Stochastic Modeling and Control of Biological Systems: The Lactose Regulation System of *Escherichia Coli*

A. Agung Julius, *Member, IEEE*, Ádám Halász, *Member, IEEE*, M. Selman Sakar, *Student Member, IEEE*, Harvey Rubin, Vijay Kumar, *Fellow, IEEE*, and George J. Pappas, *Senior Member, IEEE*

***Abstract*—In this paper, we present a comprehensive framework for stochastic modeling, model abstraction, and controller design for a biological system. The first half of the paper concerns modeling and model abstraction of the system. Most models in systems biology are deterministic models with ordinary differential equations in the concentration variables. We present a stochastic hybrid model of the lactose regulation system of *E. coli* bacteria that capture important phenomena which cannot be described by continuous deterministic models. We then show that the resulting stochastic hybrid model can be abstracted into a much simpler model, a two-state continuous-time Markov chain. The second half of the paper discusses controller design for a specific architecture. The architecture consists of measurement of a global quantity in a colony of bacteria as an output feedback and manipulation of global environmental variables as control actuation. We show that controller design can be performed on the abstracted (Markov chain) model and implementation on the real model yields the desired result.**

***Index Terms*—Abstraction, biological control systems, simulation, stochastic hybrid system, systems biology.**

I. Introduction

In THIS paper, we present a framework that consists of modeling, abstraction, and control of a biological system, namely, the lactose regulation system of the *Escherichia coli* bacteria. The conceptual idea behind the paper is captured in Fig. 1. The paper addresses two challenges regarding the lactose regulation system: 1) to develop a model that is structurally correct, based on the known biochemical processes in the system, and 2) to develop a control framework for manipulation of large-scale behavior of the system. Roughly speaking, the paper can be divided into two parts. The first part corresponds to the lower half of the hourglass in Fig. 1 that discusses modeling of the lac regulation system as a stochastic hybrid system. The model presented in this paper is a slight modification of the one presented in our earlier work [1]. We also discuss how the stochastic hybrid model can be abstracted into a two-state continuous-time Markov chain and demonstrate how this abstraction is consistent with the macroscopic behavior of a colony of bacteria.

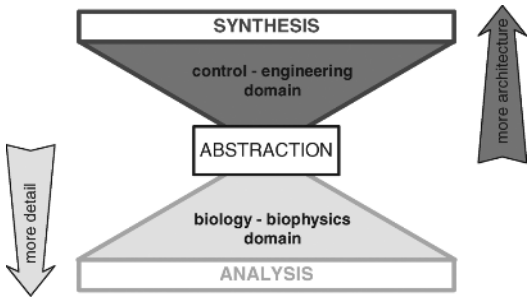


Fig. 1. Hourglass paradigm presented in this paper.

The second part of this paper pertains to the upper half of the hourglass, as it discusses feedback controller synthesis with the abstract model (two-state continuous-time Markov chain) as the plant model. We also demonstrate that the controller designed for the abstract model yields the desired behavior when implemented on the actual system, which is a colony of bacteria. The control goal, in this case, is to drive a certain amount of bacteria to an induced state by manipulating some chemical concentration in the environment.

The first part of the paper thus resides in the domain of biology/biophysics, where modeling is performed. The second part of the paper is of a control/engineering nature, where the problem of controller design and synthesis for a particular control problem is discussed. The abstract model can be viewed as an interface between the two domains, where controller design can be done with the abstract model, without the control engineer having to know about the biology/biophysics aspect of the system.

The extensive variety of species of living organisms makes it impossible for biologists to study every one of them extensively. The fact that many organisms actually share the same operational principles has enabled biologists to focus on just a handful of model organisms. The lactose regulation system in *E. coli* [2] is one of the most extensively studied examples of positive feedback in a naturally occurring gene network. In synthetic biology, the lac operon, which encodes the lactose control system, is often used as a switch to control genes in genetically engineered systems [3], [4]. The lactose regulation system can thus be considered as a model for molecular biological systems that exhibits

Manuscript received January 18, 2007; revised August 20, 2007. This work was supported in part by the National Science Foundation through Presidential Early CAREER (PECASE) Grant 0132716, the NIH-NLM through the Individual Biomedical Informatics Fellowship Award 5-F37-LM008343-1, the Army Research Office through Grant W911NF-05-1-0219, and by a grant from the Penn Genomics Institute. This paper was recommended by Editor M. Vidyasagar.

A. A. Julius, M. S. Sakar, and G. J. Pappas are with the Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia PA 19104 USA (e-mail: agung@seas.upenn.edu; sakarmah@seas.upenn.edu; pappasg@seas.upenn.edu).

A. Halász and V. Kumar are with the Department of Mechanical Engineering and Applied Mechanics, University of Pennsylvania, Philadelphia PA 19104 USA (e-mail: {halasz@seas.upenn.edu; kumar@seas.upenn.edu).

H. Rubin is with the School of Medicine, University of Pennsylvania, Philadelphia PA 19104 USA (e-mail: rubinh@mail.med.upenn.edu).

Digital Object Identifier 10.1109/TAC.2007.911346

/ © 2008 IEEE

multistability. Therefore, the modeling and control framework that we develop in this paper is meant as a model system for other biological systems with similar properties. Other important phenomena that can be attributed to random behavior of multistable biomolecular systems include the bacterial persistence to antibiotics [5] and the trigger mechanism of pathogenic behavior in yeast [6].

A. Control and Cell Biology

Biological systems exhibit many features of complex engineering systems. The origin of system complexity is generally the presence of multiple regulatory mechanisms such as feedback. Among several possible control strategies, feedback seems to be favored in biological systems [7]. Hierarchies of feedback loops result in system robustness, performance, and noise rejection, which are the properties of almost every biological system. Another key feature of biological systems is *emergence*, aggregate behaviors that may not be predicted by only investigating the individual components or subsystems. This is the case in the lac system discussed in this study. The existence of emergent properties can also be explained by the presence of control mechanisms [8].

Khammash and El-Samad [7] reviewed models for two different biological systems and showed the necessity of control actions in order to explain their behavior. Tomlin and Axelrod [9], referring to [10], stated that, if the heat shock mechanism can be described and understood in terms of engineering control principles, it will be informative to apply these principles to other cellular regulatory mechanisms. Indeed, Yang and Iglesias [11] showed that a possible source for the biphasic response in the chemotaxis pathway of *Dictyostelium* is the presence of a positive feedback loop.

Synthetic biologists use similar molecular feedback mechanisms to construct toggle switches [3], [12] and oscillators [13].

Feedback theory and other control systems concepts play an increasing role in *personalized medicine*, where external controls are used to influence cellular processes. The need for treatment procedures that are widely applicable and give satisfactory clinical results for a variety of individuals with different steady-state and dynamic responses to the same treatment emphasizes the use of feedback mechanisms [14]. For every treatment policy, there is an inevitable tradeoff between cost function weighting of pathogens, organ health, and the use of therapeutics. This makes optimal control solutions attractive. Souza *et al.* [15] and Kirschner *et al.* [16] offer an optimal control approach to HIV treatment. Jung *et al.* [17] applies optimal control theory to the treatments in a two-strain tuberculosis model. Stengel *et al.* [18] demonstrated optimal control solutions to the innate immune response. Another control method, model-based predictive control, is also advantageous if accurate models are used in controller synthesis. This is so because predictive control provides us an estimate of future behavior. For instance, Parker *et al.* [19] made a review of control algorithms for noninvasive monitoring and regulation in type-I diabetic patients and showed that model-based predictive control is an attractive choice for blood glucose concentration regulation. There is a recent attempt to apply external control to neuropharmacology [20]. Finally, we also find applications of control laws to build effective strategies in gene therapies and tissue engineering [21].

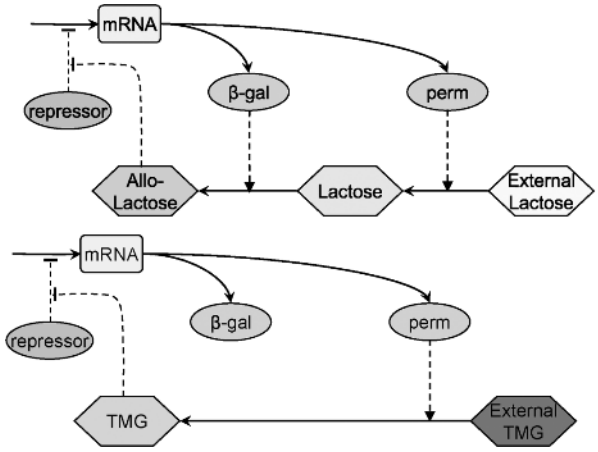


Fig. 2. Lactose network (top) and its modification with the gratuituous inducer *thio-methyl galactosidase* (TMG) (bottom).

B. Lactose Metabolism Control Network

Refer to the diagram in the upper panel of Fig. 2. Two of the lac operon's component genes encode enzymes (β-galactosidase and permease) which contribute to lactose uptake respectively to the synthesis of allolactose. In turn, allolactose acts as an inducer for the operon itself.

Hysteresis and bistability on the level of the entire bacterial population was identified early on by Monod and Pappenheimer [22]. Novick and Weiner [23] discovered bistability at the level of individual cells by studying the expression of β-galactosidase in a population of identical *E. coli* cells. They showed that cells were essentially in one of two discrete states: either fully induced, with enzyme levels close to maximum, or uninduced, with negligible enzyme levels. The observation of intermediate activity on the level of the entire population reflects comparably sized subpopulations of induced and uninduced bacteria. The population heterogeneity was interpreted by Novick and Weiner as a result of a bistability of the gene expression mechanism of individual cells combined with stochastic fluctuations inherent to biomolecular processes involving few molecules.

The notion of positive feedback in a genetic network in the form of autocatalytic gene expression in the lac system has motivated significant work in the context of dynamical models, starting from the early sixties (see the review in [24]). It was well known that positive feedback and delays can result in multiple stable equilibrium points and limit cycles, as is the case in many other biological models. Different models were proposed to study the conditions for stability, the possibility of oscillations, and the effect of time delays due to transcription and translation in the lac system (see [25] for a review). However, a significant fraction of this early work would now be categorized as theoretical, as authors were modeling the interactions with functional dependencies and parameter values that were not grounded in measurements. As the quality and scope of available experimental data improved, this direction of research gradually led to more detailed dynamical models, which explicitly incorporated all relevant biochemical processes along with experimentally motivated kinetic constant values. The work of

Yildirim and Mackey [26] is an example of this new generation of experimentally grounded dynamical modeling.

While much of the modeling of biochemical reactions is based on deterministic models with ordinary differential equations (ODEs), it is well known that these models do not satisfactorily explain the behavior of systems with very low concentrations in which the continuum model is not applicable. Methods for stochastic simulations of biochemical reactions have been developed [27]–[29]. Autocatalytic gene expression has also been studied in the stochastic context for example [30], [31]. In the past few years, it has been recognized that stochastic phenomena may have a crucial role in the fate of individual cells [32]. Multistability and stochastic transitions between equilibrium states were found to have a role in the phenomenon of bacterial persistence [5] and, more generally, are seen as an evolutionary strategy for survival in varying environments [33].

The motivation for a detailed stochastic model of the lactose induction mechanism comes from the work of the group of van Oudenaarden [34], [35], in which they used fluorescent labeling techniques allowing *in vivo* observation of individual cells. They showed a distinct bimodal distribution of the activity of the *lac* operon in individual cells. Thus, population-averaged continuous changes with inducer concentration reflect changes in the relative size of the induced and uninduced populations, reinforcing the conclusions of Novick and Weiner [23]. The experimental results of Mettetal *et al.* were well summarized by a simple empirical model. A more elaborate model, highlighting the role of individual biochemical processes, is desirable. A promising starting point is the model by Yildirim and Mackey [26]. While it predicts bistability and explains the hysteretic switching between steady states, it does not explain the bimodal distribution of β-galactosidase [23] and lactose operon activity [34] that have been experimentally observed.

This experimental picture raises questions about the meaning of ODE-based models such as the Yildirim–Mackey model. This model accounts for the molecular processes responsible for the observed bistability of the *lac* operon and consequently predicts bistability which is indeed exhibited, but on the level of individual cells. However, for intermediate lactose concentrations, both induced and uninduced cells were observed in hysteresis experiments. This indicates that cells are never truly confined to either of the two available steady states, but can "spontaneously" escape and transition into the other equilibrium. This calls for taking into account the stochastic nature of molecular processes in the Yildirim–Mackey model if it is to be applied to individual cells.

However, the parameter values of the Yildirim–Mackey model are not necessarily the correct ones to describe the behavior of individual cells, as the authors themselves caution at the end of their paper. While some model parameters were taken from independent measurements, others were fitted to experimental results. These results, similarly to some of the independent parameter measurements, were obtained on large populations of cells. Therefore, the Yildirim–Mackey model (with its original parameter values) was meant to be a model of *bulk* behavior rather than of individual cells. This bulk interpretation is also problematic. Due to the fact that cells transition spontaneously between available steady states, bulk cell populations do not in fact exhibit true bistability; rather, the two populations eventually equilibrate, leading to a *unique* stable bulk steady state for any external lactose concentration, in stark disagreement with the original Yildirim–Mackey model.

We are interested in obtaining a model that is consistent with both the cell-level and the macroscopic phenomenology. This is a current topic that attracts many systems biologists. See, for example, the recent work reported in [36]. In this paper, we construct a stochastic hybrid model for the lactose regulation system using the work of Yildirim and Mackey as a starting point. Stochasticity in the system naturally arises due to the low copy numbers of molecules involved in the reaction within the cell [10], [37]. In this situation, a model where the reactions are viewed as discrete Poisson random processes is more accurate than a deterministic one [27]. However, due to the incurred computational cost, we choose to use a model in which only those reactants with small copy numbers are modeled as discrete quantities, while the others are modeled as continuous concentrations.

We show that our stochastic hybrid model is able to reproduce the spontaneous transitions that are impossible to capture in any deterministic model. Further, the steady-state behavior of a bulk of cells simulated with our model demonstrates agreement with the predicted equilibria of the Yildirim–Mackey model. Furthermore, we extract a finite-state abstraction of the hybrid stochastic model, which is structured as a two-state continuous time Markov chain [38]. We demonstrate that, despite the simplicity of the abstraction, it can describe the average (macroscopic) behavior of a colony of *E. coli* bacteria, each of which is simulated with the hybrid stochastic model.

It has been shown experimentally [39] that, by manipulating the glucose concentration, the bistable regime can be swapped for a graded response. Investigation of a mathematical model of the lactose system [40] revealed that changes in the basal transcription rate of the lac operon can also modify the bistable behavior. Increased basal transcription leads to a graded response, but, perhaps more interestingly, a reduced basal rate leads into a regime where an induced state exists but is "classically" unreachable, that is, the threshold for induction by lactose is infinite.

The possibility of converting the bistable response into a graded one by manipulating the kinetic properties of the switching mechanism may not be satisfactory in all synthetic biological applications. We would like to investigate whether we can fix the percentage of induced cells at the level of a bacterial population by a macroscopic controller design while maintaining the underlying bistable behavior on the level of individual cells. Here, to satisfy the design criteria, we would like to adjust the amount of external TMG concentration as if it was a therapeutic agent.

The remainder of this paper is organized as follows. In Section II, we present the deterministic and stochastic mathematical models of the lactose regulation system. In Section III, we discuss the construction of the finite-state abstraction of the full model presented in Section II. The feedback control architecture is presented in Section IV, followed by the proposed control algorithms and their simulation results. We conclude the paper with Section V, where we present a few potential future research directions.

II. Modeling the Lactose Regulation System

A. Deterministic Continuous Model

Our starting point is a model of the lactose system due to Yildirim and Mackey [1], [26]. Briefly, the mRNA (M) transcribed from the lactose operon is translated into three different gene products, among them permease (P) and β-galactosidase (B). Permease facilitates the influx of lactose (L) from the exterior and also an opposing process, equilibrating the concentration of lactose inside the cell with the external lactose. The enzyme β-galactosidase has a dual role; it converts lactose to allolactose (A) and converts allolactose further to glucose and galactose. The control loop is closed by the effect of allolactose (A) on the transcription of the *lac* operon. This complicated relationship involves substances not explicitly considered in the Yildirim–Mackey model and results in the nonlinear activation function summarized by the first and second terms in (1a).

The available experimental results, including those used to validate the Yildirim–Mackey model, refer to "gratuitous" induction by substances similar to lactose such as TMG. Such gratuitous inducers, which are not processed by the cell, are often preferred in experimental settings because their presence does not lead to increased growth rate. From a modeling perspective, using TMG instead of lactose also breaks one of the feedback loops in the Yildirim-Mackey model, since β-galactosidase does not act on TMG, and TMG itself can play the inducer role played by allolactose in the full Yildirim–Mackey model, as ilustrated in the lower panel of Fig. 2. The equations of motion for induction by TMG (T) are as follows:

$$\frac{dM}{dt} = \alpha_M \frac{1 + K_1 \left(e^{-\mu\tau_M} T(t-\tau_M)\right)^n}{K + K_1 \left(e^{-\mu\tau_M} T(t-\tau_M)\right)^n} + \Gamma_0 - \tilde{\gamma}_M M \quad (1a)$$

$$\frac{dB}{dt} = \alpha_B e^{-\mu\tau_B} M(t-\tau_B) - \tilde{\gamma}_B B \quad (1b)$$

$$\frac{dT}{dt} = \alpha_L P \frac{T_e}{K_{T_e} + T_e} - \beta_L P \frac{T}{K_{L_1} + T} - \tilde{\gamma}_L T \quad (1c)$$

$$\frac{dP}{dt} = \alpha_P e^{-\mu(\tau_P+\tau_B)} M(t-\tau_P-\tau_B) - \tilde{\gamma}_P P. \quad (1d)$$

We take into account time delays due to transcription and translation. Variables without an argument are taken at time t, and time delays are indicated by an explicit argument, e.g., $M(t-\tau_B)$ is the value of the variable M delayed with τ_B.

The symbol T_e in (1c) signifies the external TMG concentration. If the system is to be viewed as an input-state system, then T_e can be thought of as an input to the system, while the other four concentrations are the state variables. The variable B, which represents β-galactosidase, is effectively decoupled from the equations of motion and has no effect on the dynamics of the remaining three variables. However, it is the experim ental quantity that is traditionally used as the observable, for example, in [23]. The other symbols in the equation are constant parameters given by Table I, together with the following relations:

$$\tilde{\gamma}_M = \gamma_M + \mu, \tilde{\gamma}_B = \gamma_B + \mu \quad (2)$$

$$\tilde{\gamma}_A = \gamma_A + \mu, \tilde{\gamma}_P = \gamma_P + \mu \quad (3)$$

where μ is the growth rate. The values of the constants are based on those in [26] but have been modified to give consistent behavior to the TMG model in the limit of a large but finite cell population.

TABLE I
Constant Parameter Values

	Value	Unit		Value	Unit
μ	$2.26 \cdot 10^{-2}$	min^{-1}	K_{T_e}	$6.5 \cdot 10^{-4}$	mM
γ_M	0.411	min^{-1}	γ_B	$8.33 \cdot 10^{-4}$	min^{-1}
γ_A	0.52	min^{-1}	Γ_0	$1.0 \cdot 10^{-6}$	mM/min
K	7200		α_M	$9.97 \cdot 10^{-4}$	mM/min
τ_B	2.0	min	K_1	$6.3 \cdot 10^{5}$	$(\text{mM})^{-2}$
K_{L_1}	0.36	mM	α_B	$1.66 \cdot 10^{-2}$	min^{-1}
τ_P	0.83	min	β_L	546.32	min^{-1}
τ_M	0.1	min	α_P	10.0	min^{-1}
γ_L	1.52	min^{-1}	γ_P	0.6274	min^{-1}
α_L	81	min^{-1}	n	2	

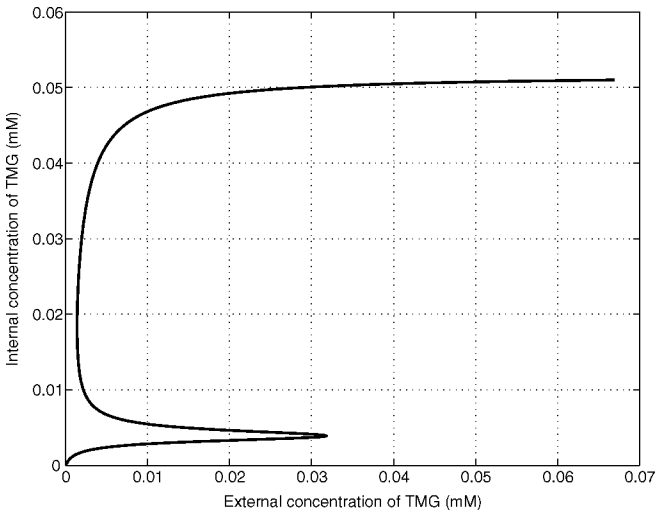


Fig. 3. Equilibria of the system given by (1). The middle range of T_e has three branches of equilibria.

When the value of T_e is maintained between 1.4–32 μM, the system has three equilibria. Two of these equilibria are stable, giving rise to bistability of the system. Also, varying the value of T_e causes a hysteresis behavior. See Fig. 3 for the illustration. The model equation (1) qualitatively reproduces the observed experimental behavior. The higher dimensional version defined in [26] and discussed in [1] very closely approximates that behavior. However, that model (with its original parameter set) is correct only as a description of the bulk behavior of a large number of cells described as a single "reactor." This is because stochastic behavior on the level of individual cells is ignored both in its construction and in its validation.

It has been recognized early on that the observed concentrations on the level of a very large number of cells is actually an average over two distinct subpopulations of cells, whose β-galactosidase level takes one of two extreme values. This microscopic bistability was termed the "all-or-none" phenomenon [23].

There is an apparent discrepancy between the macroscopic and microscopic behavior in the *lac* system. Hysteresis observed on the macroscopic level is indicative of an underlying bistability. However, a closer examination of traditional induction experiments reveals that the observed two extreme states, completely induced and uninduced, are in fact transient, and

only an intermediate concentration can be maintained indefinitely on the bulk level. Bistability is the determining feature on the level of individual cells, which spend relatively little time in intermediate induction states. The observed macroscopic timescale of switching reflects the rate of occurrence (low probability density) of fast stochastically triggered transitions of individual cells rather than the time spent *en route* between the two microscopic steady states.

To reconcile the discrepancy, we need to take into account the stochastic nature of the behavior of individual cells. The macroscopic model needs to be modified in two respects: 1) introduce stochastic phenomena and 2) adjust the kinetic parameters of the model to obtain the correct *macroscopic* phenomenology. A similar program has been carried out by Mettetal *et al.* [35], without connecting the resulting model to kinetics of individual biochemical processes in the way the Yildirim–Mackey model does.

B. Stochastic Model

While stochasticity is sometimes thought of as leading to small deviations from the ODE prediction, it actually may often lead to qualitatively different behavior [41]. There are several sources of stochasticity in the biochemistry of individual cells [42]. In this paper, we will focus on one major source, intrinsic noise generated by *low copy numbers* of molecules. The deterministic ODE model relies on continuously varying concentrations, which is a good approximation when the substances are available in huge molecule numbers. If we consider chemical reactions within a cell, whose volume is of the order of 10^{-16} l [43], the number of molecules involved in the reaction may not be too large. This is especially the case if the concentration of the chemical substances is low.

Chemical reactions, at the microscopical level, amount to creation and breaking up of chemical molecules. These processes can be modeled as Poisson random processes [27], [30], whose rates depend on the state of the system, i.e., the number of molecules in the reaction. In fact, the reaction rates given by the ODE can be considered as the rates of the Poisson processes. This is not the only way to introduce stochasticity to the system. Another approach is to use an ordinary differential model perturbed by stochastic noise [44]. However, we argue that modeling the chemical reactions as Poisson processes is more physically founded. There has also been previous work where stochasticity is introduced by modeling chemical reaction as Poisson processes [45]. However, the underlying reaction model is based on empirical observation rather than physical modeling as ours.

Construction of stochastic models for biochemical processes is a well-established procedure supported by a large amount of literature [46].

We develop a hybrid stochastic model for the system. The model is based on the idea that the messenger RNA (M) and the β-galactosidase (B) are expressed as molecule counts that evolve following some Poisson processes, while the other substances, internal TMG (T) and permease (P), are expressed as chemical concentrations that evolve following deterministic ODE. A similar approach, i.e., part stochastic and part deterministic simulation for chemical processes, is reported in [37]. The reason behind this idea is that a fully stochastic model is computationally expensive, while a hybrid model already demonstrates the stochastic noise that is lacking in the deterministic model.

We are interested in the phenomenology of a model with the structure of that in [47], incorporating the presence of realistic level of noise. The relative importance of stochastic fluctuations of one concentration will be the largest for the species with the lowest concentrations. We choose to discretize M and B, the two species with the lowest steady-state concentrations. This computationally less expensive approach allows us to perform quasi-simultaneous simulations for many cells. The connection between the ODE and the stochastic description is through the conversion constant C_N as

$$\begin{aligned} C_N &= 10^{-16} \cdot 6.023 \cdot 10^{23} \frac{\text{molecules}}{\text{mole}} \cdot 10^{-3} \frac{\text{M}}{\text{mM}} \\ &= 6.023 \cdot 10^{4} \frac{\text{molecules}}{\text{mM}}. \end{aligned}$$

In terms of stochastic differential equations, our hybrid stochastic model can be written as follows:

$$dM_t = d\hat{M}_t - d\tilde{M}_t \tag{4a}$$

$$dB_t = d\hat{B}_t - d\tilde{B}_t \tag{4b}$$

$$\frac{dT_t}{dt} = \frac{T_e \alpha_L P_t}{K_{L_e} + T_e} - \frac{\beta_L P_t T_t}{K_{L_1} + T_t} - \tilde{\gamma}_L T_t \tag{4c}$$

$$\frac{dP_t}{dt} = \alpha_P e^{-\mu(\tau_P + \tau_B)} \frac{M_{(t-\tau_P-\tau_B)}}{C_N} - \tilde{\gamma}_P P_t. \tag{4d}$$

Here, the processes $\hat{M}_t$ and $\tilde{M}_t$ are the Poisson processes that are responsible for the creation and breaking up of the messenger RNA molecules, respectively. Similarly, $\hat{B}_t$ and $\tilde{B}_t$ are the Poisson processes that are responsible for the creation and breaking up of the β–galactosidase molecules, respectively. The rates of these processes are state dependent and are given as follows:

$$\lambda_{\hat{M}}(t) = C_N \left[\alpha_M \frac{1 + K_1 \left(e^{-\mu \tau_M} T_{(t-\tau_M)} \right)^n}{K + K_1 \left(e^{-\mu \tau_M} T_{(t-\tau_M)} \right)^n} + \Gamma_0 \right] \tag{5a}$$

$$\lambda_{\tilde{M}}(t) = \tilde{\gamma}_M M_t \tag{5b}$$

$$\lambda_{\hat{B}}(t) = \alpha_B e^{-\mu \tau_B} M_{(t-\tau_B)} \tag{5c}$$

$$\lambda_{\tilde{B}}(t) = \tilde{\gamma}_B B_t. \tag{5d}$$

C. Stochastic Simulation Algorithm

We simulate the stochastic model (4) using a numerical scheme similar to the explicit tau-leaping method for Gillespie simulation [28], [48]. We pick a constant integration step δ and discretize (4). We use the following notation:

$$\begin{aligned} M[k] &:= M_{t=k\delta} \\ B[k] &:= B_{t=k\delta} \\ P[k] &:= P_{t=k\delta} \\ T[k] &:= T_{t=k\delta}. \end{aligned}$$

Since we also have to discretize the time delay, we define

$$k_M := \left\lfloor \frac{\tau_M}{\delta} \right\rfloor$$
$$k_B = \left\lfloor \frac{\tau_B}{\delta} \right\rfloor$$
$$k_P := \left\lfloor \frac{\tau_B + \tau_P}{\delta} \right\rfloor \tag{6}$$

where

$$\forall x \in \mathbb{R}_+, \lfloor x \rfloor := \max\{n \in \mathbb{Z} | n \leq x\}.$$
$$M[k+1] = M[k] + \Delta\hat{M}[k] - \Delta\tilde{M}[k] \tag{7}$$
$$B[k+1] = B[k] + \Delta\hat{B}[k] - \Delta\tilde{B}[k] \tag{8}$$
$$T[k+1] = T[k] + \delta \left[\frac{\alpha_L T_e[k] P[k]}{K_{L_e} + T_e[k]} - \frac{\beta_L P[k] T[k]}{K_{L_1} + T[k]} - \tilde{\gamma}_L T[k] \right] \tag{9}$$
$$P[k+1] = P[k] + \delta\alpha_P e^{-\mu(\tau_P + \tau_B)} \frac{M[k - k_P]}{C_N} - \delta\tilde{\gamma}_P P[k]. \tag{10}$$

The terms $\Delta\hat{M}[k]$, $\Delta\tilde{M}[k]$, $\Delta\hat{B}[k]$, and $\Delta\tilde{B}[k]$ are the approximation of the increments of the Poisson processes in (25). For example

$$\Delta\hat{M}[k] \approx \int_{k\delta}^{(k+1)\delta} d\hat{M}_t. \tag{11}$$

In the approximation, these terms are modeled as random variables with Erlang distribution, whose expectations are given as follows:

$$E\left[\Delta\hat{M}[k]\right] = \delta C_N \alpha_M \frac{1 + K_1 \left(e^{-\mu\tau_M} A[k - k_M]\right)^n}{K + K_1 \left(e^{-\mu\tau_M} A[k - k_M]\right)^n} + \delta C_N \Gamma_0 \tag{12}$$
$$E\left[\Delta\tilde{M}[k]\right] = \delta\tilde{\gamma}_M M[k] \tag{13}$$
$$E\left[\Delta\hat{B}[k]\right] = \delta\alpha_B e^{-\mu\tau_B} M[k - k_B] \tag{14}$$
$$E\left[\Delta\tilde{B}[k]\right] = \delta\tilde{\gamma}_B B[k]. \tag{15}$$

The probability density function of a random variable X with Erlang distribution and expectation $\bar{X}$ is given as

$$Pr\{X = n\} = \frac{\bar{X}^n e^{-\bar{X}}}{n!} \qquad \forall n \in \mathbb{Z}_+.$$

Algorithm 1: The simulation algorithm can be summarized as follows.

Step 1) Initialize with initial conditions $M[0]$, $B[0]$, $T[0]$, $P[0]$, $k = 0$.

Step 2) Draw the independent Erlang random variables $\Delta\hat{M}[k]$, $\Delta\tilde{M}[k]$, $\Delta\hat{M}[k]$, and $\Delta\tilde{M}[k]$ from the distribution (12)–(15).

Step 3) Compute $M[k+1], B[k+1], T[k+1], P[k+1]$ using (7)–(10).

Step 4) Increase the counter k. Return to step 2).

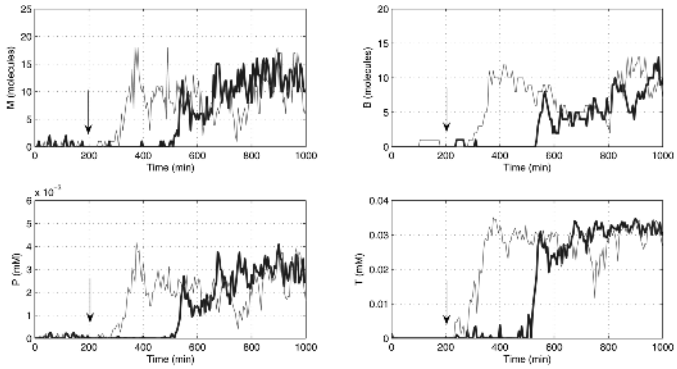


Fig. 4. Simulation results of two cells. In this plot, the external concentration of TMG is increased at $t = 200$ min (marked by the arrows). We observe that spontaneous inductions occur approximately 200 and 400 min later.

A pair of typical simulation traces for individual cells are shown in Fig. 4. In the simulation, we initialize the system at the low equilibrium of $T_e = 1.4$ μM. For the first 200 min of the simulation, the external TMG concentration T_e is kept at 1.4 μM. At $t = 200$ min, T_e is increased to 2.0 μM. We can observe that the time it takes for each cell to induce varies significantly. The aggregate for many simulations of individual cells under the same circumstances shows a gradual increase of average enzyme activity extending over a much longer time interval.

III. Finite-State Abstraction of the Stochastic Model

Here, we discuss a finite-state abstraction of the stochastic model (4). Our goal is to construct an abstraction of the stochastic model that is simple enough to allow for fast computation. This is particularly desirable, for example, when we want to simulate the behavior of a colony of bacteria. Without the abstraction, we would have to run multiple copies of the stochastic simulation described in the previous section, which can be computationally prohibitive.

Our choice of abstraction is guided by the fact that the system dwells in one of the stable equilibria in the time intervals between spontaneous transitions. Assuming that the transitions take place in a shorter time scale than the dwell time, we choose an abstraction in the form of a two-state continuous-time Markov chain [38]. The states of the Markov chain correspond to the low and high stable equilibria of the systems. The rates of switching between the two states are given as a function of the external TMG concentration T_e. See Fig. 5 for a diagram of the system.[1]

This simple model can closely reproduce the bulk behavior of a large number of cells. We run the simulation of the full model [given in (4)], where the system is initialized at the low equilibrium of $T_e = 2.0$ μM. The external concentration of TMG is kept at $T_e = 2.0$ μM. We simulate a colony of 1000 cells. The macroscopic behavior of the colony, which is computed as the average across the 1000 samples are plotted in Fig. 6. It is this

[1]A method for approximately abstracting stochastic hybrid systems is presented in [49]. The nonlinear dynamics in this paper makes the implementation of the method challenging. However, it is noteworthy that there are more systematic ways of abstracting stochastic hybrid systems.

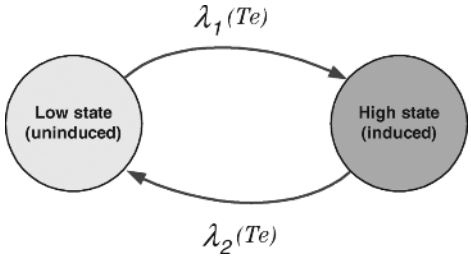


Fig. 5. Two-state continuous-time Markov chain model.

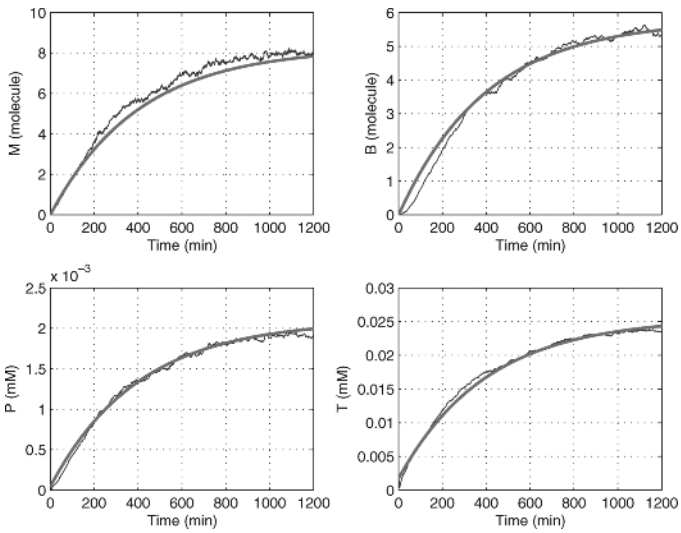


Fig. 6. Average behavior of a colony with 1000 bacteria from stochastic simulations. The exponential curve is plotted to show that the macroscopic behavior can be fitted quite well with linear dynamics [given in (16)].

colony-level average behavior that is observed in macroscopic experiments such as that of Novick and Weiner.

Given the continuous-time Markov chain model as in Fig. 5, we can compute the probability distribution of the states as follows. Define $x_{\rm lo}(t)$ and $x_{\rm hi}(t)$ as the probability of finding the system at time t in the low and high state, respectively. The probability distribution satisfies the following differential equation:

$$\frac{d}{dt}\begin{bmatrix} x_{\rm lo} \\ x_{\rm hi} \end{bmatrix} = \begin{bmatrix} -\lambda_1(T_e) & \lambda_2(T_e) \\ \lambda_1(T_e) & -\lambda_2(T_e) \end{bmatrix} \begin{bmatrix} x_{\rm lo} \\ x_{\rm hi} \end{bmatrix}. \quad (16)$$

The switching rates λ_1 and λ_2 are the inverse of the average switching time from uninduced to induced and vice versa. In this paper, we compute these rates by running a large number of Monte Carlo simulations described in the previous section. From these simulations, we compute the average switching times and take the inverse to obtain the switching rates.

Let the concentration of one of the substances, for example, the mRNA (M), in the low state and high state be denoted by $M_{\rm lo}$ and $M_{\rm hi}$, respectively. Further, let the random process $M_t^{\rm av}$ be the average of the value of the mRNA concentrations in the colony of 1000 cells. Denote the concentration of mRNA in the ith cell by M_t^i. Since the processes $\{M_t^i\}$, $1 \le i \le 1000$, are mutually independent and identically distributed, $M_t^{\rm av}$ is an unbiased estimator for $E[M_t^i]$, which is given as follows:

$$E\left[M_t^i\right] = \begin{bmatrix} M_{\rm lo} \\ M_{\rm hi} \end{bmatrix}^T e^{\begin{bmatrix} -\lambda_1(T_e) & \lambda_2(T_e) \\ \lambda_1(T_e) & -\lambda_2(T_e) \end{bmatrix} t} \begin{bmatrix} x_{\rm lo}(0) \\ x_{\rm hi}(0) \end{bmatrix} \quad (17)$$

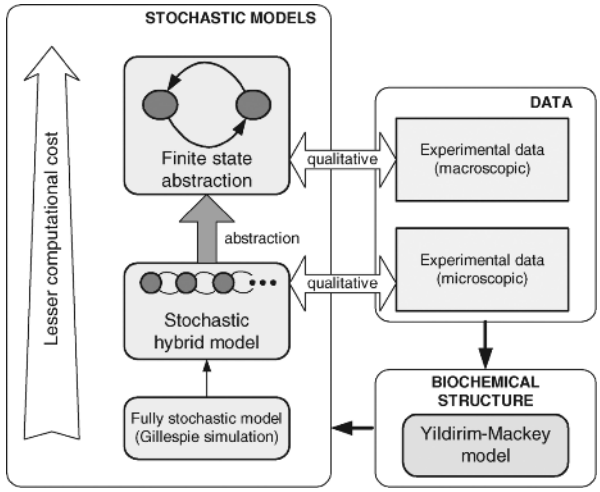


Fig. 7. Block diagram summarizing the models presented in this paper. Our stochastic hybrid model and its finite-state abstraction qualitatively reproduce experimental observations. These stochastic models are based on the biochemically founded deterministic ODE model of Yildirim and Mackey.

which tells us that it is a linear combination of two exponential functions.

In Fig. 6, we compare the simulated average curves and the computed expectation based on (17). We observe that the average of many individual simulations is well matched by the two-exponential behavior. This is a strong argument for the validity of our abstraction.

The models that we have constructed in this paper can be summarized in Fig. 7. Our stochastic hybrid model matches experimental data qualitatively, in the sense that it reproduces the spontaneous transitions between stable equilibria in the bistable zone. This is not possible in the deterministic model [26]. The proposed finite-state Markov chain abstraction is faithful to the stochastic hybrid model, in the sense that it captures the macroscopic dynamics of the stochastic hybrid model. Moreover, a similar exponential-like behavior is also observed in experiments, as reported in [23], [50], and [51].

Remark 2: Notice that we begin the discussion in this paper with an ODE model (1). We then proceed by arguing that the small size of the cells requires stochasticity in the model and proposed a stochastic hybrid model (4). Subsequently, we propose another model that results in another ODE model (16), which is suitable for modeling the macroscopic behavior of the system. To avoid any confusion, we would like to clarify that we do not return to the starting point of the discussion. Indeed, the ODEs (1) and (16) are fundamentally different. While (1) basically describes the dynamics of concentration in a *single large cell* and thus unable to explain the spontaneous transitions between the stable equilibria, the model (16) described the dynamics of a *large colony of small cells* and use the spontaneous transitions to account for the dynamics.

As mentioned earlier, the goal of having the abstraction is to have a model simple enough to be used for computation and design for a large-scale system, that is, a system involving a large number of bacteria. In this macroscopic point of view, a bacterium is represented as a simple Markov chain (Fig. 5) rather than the complicated stochastic hybrid model (4). This idea is captured in Fig. 1.

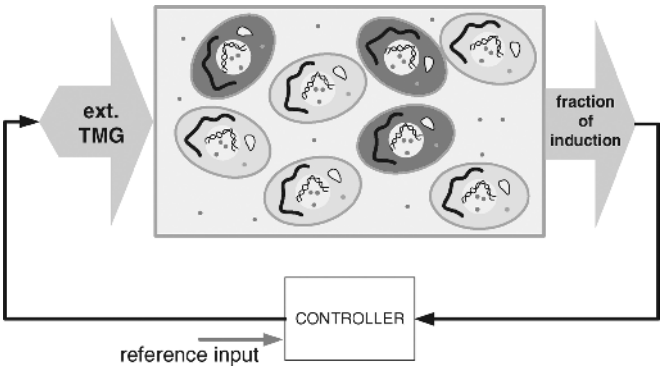


Fig. 8. Control block diagram.

We will rely on the two-level abstraction (16) to design a control strategy for obtaining a graded response from a colony of individual bacteria by modulating the external TMG. We will use simulation on the full model for validation.

IV. Controlling the System

A. Control Architecture

The architecture of the control system that we discuss in this paper is illustrated in Fig. 8. The plant to be controlled is a large colony of *E. coli* bacteria. The controller affects the plant by adjusting of the external concentration of TMG in the environment. Feedback information is read from the plant in the form of a global quantity, which we consider as the *output* of the control system. By this, we mean the controller does not have any information about the individual cells in the colony. Rather, the controller relies on sensing a global quantity, for example, the fraction of induced cells in the population. The control goal is to make the output track a given reference trajectory or attain a desired level. Similar control architecture, where feedback control of a group of Markov chains by adjusting the transition rates has been studied, for example in [52]. There, the plant is a group of artificial muscle cells that can switch between contracting and noncontracting states.

Control actuation by means of adjusting the external concentration of TMG in the environment can be realized as follows. Increasing the concentration can be done, for example, by injecting the enzyme into the plant. There are a number of limitations associated with this method. First, the concentration cannot be made arbitrarily high since it can only be as high as the concentration of the injected enzyme. Second, the concentration cannot evolve arbitrarily fast. Decreasing the external concentration can be done, for example through dilution of the enzyme in the plant. Real state-of-the-art implementation of such a controlled environment can be done on a microfluidic chip (see [53]).

Sensing activity level of the colony can be done through sensing of certain protein concentrations in the cells. A certain protein called the gfp (green fluorescent protein) can be encoded in the lac operon. When the genes in the operon are expressed, gfp is also produced. Thus, the concentration of gfp in the cell can be used as an indicator for the activity of the cell. The protein gfp emits green light. Therefore, we can use the luminescence of the cells as a way to measure its level of activity. This is actually a standard procedure in synthetic biology [3], [39].

Recall the two state Markov chain model of the bacteria. We denote the probability of finding the cell at time t in the induced state as $x_{\rm hi}(t)$, and in the uninduced state as $x_{\rm lo}(t)$. The evolution of the variables $x_{\rm hi}$ and $x_{\rm lo}$ satisfies differential equation (16).

Suppose that we have N cells, and, for each cell, we introduce an output/observation map

$$y_t^i = \begin{cases} y_{\rm hi}, & \text{if the } i\text{th cell is induced at time } t \\ y_{\rm lo}, & \text{if the } i\text{th cell is uninduced at time } t \end{cases} \tag{18}$$

where y_{hi} and y_{lo} are both fixed real numbers. Obviously, $\{y_t^i\}_{i\in\{1,\ldots,N\}}$ are random processes. Furthermore, we denote the average output across the population as another random process $\bar{y}_t^N$, defined as

$$\bar{y}_t^N := \frac{1}{N}\sum_{i=1}^{N} y_t^i.$$

Notice that we explicitly write down the dependency of the average output on the size of the population.

Suppose that we are given a control problem, namely, we want to make $\bar{y}_t$ track a certain given trajectory $\eta(t)$. We propose the following solution. Consider the following model of control system:

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -\lambda_1(u) & \lambda_2(u) \\ \lambda_1(u) & -\lambda_2(u) \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$
$$y(t) = y_{lo}x_1(t) + y_{hi}x_2(t).$$

Suppose that we have an output feedback law

$$u(t) = f\left(y(t)\right) \tag{19}$$

such that the closed-loop system

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -\lambda_1\left(f(y)\right) & \lambda_2\left(f(y)\right) \\ \lambda_1\left(f(y)\right) & -\lambda_2\left(f(y)\right) \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{20}$$
$$\begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \begin{bmatrix} x_{1,0} \\ x_{2,0} \end{bmatrix}$$
$$y(t) = y_{lo}x_1(t) + y_{hi}x_2(t) \tag{21}$$

produces the output trajectory

$$y(t) = \eta(t). \tag{22}$$

We can state the following theorem about the effect of a feedback control law on the behavior of the system.

Theorem 3: If we apply the following feedback:

$$T_e(t) = f(\bar{y}_t),$$

on a colony of N cells with independent and identically distributed (i.i.d.) initial states

$$\begin{bmatrix} x_{lo}(0) \\ x_{hi}(0) \end{bmatrix} = \begin{bmatrix} x_{1,0} \\ x_{2,0} \end{bmatrix}$$

then the expected value of $\bar{y}_t$ satisfies

$$\lim_{N\to\infty} E\bar{y}_t^N = \eta(t). \tag{23}$$

Proof: Denote the state of the ith cell at time t as q_t^i, i.e.,

$$q_t^i \in \{\text{hi}, \text{lo}\}$$

which correspond to the cell being in the induced and uninduced states, respectively. The evolution of the probability densities for each cell satisfies

$$\Pr\left\{q_{t+\delta}^i = \text{hi} | q_t\right\} = \begin{cases} (1-\lambda_2(\bar{y}_t))\,\delta + O(\delta^2), & q_t^i = \text{hi} \\ \lambda_1(\bar{y}_t) + O(\delta^2), & q_t^i = \text{lo} \end{cases}$$
$$\Pr\left\{q_{t+\delta}^i = \text{lo} | q_t\right\} = \begin{cases} (1-\lambda_1(\bar{y}_t))\,\delta + O(\delta^2), & q_t^i = \text{lo} \\ \lambda_2(\bar{y}_t) + O(\delta^2), & q_t^i = \text{hi}. \end{cases}$$

Now, let n_t be the random processes that correspond to the fraction of population that is induced at time t. The output $\bar{y}_t$ can then be written as a function of n_t as

$$\bar{y}_t = y_{hi} n_t + y_{lo}(1-n_t) \tag{24}$$

and, for brevity, we shall express the dependence of λ_1 and λ_2 on n_t as $\lambda_1(n_t)$ and $\lambda_2(n_t)$, respectively. Since the transition of the cells are independent, it can be proven that

$$\Pr\left\{|n_{t+\delta} - n_t| > \frac{1}{N} | n_t\right\} = O(\delta^2).$$

Moreover, the transition densities are given by

$$\begin{aligned} \Pr\left\{n_{t+\delta} = n_t + \frac{1}{N} | n_t\right\} &= (1-n_t)N\lambda_1(n_t)\delta + O(\delta^2) \\ \Pr\{n_{t+\delta} = n_t | n_t\} &= (1-(1-n_t)N\lambda_1(n_t) \\ &\quad -n_t N\lambda_2(n_t))\,\delta + O(\delta^2) \\ \Pr\left\{n_{t+\delta} = n - \frac{1}{N} | n_t\right\} &= n_t N\lambda_2(n_t)\delta + O(\delta^2). \end{aligned}$$

Define a probability density function ρ as

$$\rho(x,t) := \Pr\{n_t = x\}.$$

Obviously, $\rho(x,t)$ can be nonzero only if $x \in \{0/N, 1/N, \ldots, N/N\}$. From the transition densities, we can compute the evolution of $\rho(x,t)$ as follows:

$$\begin{aligned} \rho(x,t+\delta) &= ((1-x)N+1)\,\lambda_1\left(x-\frac{1}{N}\right)\rho\left(x-\frac{1}{N},t\right)\delta \\ &\quad + (1-(1-x)N\lambda_1(x) - xN\lambda_2(x))\,\rho(x,t)\delta \\ &\quad + (xN+1)\lambda_2\left(x+\frac{1}{N}\right)\rho\left(x+\frac{1}{N},t\right)\delta \\ &\quad + O(\delta^2) \\ \frac{\partial\rho(x,t)}{\partial t} &= \lim_{\delta\to 0}\frac{\rho(x,t+\delta)-\rho(x,t)}{\delta} \\ &= ((1-x)N+1)\,\lambda_1\left(x-\frac{1}{N}\right)\rho\left(x-\frac{1}{N},t\right) \\ &\quad + (-(1-x)N\lambda_1(x) - xN\lambda_2(x))\,\rho(x,t) \\ &\quad + (xN+1)\lambda_2\left(x+\frac{1}{N}\right)\rho\left(x+\frac{1}{N},t\right). \end{aligned}$$

By rearranging the equation, we obtain

$$\begin{aligned} &\frac{\partial\rho(x,t)}{\partial t} \\ &= (1-x)N\left(\lambda_1\left(x-\frac{1}{N}\right)\rho\left(x-\frac{1}{N},t\right) - \lambda_1(x)\rho(x,t)\right) \\ &\quad + xN\left(\lambda_2\left(x+\frac{1}{N}\right)\rho\left(x+\frac{1}{N},t\right) - \lambda_1(x)\rho(x,t)\right) \\ &\quad + \lambda_1\left(x-\frac{1}{N}\right)\rho\left(x-\frac{1}{N},t\right) \\ &\quad + \lambda_2\left(x+\frac{1}{N}\right)\rho\left(x+\frac{1}{N},t\right). \end{aligned}$$

By taking the limit of N approaching infinity, we can interpret ρ as continuous density function, and we obtain

$$\begin{aligned} \frac{\partial\rho(x,t)}{\partial t} &= (x-1)\frac{\partial(\lambda_1\rho)}{\partial x} + x\frac{\partial(\lambda_2\rho)}{\partial x} + (\lambda_1+\lambda_2)\rho \\ &= \frac{\partial((x-1)\lambda_1\rho + x\lambda_2\rho)}{\partial x}. \end{aligned} \tag{25}$$

Seeing (25) as a diffusionless Fokker–Planck equation, we can infer that the fraction of induced cells n_t evolves according to

$$\frac{dn_t}{dt} = (1-n_t)\lambda_1(n_t) - n_t\lambda_2(n_t). \tag{26}$$

Now, we are going to bring (26) to a form similar to (20). By introducing an auxilliary variable $m_t = 1 - n_t$, which is the fraction of uninduced cells, we have

$$\begin{aligned} \frac{d}{dt}\begin{bmatrix} m_t \\ n_t \end{bmatrix} &= \begin{bmatrix} -\lambda_1(n_t) & \lambda_2(n_t) \\ \lambda_1(n_t) & -\lambda_2(n_t) \end{bmatrix}\begin{bmatrix} m_t \\ n_t \end{bmatrix} \\ \begin{bmatrix} Em_0 \\ En_0 \end{bmatrix} &= \begin{bmatrix} x_{1,0} \\ x_{2,0} \end{bmatrix}. \end{aligned}$$

Since we also have

$$\lim_{N\to\infty} E\bar{y}_t^N = y_{lo}Em_t + y_{hi}En_t$$

and, by the construction of the feedback law, we can obtain (23). ∎

Theorem 3 provides us with a guarantee that if we design a suitable output feedback law based on the model (16), implementing the feedback law on the colony of N cells will make the expected value of the average output $E\bar{y}_t^N$ as N tends to infinity.

B. Controller Design and Simulation Results

In the remainder of this paper, we shall address the following control problem. Given the control architecture in (8), we want to design a controller such that the fraction of induced cells attains a certain level, for example, 50%. In order to achieve this task, we first design and implement the controllers on the abstract model and evaluate their performance. Subsequently, we implement the controllers on the stochastic model. Theorem 3 provides us with a guarantee that, if we design a suitable output feedback law based on the model (16), implementing the feedback law on the colony of N cells will make the expected value

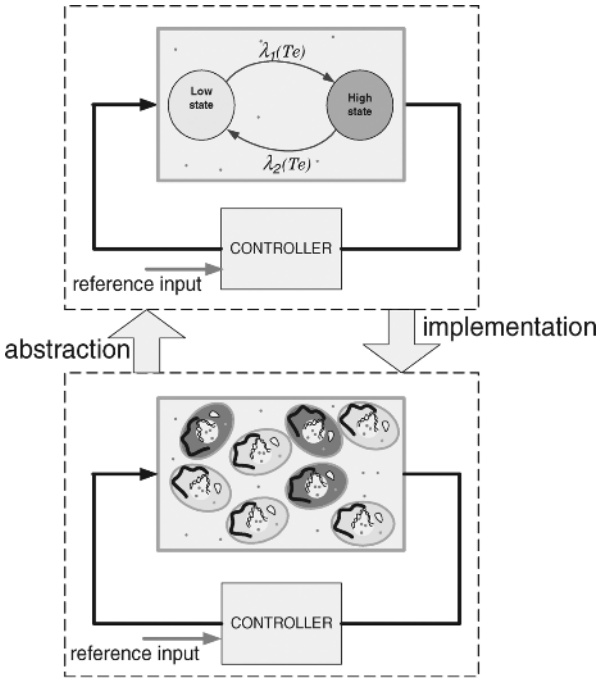


Fig. 9. Control design is performed for the abstract model. The resulting controller is then implemented for the stochastic system.

of the average output $E\bar{y}_t^N$ as the number of cells become larger. The general procedure may be summarized in Fig. 9.

Before we proceed to propose a feedback control algorithm, first notice that the state space model of the abstraction is given in (16) and the fraction of induced cells is given by $x_{hi}/(x_{hi}+x_{lo})$. In addition, the equilibria of (16) is given by

$$\begin{aligned}\frac{d}{dt}\begin{bmatrix}x_{\rm lo}\\ x_{\rm hi}\end{bmatrix} &= \begin{bmatrix}-\lambda_1(T_e) & \lambda_2(T_e)\\ \lambda_1(T_e) & -\lambda_2(T_e)\end{bmatrix}\begin{bmatrix}x_{\rm lo}\\ x_{\rm hi}\end{bmatrix}\\ &=0\\ x_{\rm hi} &= \frac{\lambda_1(T_e)}{\lambda_2(T_e)}x_{\rm lo}\end{aligned}$$

such that the fraction of induced cells at the equilibria is given by

$$y = \frac{\lambda_1(T_e)}{\lambda_1(T_e)+\lambda_2(T_e)}. \tag{27}$$

Fig. 10 captures the relation between the transition rates λ_1 and λ_2 and the external concentration of TMG. Notice that λ_1 is a monotonously increasing function of T_e, while λ_2 is monotonously decreasing. Also notice that, at $T_e = 1.4\ \mu$M, λ_2 is about nine times larger than λ_1, while at $T_e = 2\ \mu$M λ_1 is about 4.5 times larger than λ_2. Therefore, if the external concentration of TMG is kept at 1.4 μM, the fraction of induced cells is going to converge to around 10%, while, if the external concentration of TMG is kept at 2 μM, the fraction of induced cells is going to converge to around 80%.

Based on this knowledge, we can design our controller to work in the region where the external concentration of TMG is between 1.4 and 2 μM. To identify the functions $\lambda_1(T_e)$ and $\lambda_2(T_e)$ in the whole interval, we take fifth-order polynomial interpolations of the data points for both the induction and deinduction rates. The interpolated curves are shown in Fig. 11. The

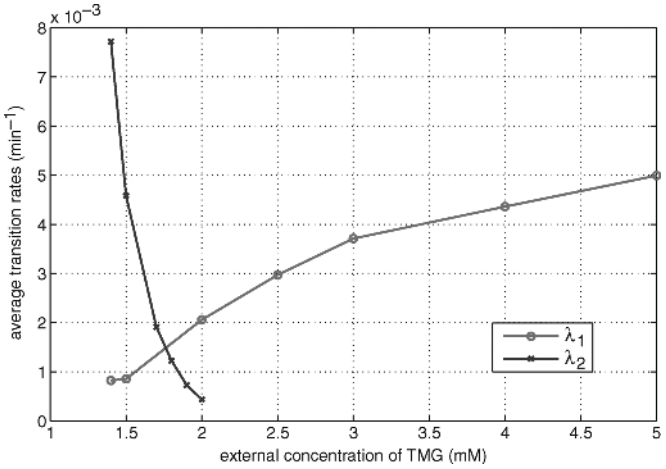


Fig. 10. Relationship between the external concentration of TMG (T) and the average transition rates (induction (λ_1) and deinduction (λ_2)). The points are data obtained from Monte Carlo simulations.

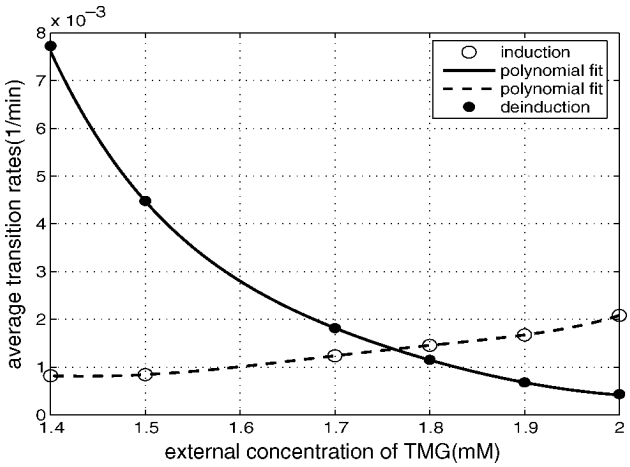


Fig. 11. Polynomial interpolation of the transition rates as functions of the external concentration of TMG in the interval between 1.4 and 2 μM.

number of data points used is a clear limitation on the order of interpolating polynomial. As a rule of thumb, more data points will result in better approximation. On the other hand, obtaining those points from Monte Carlo simulations is time consuming, which results in a tradeoff.

Now we can proceed to the design and implementation of the controllers.

On–Off Controller: The on–off controller is a basic switching controller with deadzone, where we assume that the control input (the external concentration of TMG) can assume only two values $T_{\rm lo} = 1.4\ \mu$M and $T_{\rm hi} = 2\ \mu$M. If the fraction of induced cells $\bar{y}_t$ is higher than 0.52, then $T_e = T_{lo}$. If $\bar{y}_t$ is less than 0.48, then $T_e = T_{hi}$. If $\bar{y}_t$ is between 0.48 and 0.52, then T_e is kept at its current value. The scheme of the dynamics is shown in Fig. 12. We therefore create a deadzone that will prevent the controller from switching indeterminately around the desired level of $\bar{y}_t = 0.5$.

Using the abstract model, we can predict that, when $T_e = T_{\rm lo}$, the fraction of induction is going to converge exponentially towards 10%. Similarly, when $T_e = T_{\rm hi}$, the fraction of induction is going to converge exponentially towards 80%. Obviously, the

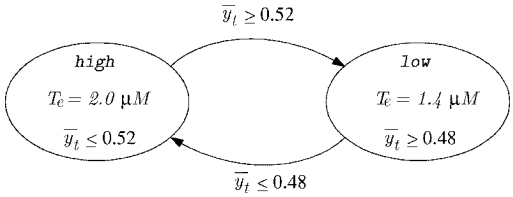


Fig. 12. Hybrid automaton representing the on–off controller.

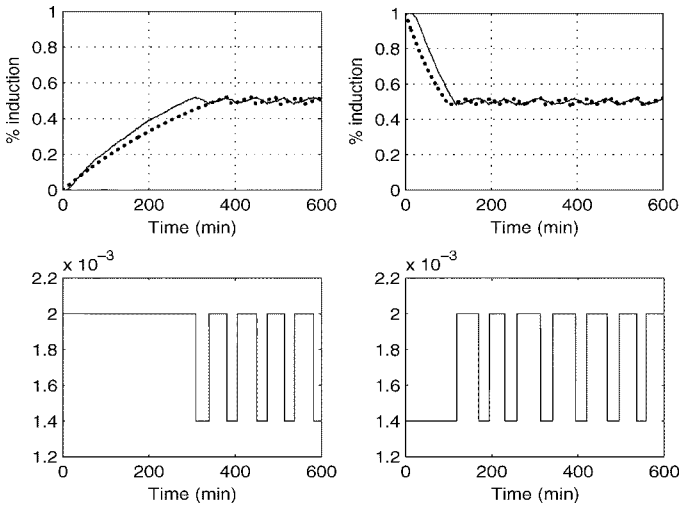


Fig. 13. Simulation results using the on–off controller. Top: fraction of induced cells. The dashed lines are the results obtained from the abstract model. The solid lines are the results from the stochastic model with 5000 cells. Bottom: the level of T_e for both simulations (in mM).

proposed on–off controller is able to steer and keep the output of the abstract model in the deadzone range, that is, 48%–52%.

Fig. 13 shows the comparative simulation results of both the abstract model and the stochastic model of the colony of bacteria with 5000 cells for two initial conditions, fully induced and fully uninduced, when the on–off controller is used. We can see that the desired fraction of activation of 50% can be attained and maintained within the deadzone. On the bottom of Fig. 13, we can see the level of T_e switches between $T_{\rm lo}$ and $T_{\rm hi}$ in both simulations.

Although the general behavior is the same, the closed-loop abstract model and closed-loop stochastic model do not match as well as their open-loop counterparts (see Fig. 6). The source of the mismatch is the error in the identified transition rates. Recall that the transition rates in the abstract model are constructed from polynomial interpolation of the data points obtained through Monte Carlo simulations of the stochastic model. Therefore, there are two sources of error: the interpolation error and the statistical deviation in the Monte Carlo simulation itself.

In Fig. 14, we can see a dynamic histogram that shows the distribution of the internal concentration of TMG in the cells, when the on–off controller is used with fully uninduced initial condition. We can see that initially (at time $t = 0$) the distribution is concentrated at the bottom level. As time progresses, a second cluster, which corresponds to the induced cells, appears. After $t = 300$ min, we can see that the higher cluster moves up and down because of the modulation of T_e, as it is also shown in Fig. 13.

Flow Controller: The on–off controller algorithm assumes that the external concentration of TMG can change between $T_{\rm lo}$ and $T_{\rm hi}$ instantaneously. This is not physically feasible if the controller is to be actually implemented. We therefore propose another controller that is more feasible, where T_e changes its value gradually.

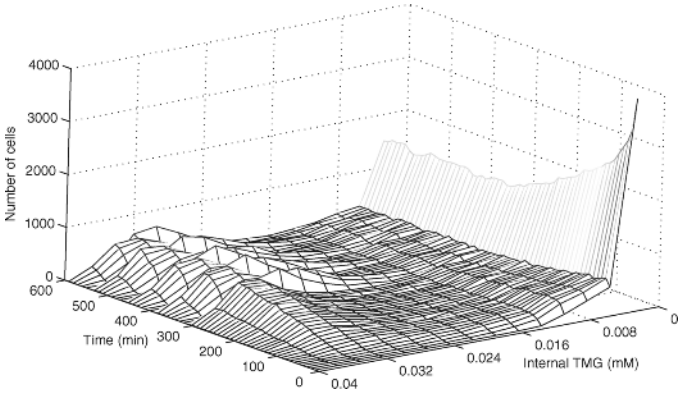


Fig. 14. Dynamic histogram of the distribution of the internal concentration of TMG in the cells when the on–off controller is used.

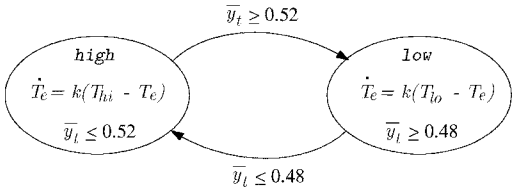


Fig. 15. Hybrid automaton representing the flow controller. The value of k is chosen such that the time constant of the exponential is 10 min ($k = 0.1$).

The controller that we propose is essentially a hybrid system with two modes of dynamics. The continuous dynamics of the first mode, `high`, is such that the concentration of T_e converges exponentially to $T_{\rm hi}$, while in the other mode, `low`, T_e converges exponentially to $T_{\rm lo}$. The scheme of the dynamics is shown in Fig. 15. If the fraction of induced cells $\bar{y}_t$ is higher than 0.52, then the controller is switched to the `low` mode. If $\bar{y}_t$ is less than 0.48, then the controller is switched to the `high` mode. If $\bar{y}_t$ is between 0.48 and 0.52, then T_e is kept at its current value. Again, here we create a deadzone that will prevent the controller from switching indeterminately around the desired level of $\bar{y}_t = 0.5$.

The flow controller basically realizes the same idea as the on–off controller. The main difference here is that the external concentration of TMG is assumed to change gradually. It converges to the desired level ($T_{\rm lo}$ or $T_{\rm hi}$) exponentially, with a time constant of 10 min. We believe that this is a realistic assumption.

We implemented the flow controller on the abstract model and the stochastic model. The simulations results are shown in Fig. 16. As expected, in the abstract model, we can see that the desired fraction of activation of 50% can be attained and maintained close to the deadzone. We can also observe that the stochastic model behaves similarly to the abstract model.

Fig. 17 shows a dynamic histogram of the internal concentration of TMG in the cells, when the flow controller is used with fully uninduced initial condition.

Both Figs. 16 and 17 show that the output variation of the flow controller is larger than that of the on–off controller. This result is expected, because the flow controller is more sluggish

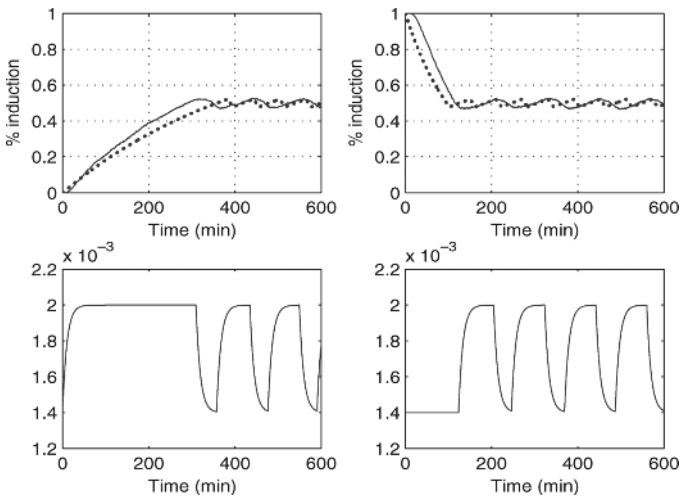


Fig. 16. Simulation results using the flow controller. Top: fraction of induced cells. The dashed lines are the results obtained from the abstract model. The solid lines are the results from the stochastic model with 5000 cells. Bottom: the level of T_e for both simulations (in mM).

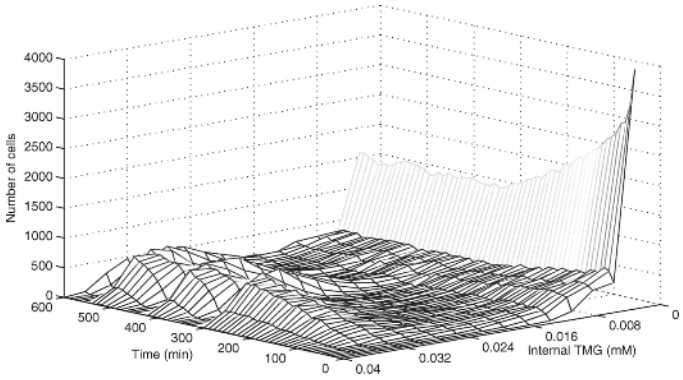


Fig. 17. Dynamic histogram of the distribution of the internal concentration of TMG in the cells when the flow controller is used.

than the on–off controller. In order to decrease the steady-state tracking error and the output variation, we propose another control algorithm.

Hybrid PI Controller: In order to improve the performance of the system in the deadzone, we propose a change in the control algorithm. In particular, we propose new controller dynamics for the deadzone range. We introduce integral action as a new controller, which is active only in the deadzone range. Therefore, we now have three modes of dynamics. If the fraction of induced cells $\overline{y}_t$ is higher than 0.52, then the controller is switched to the `low` mode. If $\overline{y}_t$ is less than 0.48, then the controller is switched to the `high` mode. However, if $\overline{y}_t$ is between 0.48 and 0.52, then the controller is switched to the `PI` mode where we activate the PI controller. The block diagram of the hybrid automaton is shown in Fig. 18.

We argue that, for the abstract model, the hybrid PI controller is robust against modeling error in terms of errors in the identification of the transition rates $\lambda_1(T_e)$ and $\lambda_2(T_e)$. In fact, we only need the fact that, in the interval $T_e = 1.4\,\mu\text{M} - 2.0\,\mu\text{M}$, $\lambda_1(T_e)$ is monotonously decreasing and $\lambda_2(T_e)$ is monotonously increasing.

The integral action basically works by increasing the control input T_e if the fraction of induction is less than the desired value

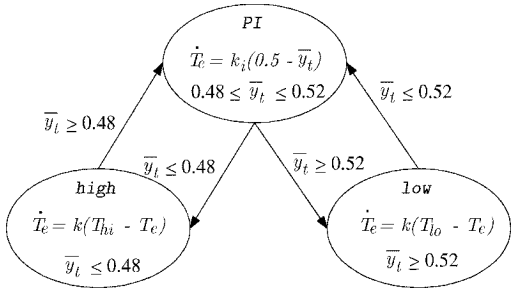


Fig. 18. Hybrid automaton representing the hybrid PI controller. The value of the integral gain $k_i = 0.01$ is chosen such that the steady-state error is small and the actuation of the controller is feasible.

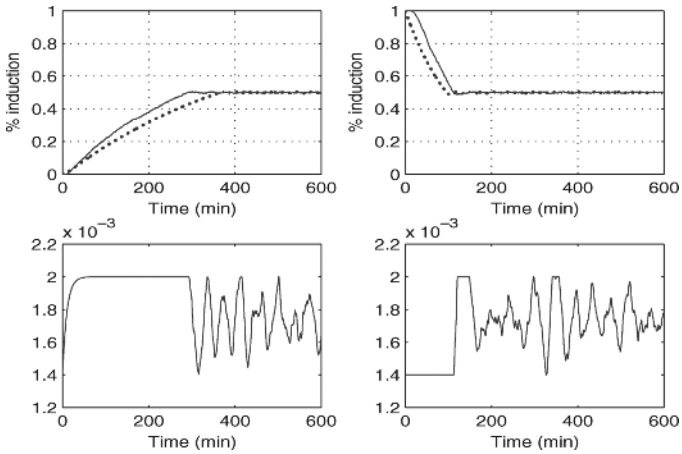


Fig. 19. Simulation results using the hybrid PI controller. Top: fraction of induced cells. The dashed lines are the results obtained from the abstract model. The solid lines are the results from the stochastic model with 5000 cells. Bottom: the level of T_e for both simulations (in mM).

and by reducing T_e if the fraction of induction is higher than the desired value. If this adjustment is done sufficiently slowly (k_i has to be small enough), we can assume that the output of the system has a much faster dynamics. Consequently, we can perform a quasi-steady-state analysis by assuming that the system is always at its equilibrium and that adjusting the control input effectively means adjusting the equilibrium. We then use the fact that the fraction of induction at the equilibrium of the abstract model is given by $\lambda_1(T_e)/(\lambda_1(T_e) + \lambda_2(T_e))$, which is a monotonously increasing function of control input T_e if $\lambda_1(T_e)$ is monotonously decreasing and $\lambda_2(T_e)$ is monotonously increasing. Therefore, the controller can drive the output to the desired value.

In the implementation on the stochastic system, k_i cannot be arbitrarily small, since the adaptation of the control input must be fast enough to react against the stochastic fluctuation of the system. We solve the problem of choosing the value of k_i through simulation using the abstract model. As the result, we pick $k_i = 0.01$.

The results of the simulations of the implementation of the hybrid PI controller are shown in Fig. 19. Compared with the flow and on–off controller, the tracking error is much smaller as expected.

In Fig. 20, we can see a dynamic histogram that shows the distribution of the internal concentration of TMG in the cells, when the hybrid PI controller is used with a fully uninduced initial condition. By looking at the histogram, we can again conclude

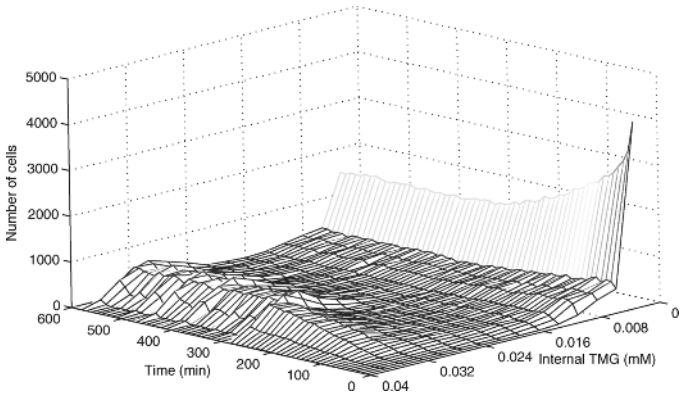


Fig. 20. Dynamic histogram of the distribution of the internal concentration of TMG in the cells when the hybrid PI controller is used.

that the hybrid PI controller is superior to the other two strategies, as the oscillations in the distribution of cells is smaller.

For all of the histograms shown above, the colony size is 5000 cells. At each time instant, the classes in the histogram are constructed using 15 equal-length intervals of internal TMG concentration between 0 and 40 μM.

As a last remark, we reiterate that actuation in the flow and hybrid PI controller is physically feasible, since we assume that the change in the external concentration of TMG is done gradually in the order of minutes. Sensing of the fraction of induction can be implemented by monitoring the luminescence of the cells, due to the production of gfp in the cells.

V. Concluding Remarks

In this paper, we present a comprehensive framework for abstraction and controller design for the lactose regulation system of the *E. coli* bacteria. The abstraction framework is based on the idea that two stable equilibria of the systems can be thought of as states of a continuous-time Markov chain and that the transition rates of the Markov chain can be obtained through Monte Carlo simulations of the actual system.

Because of the simplicity of the abstract model and its demonstrated accuracy in predicting the average behavior of a colony with a large number of cells, we can use the abstract model as a *building block* for designing, for example, a feedback control system for the biological systems. By feedback control, here we mean influencing the average behavior of the colony using an environmental variable (external concentration of TMG) as control actuation.

We presented three control algorithms: the on–off controller, the flow controller, and the hybrid PI controller. The flow controller is an improvement over the on–off controller in terms of implementability. The hybrid PI controller is an improvement over the flow controller in terms of tracking performance. In addition, we also show that the hybrid PI controller is robust with respect to plant model identification error. In the future, we plan to explore other feedback control methodologies, such as model predictive control (MPC).

In the (future) implementation, the role of the controller may be played by another genetically engineered module, such as a toggle switch [3], that results in the production or consumption of the inducer. This may be implemented in the same organism or in another strain which is present in the same bioreactor. Thus, one might be able to construct a network using specifically engineered organisms as circuit elements. A related idea, constructing a feedback motif in a synthetic genetic network, has been pursued and reported in the recent work by Guido *et al.* [54].

The control architecture that we propose in this paper, which is based on control of a large number of Markov chains by adjusting the transition rates, is a versatile framework. Another application of this control architecture in controlling artificial muscle fibers is reported in [52]. Considering the generality of the framework, we see potential application of it in other fields such as active materials and networked engineering systems with a large number of autonomous agents such as sensor networks and robotic swarms [55].

References

[1] A. A. Julius, A. Halasz, V. Kumar, and G. J. Pappas, "Finite state abstraction of a stochastic model of the lactose regulation system of *Escherichia coli*," in *Proc. IEEE Conf. Decision Control*, San Diego, CA, 2006, pp. 9–14.

[2] J. Beckwith, *The Lactose Operon*, F. C. Neidhart, Ed. Washington, DC: Amer. Soc. for Microbiology, 1987, ch. 4, pp. 1444–1452.

[3] T. S. Gardner, C. R. Cantor, and J. J. Collins, "Construction of a genetic toggle switch in *Escherichia coli*," *Nature*, vol. 403, pp. 339–342, 2000.

[4] E. Andrianantoandro, S. Basu, D. K. Karig, and R. Weiss, "Synthetic biology: New engineering rules for an emerging discipline," *Molec. Syst. Biol.*, vol. 2, pp. 2006–0028, 2006.

[5] N. Q. Balaban, J. Merrin, R. Chait, L. Kowalik, and S. Leibler, "Bacterial persistence as a phenotypic switch," *Science*, vol. 305, pp. 1622–1625, 2004.

[6] G. Huang, H. Wang, S. Chou, X. Nie, J. Chen, and H. Liu, "Bistable expression of WOR1, a master regulator of white-opaque switching Candida albicans," *Proc. Nat. Acad. Sci.*, vol. 103, no. 34, pp. 12813–12818, 2006.

[7] M. Khammash and H. El-Samad, "Systems biology: From physiology to gene regulation," *IEEE Control Syst. Mag.*, vol. 24, no. 4, pp. 62–76, Aug. 2004.

[8] E. L. Berg, E. J. Kunkel, and E. Hytopoulos, "Biological complexity and drug discovery: A practical systems biology approach," *Syst. Biol.*, vol. 152, no. 4, pp. 201–206, Dec. 2005.

[9] C. J. Tomlin and J. D. Axelrod, "Understanding biology by reverse engineering the control," in *Proc. Nat. Acad. Sci.*, Mar. 2005, vol. 102, pp. 4219–4220.

[10] H. El-Samad, H. Kurata, J. C. Doyle, C. A. Gross, and M. Khammash, "Surviving heat shock: Control strategies for robustness and performance," *Proc. Nat. Acad. Sci.*, vol. 102, no. 8, pp. 2736–2741, 2005.

[11] L. Yang and P. A. Iglesias, "Feedback induced biphasic response in the chemotaxis pathway of Dictyostelium," in *Proc. Amer. Control Conf.*, Jun. 2005, pp. 4393–4398.

[12] B. P. Kramer, C. Fischer, and M. Fussenegger, "BioLogic gates enable logical transcription control in mammalian cells," *Biotechnol. Bioeng.*, vol. 87, pp. 478–484, 2004.

[13] M. R. Atkinson, M. A. Savageau, J. T. Myers, and A. J. Ninfa, "Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in *Escherichia coli*," *Cell*, vol. 113, pp. 597–607, 2003.

[14] P. Wellstead, R. Middleton, and O. Wolkenhauer, " Working notes on control sytems concepts in personalized, predictive and combinatorial intervention," published as arXiv :q-bio.TO/0603032 vi, Mar. 2006.

[15] J. A. M. F. de Souza, M. A. L. Caetano, and T. Yoneyama, "Optimal control theory applied to the anti-viral treatment of AIDS," in *Proc. IEEE Conf. Decision Control*, 2000, pp. 4839–4844.

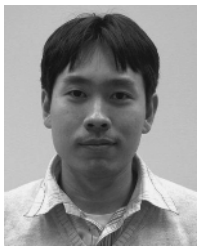
[16] D. Kirschner, S. Lenhart, and S. Serbin, "Optimal control of the chemotheraphy of HIV," *J. Math. Biol.*, vol. 35, pp. 775–792, 1997.

[17] E. Jung, S. Lenhart, and Z. Feng, "Optimal control of treatments in a two-strain tuberculosis model," *Discrete Contin. Dynam. Syst. B*, vol. 2, pp. 473–482, Nov. 2002.

[18] R. F. Stengel, R. Ghigliazza, N. Kulkarni, and O. Laplace, "Optimal control of innate immune response," *Optim. Control Appl. Meth.*, vol. 23, pp. 91–104, 2002.

[19] R. S. Parker, F. J. Doyle, III, and N. A. Peppas, "The intravenous route to blood glucose control," *IEEE Eng. Med. Biol.*, pp. 65–73, 2001.

[20] P. Erdi, T. Kiss, J. Toth, B. Ujfalussy, and L. Zalanyi, "From systems biology to dynamical neuropharmacology: Proposal for a new methodology," *Syst. Biol.*, vol. 153, no. 4, pp. 299–308, Jul..
[21] Y. Liu, H. B. Sun, and H. Yokota, "Regulating gene expression using optimal control," in *Proc. 3rd IEEE Symp. Bioinformatics and Bioengineering*, Mar. 2003, pp. 313–318.
[22] J. Monod and A. M. Pappenheimer, "The kinetics of the biosynthesis of beta-galactosidase in *Escherichia coli* as a function of growth," *Biochim. et Biophys. Acta*, vol. 9, p. 648, 1952.
[23] A. Novick and M. Weiner, "Enzyme induction as an all-or-none phenomenon," *Proc. Nat. Acad. Sci. USA*, vol. 43, pp. 553–566, 1957.
[24] J. M. Mahaffy and E. Simeonov, "Stability analysis for a mathematical model of the lac operon," *Q. Appl. Math.*, vol. 57, pp. 37–53, 1999.
[25] N. Yildirim, M. Santillan, D. Horike, and M. C. Mackey, "Dynamics and bistability in a reduced model of the lac operon," *Chaos*, vol. 14, pp. 279–292, June.
[26] N. Yildirim and M. C. Mackey, "Feedback regulation in the lactose operon: A mathematical modeling study and comparison with experimental data," *Biophys. J.*, vol. 84, pp. 2841–2851, 2003.
[27] D. T. Gillespie, "A general method for numerically simulating the stochastic time evolution of coupled chemical reactions," *J. Comput. Phys.*, vol. 22, pp. 403–434, 1976.
[28] D. T. Gillespie, "Approximate accelerated stochastic simulation of chemically reacting systems," *J. Chem. Phys.*, vol. 115, no. 4, p. 1716, 2001.
[29] D. Gillespie and L. Petzold, "Improved leap-size selection for accelerated stochastic simulation," *J. Chem. Phys.*, vol. 119, pp. 8229–8234, 2003.
[30] T. B. Kepler and T. C. Elston, "Stochasticity in transcriptional regulation: Origins, consequences, and mathematical representations," *Biophys. J.*, vol. 81, pp. 3116–3136, Dec. 2001.
[31] L. S. Tsimring, D. Volfson, and J. Hasty, "Stochastically driven genetic circuits," *Chaos*, vol. 16, no. 026103, 2006.
[32] A. Arkin, J. Ross, and H. H. McAdams, "Stochastic kinetic analysis of developmental pathway bifurcation on phage lambda-infected escherichia coli cells," *Genetics* vol. 149, pp. 1633–1648, 1998.
[33] E. Kussell and S. Leibler, "Phenotypic diversity, population growth, and information in fluctuating environments," *Science*, vol. 309, p. 2075, 2005.
[34] E. M. Ozbudak, M. Thattal, H. N. Lim, B. I. Shraiman, and A. van Oudenaarden, "Multistability in the lactose utilization network of *Escherichia coli*," *Nature*, vol. 427, pp. 737–740, Feb. 2004.
[35] J. T. Mettetal, D. Muzzey, J. M. Pedraza, E. M. Ozbudak, and A. van Oudenaarden, "Predicting stochastic gene expression dynamics in single cells," *Proc. Nat. Acad. Sci.*, vol. 103, no. 19, pp. 7304–7309, 2006.
[36] D. Longo and J. Hasty, "Dynamics of single-cell gene expression," *Molec. Syst. Biol.* vol. 2, no. 64, 2006.
[37] N. A. Neogi, "Dynamic partitioning of large discrete event biological systems for hybrid simulation and analysis," in *Hybrid Systems: Computation and Control*, ser. Lecture Note in Compuer Science. Berlin, Germany: Springer-Verlag, 2004, vol. 2993, pp. 463–476.
[38] C. G. Cassandras and S. Lafortune, *Introduction to Discrete Event Systems*. Boston, MA: Kluwer, 1999.
[39] E. M. Ozbudak, M. Thattai, H. N. Lim, B. I. Shraiman, and A. van Oudenarden, "Multistability in the lactose utilization network of *Escherichia coli*," *Nature*, vol. 427, pp. 737–740, 2004.
[40] A. Halasz, V. Kumar, M. Imielinski, C. Belta, O. Sokolsky, and S. Pathak, "Analysis of lactose metabolism in *E. coli* using reachability analysis of hybrid systems," *IET Syst. Biol.*, vol. 1, no. 2, pp. 130–148, 2007.
[41] M. Samoilov, S. Plyasunov, and A. P. Arkin, "Stochastic amplification and signaling in enzymatic futile cycles through noise-induced bistability with oscillations," *Proc. Nat. Acad. Sci.*, vol. 102, no. 7, pp. 2310–2315, 2005.
[42] M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain, "Stochastic gene expression in a single cell," *Science*, vol. 297, pp. 1183–1186, 2002.
[43] R. A. Cox, "Quantitative relationships for specific growth rates and macromolecular compositions of *Mycobacterium tuberculosis, Streptomyces coelicolor* a3(2) and *Escherichia coli* B/r: An intergrative theoretical approach," *Microbiology*, vol. 150, pp. 1413–1426, 2004.
[44] T. Tian and K. Burrage, "Stochastic models for regulatory networks of the genetic toggle switch," *Proc. Nat. Acad. Sci.*, vol. 103, no. 22, pp. 8372–8377, 2006.
[45] J. M. G. Vilar, C. C. Guet, and S. Leibler, "Modeling network dynamics: The lac operon, a case study," *J. Cell Biol.*, vol. 161, no. 3, pp. 471–476, 2003.
[46] Z. Szallasi, J. Stelling, and V. Periwal, *System Modeling in Cellular Biology: From Concepts to Nuts and Bolts*. Cambridge, MA: MIT Press, 2006.
[47] N. Yildirim and M. C. Mackey, "Feedback regulation in the lactose operon: A mathematical modeling study and comparison with experimental data," *Biophys. J.*, vol. 84, pp. 2841–2851, May 2003.
[48] Y. Cao, D. T. Gillespie, and L. Petzold, "Avoiding negative populations in explicit Poisson tau-leaping," *J. Chem. Phys.*, vol. 123, pp. 054104–054112, 2005.
[49] A. A. Julius, "Approximate abstraction of stochastic hybrid automata," in *Hybrid Systems: Computation and Control*, ser. Lecture Notes Computer Science. Berlin, Germany: Springer-Verlag, 2006, vol. 3927, pp. 318–332.
[50] W. A. Knorre, "Oscillation of the rate of synthesis of beta-galactosidase in *Escherichia coli* ML 30 and ML 308," *Biochem. Biophys. Res. Commun*, vol. 30, pp. 1248–1290, 1968.
[51] S. Pestka, B. L. Daugherty, V. Jung, K. Hotta, and R. K. Pestka, "Anti-mRNA: Specific inhibition of translation of single mRNA molecules," *Proc. Nat. Acad. Sci.*, vol. 81, pp. 7525–7528, 1984.
[52] J. Ueda, L. Odhnar, and H. Asada, "A broadcast-probability approach to the control of vast DOF cellular actuators," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2006, pp. 1456–1461.
[53] D. N. Breslauer, P. J. Lee, and L. P. Lee, "Microfluidics-based systems biology," *Molec. Biosyst.*, vol. 2, pp. 97–112, 2006.
[54] N. Guido, X. Wang, D. Adalsteinsson, D. McMillen, J. Hasty, C. Cantor, T. Elston, and J. Collins, "A bottom-up approach to gene regulation," *Nature*, vol. 439, pp. 856–860, 2006.
[55] S. Berman, A. Halasz, V. Kumar, and S. Pratt, "Algorithms for the analysis and synthesis of a bio-inspired swarm robotic system," in *Proc. Int. Conf. Simulation Adapt. Behav. (SAB '06)*, Rome, Italy, 2006, pp. 56–70.



A. Agung Julius (S'96–M'06) received the B.Eng. degree in electrical engineering from the Bandung Insitute of Technology in 1998, and the M.Sc. and Ph.D. degrees in applied mathematics from the University of Twente, Twente, The Netherlands, in 2001 and 2005, respectively.

He is currently a Postdoctoral Fellow with the Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia. His research interests include hybrid systems, reduction of complex systems, verification and robust testing of hybrid systems, stochastic models in systems biology, control of biological systems, mathematical systems theory, behavioral systems and control theory, interplay between theoretical computer science, and systems theory.



Adam M. Halasz (M'07) received the Diploma in physics from Bucharest University, Bucharest, Romania, in 1989, and the M.A. and Ph.D. degrees in theoretical nuclear physics from the State University of New York, Stony Brook, in 1995 and 1998, respectively.

He is currently a Research Scientist with the Department of Mechanical Engineering and Applied Mechanics, University of Pennsylvania, Philadelphia. His research interests include the molecular systems biology of cellular processes (e.g., bacterial stringent response, persistence, and metabolic networks), development of mathematical methods for systems biology inspired from engineering (e.g., automated analysis of biomolecular network models using hybrid systems theory and reachability analysis), as well as the design of biologically inspired, stochastically implemented control strategies for swarms of robots.



M. Selman Sakar (S'05) received the B.S. degree in electrical and electronics engineering from Bogazici University, Turkey, in 2005, and the M.S. degree in electrical engineering from the University of Pennsylvania, Philadelphia, in 2007. He is currently working toward the Ph.D. degree at the Department of Electrical and Systems Engineering, University of Pennsylvania.

His research interests include applications of control theory and dynamical systems in physiology and molecular biology, stochastic models in systems biology, and bio-inspired controllers for swarms of robots.



Harvey Rubin received the Ph.D. degree in molecular biology from the University of Pennsylvania, Philadelphia, in 1974, and the M.D. degree from Columbia University, New York, in 1976.

He joined the faculty at the University of Pennsylvania in 1983 and became a Professor of Medicine in 1998. He holds secondary appointments as a Professor with the Department of Microbiology, the Department of Biochemistry and Biophysics in the School of Medicine, and as a Professor of Computer and Information Sciences with the University of Pennsylvania School of Engineering and Applied Sciences. He is the founder and Director of the Institute for Strategic Threat Analysis and Response (ISTAR) at the University of Pennsylvania. The mission of ISTAR is global--addressing strategies and responses to intentional as well as unintentional threats. Some of the areas under investigation include: the dynamics of urban preparedness; international treaties and compacts for infectious diseases, protection of the physical and information infrastructures; risk assessment and public communication in times of crises; emerging infectious diseases and agents of bioterrorism; health care delivery and public health systems under crisis situations; curriculum development to teach the response to all hazards; and economic models of the consequences of pandemics and other disasters. His laboratory is in the Department of Infectious Diseases, School of Medicine, and his research is on the genetic and biochemical mechanisms involved in the pathophysiology of Mycobacterium tuberculosis, the bacteria that is responsible for the disease that causes more deaths worldwide than any other bacterial infection. In addition to his work on the basic biology of the disease, he has extended the investigations to mathematical modeling of complex biological systems. He takes care of patients in the outpatient clinic and with the inpatient infectious disease consultation service at the Hospital of the University of Pennsylvania.

Dr. Rubin has served on a number of national and international scientific review panels including the National Institutes of Health, the National Science Foundation, the National Aeronautics and Space Administration Intelligent Systems Program, the Defense Advanced Research Projects Agency, and The Medical Research Council, South Africa. He is a member of the United States National Science Advisory Board for Biosecurity.



Vijay Kumar (F'06) received the M.Sc. and Ph.D. degrees in mechanical engineering from The Ohio State University, Columbus, in 1985 and 1987, respectively.

He has been a member of the faculty of the Department of Mechanical Engineering and Applied Mechanics, with a secondary appointment with the Department of Computer and Information Science, University of Pennsylvania, Philadelphia, since 1987. He is currently the UPS Foundation Professor and the Chairman of Mechanical Engineering and Applied Mechanics. He has served as the Deputy Dean of the School of Engineering and Applied Science from 2000 to 2004. He directed the GRASP Laboratory, a multidisciplinary robotics and perception laboratory, from 1998–2004.

Dr. Kumar is a Fellow of the American Society of Mechanical Engineers (ASME). He has served on the editorial boards of the IEEE TRANSACTIONS ON ROBOTICS AND AUTOMATION, the *Journal of the Franklin Institute,* the *ASME Journal of Mechanical Design*, and the IEEE TRANSACTIONS ON AUTOMATION SCIENCE AND ENGINEERING. He was the recipient of the 1991 National Science Foundation Presidential Young Investigator Award, the Lindback Award for Distinguished Teaching, the 1997 Freudenstein Award for significant accomplishments in mechanisms and robotics, and the 2004 IEEE International Conference on Robotics and Automation Kawamori Best Paper Award. He is also a Distinguished Lecturer of the IEEE Robotics and Automation Society and an elected member of the Robotics and Automation Society Administrative Committee.



George J. Pappas (S'91–M'98–SM'04) received the Ph.D. degree in electrical engineering and computer sciences from the University of California, Berkeley, in 1998.

He is currently a Professor with the Department of Electrical and Systems Engineering and the Director of the GRASP Laboratory, University of Pennsylvania, Philadelphia. He also holds secondary appointments with the Department of Computer and Information Sciences and the Department of Mechanical Engineering and Applied Mechanics, University of Pennsylvania. His research focuses on the areas of hybrid systems, hierarchical control systems, distributed control systems, nonlinear control systems, and geometric control theory, with applications to robotics, unmanned aerial vehicles, and biomolecular networks. He coedited the book *Hybrid Systems: Computation and Control* (Springer-Verlag, 2004, ser. Lecture Notes in Computer Science).

Dr. Pappas was the recipient of a National Science Foundation (NSF) Career Award in 2002, as well as the 2002 NSF Presidential Early Career Award for Scientists and Engineers (PECASE). He received the 1999 Eliahu Jury Award for Excellence in Systems Research from the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley. His and his students' papers were finalists for the Best Student Paper Award at the IEEE Conference on Decision and Control (1998, 2001, 2004, and 2006), the American Control Conference (2001 and 2004), and the IEEE Conference on Robotics and Automation (2007). He currently serves as an Associate Editor for the IEEE TRANSACTIONS ON AUTOMATIC CONTROL.