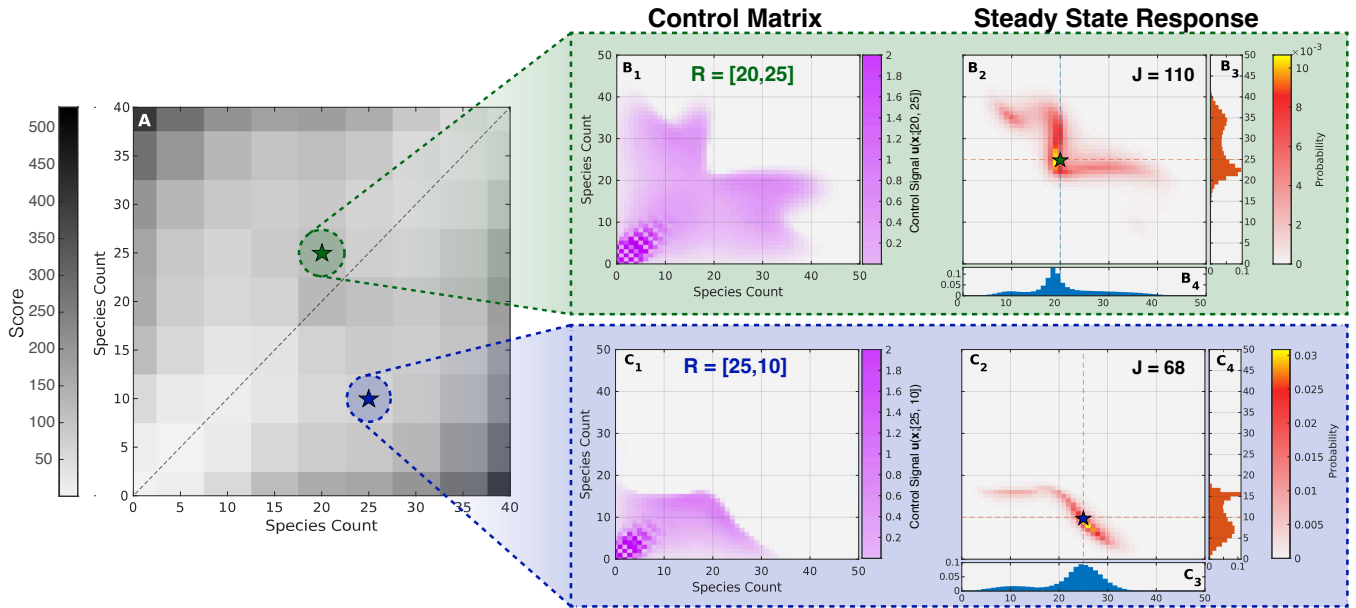


Fig. 2. **Control laws and performance for fixed reference points.** (A) Optimized scores J_r for on a 2D domain with a grid of reference points. Two specific reference points are selected ($r = [20, 25]$, green highlighted star) and ($r = [25, 10]$, blue highlighted star). (B1) Optimized control law $u(x, [20, 25])$ for the desired reference point $[20, 25]$. (B2) Steady state joint probability distribution of the system when using the control input law depicted in B1. (B3 and B4) Marginal distribution $P(x_1)$ and $P(x_2)$, calculated using the joint distribution in B2. (C1-C4) Same as for B1-B4, but for reference point of $[25, 10]$.

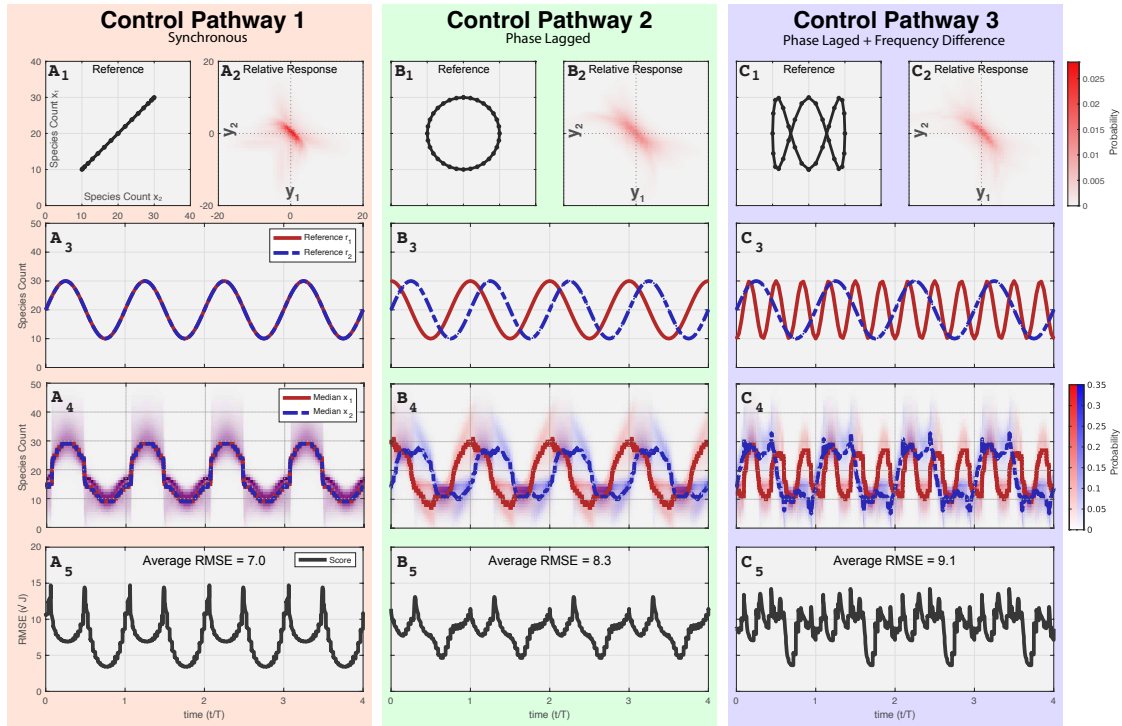


Fig. 3. **(Control performance for transient reference signals.** Three signals are considered: (Column A) synchronous ($r_1 = r_2 = 10 \sin(2\pi t/T) + 20$), (Column B) asynchronous with $\pi/2$ phase lag ($r_1 = 10 \sin(2\pi t/T) + 20$, and $r_2 = 10 \cos(2\pi t/T) + 20$), and (Column C) asynchronous and aperiodic ($r_1 = 10 \sin(2\pi t/T) + 20$, and $r_2 = -10 \cos(6\pi t/T) + 20$). These reference signals are shown in (A1,B1,C1) 2D phase space (r_1 vs. r_2) and (A3,B3,C3) each versus time. (A2,B2,C2) Time-averaged joint distribution of error (Eq. 25) shows that the response remains well centered about the reference. (A4,B4,C4) Median of x_1 (red line) and x_2 (blue line) show that the controlled periodic behavior matches well to the reference signal. FSP analyses show that probability distributions $P(x_1, t)$ (red cloud) and $P(x_2, t)$ (blue cloud) are densely-distributed about the reference trajectory. (A5,B5,C5) Plots of root-mean-square-error versus time; the time-averaged RMSE is shown at the top of each plot.

where the second cell is expected to fluctuate with three times the frequency as the first and with a phase shift.

For each of these 2D reference trajectories, $N = 32$ points

along the pathway are selected and 32 different controllers are optimized for each point on the path. The period is set to $T = 5000$ minutes. For each reference trajectory, the system is analyzed for a total of five periods, where only

the responses over last four periods are shown in each of the plots.

The synchronized reference signal (Fig. 3, Column A) consists of a symmetric reference signal, $r_1(t) = r_2(t)$. Fig 3(A4) shows $\mathbf{P}(x_1, t)$ (red cloud) and $\mathbf{P}(x_2, t)$ (blue cloud) and the median of the probability distributions over time of x_1 and x_2 (red and blue lines) using the pathway control method. The appearance of a purple probability cloud in Fig 3(A4) is due to the overlap of $\mathbf{P}(x_1, t)$ and $\mathbf{P}(x_2, t)$. The similarity between the median in Fig 3(A4) and the reference points Fig 3(A3) suggest that the system was able to be controlled to the dynamic reference signal, although with some error due to the stochastic nature of the process. Fig 3(A5) shows the root-mean-squared-error (RMSE) given by the square root of $J_{\mathbf{r}(t)}(t)$ versus $\frac{t}{T}$. The spikes in the RMSE occur at the points in time when the reference signal crosses through an unstable point of the system's macroscopic dynamics. This point, where $r_1 = r_2 = \beta = 20$ (i.e., at the half maximal state of the self-actuating promoters Hill function in Eq. 1) was also seen to be poorly controllable for fixed reference signals, and shows up as a dark patch at that location in Fig 2(A).

This process was repeated for the circular (phase-lagged) reference trajectory where $r_1(t) = 10(\sin(2\pi t/T)) + 20$, and $r_2(t) = 10(\cos(2\pi t/T)) + 20$. For the $\mathbf{P}(x_1, t)$ (red cloud), $\mathbf{P}(x_2, t)$ (blue cloud) and medians of each (red line, blue line), we find that the solutions generally match the shape and phase lag of the specified reference signals. The data taken together suggest that it is indeed possible to commit two cells to different periodic fates using noise-exploiting feedback control and a single control input. However, we note that the error is not a simple shift from one cell to the other. Rather, the error in cell 1 using clockwise motion matches that of cell 2 in a counter lock wise motion. This effect is clear also in the RMSE error (Fig 3B4) which now spikes four times per period; twice when $r_1 = \beta$ and twice when $r_2 = \beta$.

Finally, the analysis is repeated on a system where the two dynamic reference points have different phases and frequencies given by $r_1(t) = 10(\sin(2\pi t/T)) + 20$, and $r_2(t) = -10(\cos(6\pi t/T)) + 20$ (Figs 3C1,C3). For this system, $\mathbf{P}(x_1, t)$, $\mathbf{P}(x_2, t)$ and the medians are calculated in similar fashion. Looking at Fig 3(C4), we can see that the frequency of the median of x_1 is three times the frequency for x_2 , and they generally match the phase and frequency of the desired reference points. Therefore, we can confirm that two cells can generally be controlled to two different frequencies and different phases using the control method, despite the systems being identical and using only a single global light input. The RMSE error (Fig 3C4) now now spikes eight times per period; twice when $r_1 = \beta$ and six times when $r_2 = \beta$.

Plots of $\mathbf{P}_{Ave}(\mathbf{y})$ in each of the three controllers (Figs 3A2,B2,C2) show that the synchronous pathway seems to have the most probability mass centered near the origin. This controllability in the synchronous pathway can also be confirmed with the RMSE plotted versus time in (Figs 3A5,B5,C5). The plots of $\mathbf{P}_{Ave}(\mathbf{y})$ for the asynchronous

trajectories (Figs 3B2,C2) show more probability mass at larger magnitudes of error, and these results can be quantified by the higher average RMSE of 8.3 and 9.1 respectively.

IV. DISCUSSION

Our results show that a noise-exploiting controller is theoretically capable of controlling a stochastic two cell system to different asynchronous and aperiodic trajectories. To design these controllers, we derived a master equation analyses where the control single input was a function of the current random state of the observed cells $\tilde{\mathbf{x}}(t) = [\tilde{x}_1(t), \tilde{x}_2(t)]$ and a temporally changing reference signal $\mathbf{r}(t) = [r_1(t), r_2(t)]$. We note that ordinary differential equations could not have been used to model this system, as stochasticity is critical to break the system symmetry and enable control.

The controllers discussed above exclusively were optimized assuming steady state reference points and then modified to reflect the changing reference signal. Unfortunately, it takes for the CME to reach steady state upon changes in the control law, so such controllers may not be optimal for trajectories with faster periods. It would be interesting implement controllers to exploit the transient nature of the CME dynamics. For example, one could envision adaptive controllers based on counterdiabatic driving [12] to temporarily sacrifice steady state responses to drive faster transient changes.

REFERENCES

- [1] A. Baumschlager and M. Khamash, "Synthetic Biological Approaches for Optogenetics and Tools for Transcriptional Light-Control in Bacteria," *Advanced Biology*, vol. 5, no. 5, p. 2000256, 2021. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/adbi.202000256>
- [2] G. Perrino, A. Hadjimitsis, R. Ledesma-Amaro, and G.-B. Stan, "Control engineering and synthetic biology: working in synergy for the analysis and control of microbial systems," *Current Opinion in Microbiology*, vol. 62, pp. 68–75, 2021.
- [3] M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain, "Stochastic gene expression in a single cell," *Science*, vol. 297, no. 5584, pp. 1183–1186, 2002.
- [4] T. Wang and M. J. Dunlop, "Controlling and exploiting cell-to-cell variation in metabolic engineering," *Current Opinion in Biotechnology*, vol. 57, pp. 10–16, 2019.
- [5] M. May and B. Munsky, "Exploiting noise, nonlinearity, and feedback to differentially control multiple synthetic cells with a single optogenetic input," pp. 1–28, 2021.
- [6] P. Szymańska, N. Gritti, J. M. Keestra, M. Soltani, and B. Munsky, "Using noise to control heterogeneity of isogenic populations in homogenous environments," *Physical Biology*, vol. 12, no. 4, 2015.
- [7] D. T. Gillespie, "Exact stochastic simulation of coupled chemical reactions," *The journal of physical chemistry*, vol. 81, no. 25, pp. 2340–2361, 1977.
- [8] M. Foo, O. E. Akman, and D. G. Bates, "Restoring circadian gene profiles in clock networks using synthetic feedback control," *NPJ systems biology and applications*, vol. 8, no. 1, pp. 1–11, 2022.
- [9] D. A. McQuarrie, "Stochastic approach to chemical kinetics," *Journal of applied probability*, vol. 4, no. 3, pp. 413–478, 1967.
- [10] B. Munsky and M. Khamash, "The finite state projection algorithm for the solution of the chemical master equation," *Journal of Chemical Physics*, vol. 124, no. 4, 2006.
- [11] Y. Saad and M. H. Schultz, "Gmres: A generalized minimal residual algorithm for solving nonsymmetric linear systems," *SIAM Journal on scientific and statistical computing*, vol. 7, no. 3, pp. 856–869, 1986.
- [12] S. Iram, E. Dolson, J. Chiel, J. Pelesko, N. Krishnan, Ö. Güngör, B. Kuznets-Speck, S. Deffner, E. Ilker, J. G. Scott *et al.*, "Controlling the speed and trajectory of evolution with counterdiabatic driving," *Nature Physics*, vol. 17, no. 1, pp. 135–142, 2021.