Modeling Subtilin Production in *Bacillus subtilis* Using Stochastic Hybrid Systems*

Jianghai Hu, Wei-Chung Wu, and Shankar Sastry

Department of Electrical Engineering and Computer Sciences
University of California at Berkeley - Berkeley CA 94720, USA
{jianghai,wcwu,sastry}@eecs.berkeley.edu

Abstract. The genetic network regulating the biosynthesis of subtilin in *Bacillus subtilis* is modeled as a stochastic hybrid system. The continuous state of the hybrid system is the concentrations of subtilin and various regulating proteins, whose productions are controlled by switches in the genetic network that are in turn modeled as Markov chains. Some preliminary results are given by both analysis and simulations.

1 Background of Subtilin Production

In order to survive, bacteria develop a number of strategies to cope with harsh environmental conditions. One of the survival strategies employed by bacteria is the release of antibiotics to eliminate competing microbial species in the same ecosystem [15]. It is observed that the production of antibiotics in the cells is affected by not only the environmental stimuli (*e.g.* nutrient levels, aeration, *etc.*) but also the local population density of their own species [12]. Therefore, the physiological states of the cell and the external signals both contribute to the regulation of antibiotic synthesis. Our study focuses on the subtilin, an antibiotic produced by *Bacillus subtilis* ATCC 6633, because the genetics of subtilin is known and its biosynthetic pathways are well characterized [2, 7, 11].

We briefly describe the production process of subtilin in *B. subtilis*. It is shown in [19] that the production is controlled by two independent mechanisms. When the foods are abundant, the population proliferates and the cells produce very little amount (non-lethal dose) of subtilin. However, when the foods become scarce, the production of subtilin picks up as follows. First, sigma-H (SigH), a sigma factor that regulates gene expression, enables the production of SpaRK (SpaR and SpaK) proteins by binding to the promoter regions of their genes (*spaR* and *spaK*). The membrane-bound SpaK protein senses the extracellular subtilin accumulating in the environment as the cell colony becomes large, and activates the SpaR protein. The activated SpaR (SpaR$\sim p$) in turn directs the productions of the subtilin structural peptide SpaS, the biosynthesis complex SpaBTC which modifies SpaS to yield the final product subtilin, and the immunity machinery SpaIFEG which protects the cell against the killing effect of

* This research is partially supported by the National Science Foundation under Grant No. EIA-0122599.

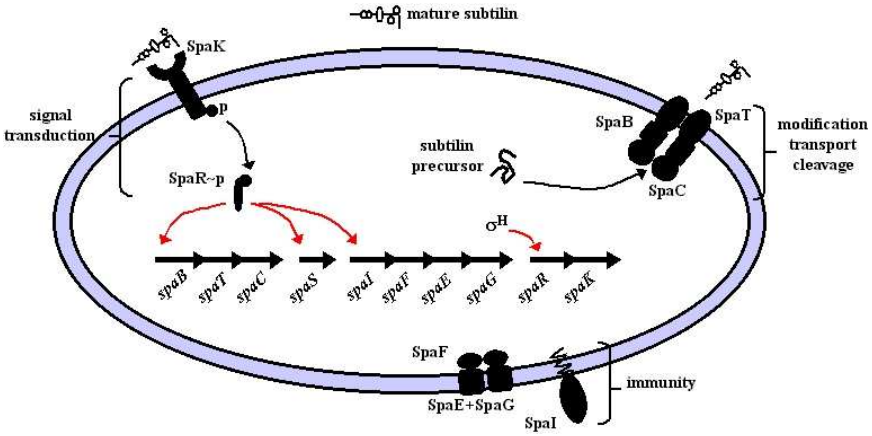

Fig. 1. Schematic representation of subtilin biosynthesis, immunity, and regulation in *Bacillus subtilis*. Subtilin prepeptide SpaS is modified, cleaved, and translocated across the cell membrane by the subtilin synthetase complex SpaBTC. The genes of subtilin are organized in an operon-like structure (*spaBTC*, *spaS*, *spaIFEG*, and *spaRK*) so that each functional unit is transcribed together. The extra-cellular subtilin functions as pheromone that regulates its own synthesis via an autoinduction feedback pathway.

subtilin. See Fig. 1 for a schematic representation of the biosynthesis process of subtilin in *B. subtilis*. In this work, a simplified version of the production process is adopted for ease of study. Namely, we ignore the dynamics in the post-translational processing of SpaS by SpaBTC to form mature subtilin and the signal transduction between SpaK and SpaR. Hence, the amount of SpaS is assumed to be equivalent to the amount of subtilin released by the cell and the SpaK and SpaR proteins are considered as one protein species.

From the above description, a dynamical model of subtilin production consists of two parts: discrete events and continuous dynamics. The discrete events include the initiations and terminations of transcriptions of various genes due to the binding and unbinding of their transcription regulators to their promoter regions, while the continuous dynamics include the accumulations and degradations of the protein species after the expressions of their genes are being switched on and off, respectively. Thus, hybrid systems can be a suitable choice for such a model.

Furthermore, in cellular networks involving coupled genetic and biochemical reactions, the expressions of genes are intrinsically non-deterministic, as is evidenced by the random fluctuations (noise) in the concentrations of protein species in the cell population, even in an isogenic culture [6]. One reason for the stochasticity in gene expressions is the small copy number of interacting molecular species (*e.g.* regulatory proteins and genes) in the relatively large cell volume [13, 14, 16], since in a chemical system with extremely low concentrations of reacting species, a reaction (collision of the reacting molecules) occurs in a

short time interval and is best viewed as a probabilistic event [9]. This is also the case in our example, as an individual *B. subtilis* cell has only one copy of each *spa* gene for the corresponding *spa* protein. Techniques such as stochastic differential equations and Monte Carlo algorithms are often used to model and simulate such biological systems with noise (see [17] for a review). Although stochastic algorithms are computationally involved, they produce a more realistic and complete description of the time-dependent behavior of biochemical systems than deterministic algorithms. In this paper, we adopt such a stochastic point of view and propose a stochastic hybrid systems approach to analyze the dynamics of the *spa* genes in the subtilin system. We also present simulation results of the subtilin regulatory network to demonstrate the distinctive behaviors of the systems using the deterministic and stochastic modeling formalisms.

2 Stochastic Hybrid Systems

A framework first proposed by the engineering community to model systems that exhibit both continuous dynamics and discrete state changes, hybrid systems have in recent years found increasingly wide applications in many practical fields. In particular, applications of hybrid systems in modeling of biological systems can be found in [8, 1], to name a few. However, most of the models proposed in the literature so far are deterministic, and are not suitable to model system with inherent randomness. This is the case, for example, in cellular processes modeling, especially when the number of participating cells is not large enough and random fluctuations within single cells cannot be ignored. Hence one needs to extend the framework to a more general class of hybrid systems with built-in randomness, namely, stochastic hybrid systems.

There have already been some work on stochastic hybrid systems (see, for example, [10, 3]). In this section we shall present a simple model suitable for our main example, the production of subtilin, to be given in the next section. Basically, the state of the stochastic hybrid system consists of two parts, a continuous one and a discrete one. The discrete state has a finite number of possible values, and evolves randomly according to a Markov chain. The continuous state, on the other hand, takes value in a certain Euclidean space, and evolves deterministically according to some ordinary differential equations. The dynamics of the discrete and the continuous states are coupled in the following sense: on one hand, the transition probabilities of the Markov chain for the discrete state depend on the value of the continuous state at the moment of jump; on the other hand, the differential equations governing the evolution of the continuous state are different when the discrete state takes on different values.

More formally, we consider the following model of a stochastic hybrid system. Its state (q, x) consists of a discrete part (mode) q taking values in a finite set Q and a continuous part x taking values in a Euclidean space $\mathbb{R}^n$, $n \geq 1$. Both q and x are functions of time t and their dynamics are specified in the following.

– (*Discrete Dynamics*) The discrete state q follows a Markov chain that jumps at epochs $t = 0, \Delta, 2\Delta, \ldots$ of constant interval $\Delta > 0$ with transition proba-

bilities

$$p_{ij}(z) \triangleq P\{q(t^+) = j | q(t^-) = i, x(t^-) = z\}, \quad \forall i, j \in Q. \tag{1}$$

Here $q(t^-)$ and $q(t^+)$ denote the values of the discrete state immediately before and after the jump at time t, respectively. Note that the transition probabilities as defined in (1) depend on the value of the continuous state at the moment of the jump.

– (*Continuous Dynamics*) For each $i \in Q$, let f_i be a vector field on $\mathbb{R}^n$. Then the dynamics of the continuous state x follows

$$\dot{x}(t) = f_i(x(t)) \tag{2}$$

in any open time interval where $q \equiv i$. Here we assume that solutions to equation (2) starting from an arbitrary initial position are well defined on all time intervals. This is automatically satisfied if each f_i is Lipschitz continuous in x.

– (*Reset Condition*) We assume that the continuous state is reset trivially at any moment t of discrete state jump, namely, $x(t^-) = x(t^+)$ for all $t = 0, \Delta, 2\Delta, \ldots$. Thus we shall simply denote them by $x(t)$.

It is obvious from the definition that solutions (or *executions*) $(q(t), x(t))$ of the stochastic hybrid system exist and are stochastic processes with continuous $x(t)$ and piecewise constant $q(t)$.

Remark 1. Instead of a discrete Markov chain jumping at fixed time intervals, we can alternatively model the discrete dynamics as a continuous Markov chain with generator $[r_{ij}(z)]_{i,j\in Q}$ satisfying $r_{ij}(z) \geq 0$ for $i \neq j$ and $\sum_{j\in Q} r_{ij}(z) = 0$, $\forall i \in Q$. Therefore, given that $x(t) = z$ and $q(t^+) = i$ right after the time t of jump, the discrete state will remain in i for a random time of exponential distribution with parameter $-r_{ii}(z)$, and then jump to a new state $j \neq i$ with probability $r_{ij}(z)/\sum_{j\neq i} r_{ij}(z)$.

3 Model of Subtilin Production

3.1 Population Growth Model

Denote by D the size of *B. subtilis* population, and by X the total amount of nutrient available in the environment, both properly normalized. The growth of *B. subtilis* population can be modeled by the logistic equation:

$$\frac{dD}{dt} = rD\left(1 - \frac{D}{D_\infty}\right). \tag{3}$$

Here D_∞ represents the equilibrium population size that can be sustained in the long term given the amount X of available nutrient. Solutions to equation (3) with a fixed D_∞ are

$$D(t) = \frac{D_\infty}{1 + \left(\frac{1}{D(0)} - 1\right)e^{-rt}}, \tag{4}$$

which, starting from any initial value $D(0) > 0$, converge to D_∞ as $t \to \infty$. In our model D_∞ changes with X, the amount of available nutrient, in the following way:

$$D_\infty = \min\{X/X_0, D_{max}\} \tag{5}$$

for some constant X_0. Thus D_∞ increases linearly with the increase of X before it saturates at a fixed level D_{max} due to space limit and competition within the population. With a time-varying X, so is D_∞, and solutions to equation (3) do not have a simple expression as in (4).

The dynamics of X is given by

$$\frac{dX}{dt} = -k_1 D + k_2 \overline{[\text{SpaS}]}, \tag{6}$$

for some constants k_1 and k_2. By equation (6), there are two factors affecting the dynamics of X: the nutrient is consumed at a rate proportional to the population size (the first term); and it is replenished at a rate proportional to the average concentration of SpaS protein in the population due to the elimination of competitors caused by SpaS in the environment. The bar over [SpaS] indicates that it is a population level average, not for a particular cell.

3.2 Single Cell Model

Within each *B. subtilis* cell, there are several modulating proteins affecting the concentration level of SpaS. In this section we shall present models of their dynamics.

First of all, SigH is a sigma factor whose production is switched on if and only if the food level X is below a certain threshold ηD_{max} for some $\eta > 0$. Its production rate is assumed to be $k_3 > 0$ when it is being produced and negligible when it is not. Thus the dynamics of its concentration [SigH] can be modeled as

$$\frac{d[\text{SigH}]}{dt} = \begin{cases} -\lambda_1[\text{SigH}] & X \geq \eta D_{max}, \\ k_3 - \lambda_1[\text{SigH}] & X < \eta D_{max}, \end{cases} \tag{7}$$

where λ_1 is the natural decaying rate of SigH.

The production of the protein SpaRK is controlled by a switch S_1 in the following way. The switch has two states 1 (on) and 0 (off), corresponding to the cases where SigH is bound and unbound to the promoter region of the gene *spaRK*, respectively. SpaRK is produced at a constant rate k_4 when the switch is on, and at a negligible rate when the switch is off. Taking into consideration its natural decaying rate of λ_2, the dynamics of [SpaRK] is given by

$$\frac{d[\text{SpaRK}]}{dt} = \begin{cases} -\lambda_2[\text{SpaRK}] & \text{if } S_1 \text{ is off}, \\ k_4 - \lambda_2[\text{SpaRK}] & \text{if } S_1 \text{ is on}. \end{cases} \tag{8}$$

The switch S_1 is modeled as evolving randomly at constant time interval $\Delta > 0$ according to a Markov chain with a probability transition matrix dependent

on the concentration level of SigH as follows:

$$A([\text{SigH}]) = \begin{pmatrix} 1 - a_0([\text{SigH}]) & a_0([\text{SigH}]) \\ a_1([\text{SigH}]) & 1 - a_1([\text{SigH}]) \end{pmatrix}. \tag{9}$$

Here $a_0([\text{SigH}])$ and $a_1([\text{SigH}])$ denote the probabilities that S_1 switches from off to on and from on to off, respectively. In practice one may not know exactly the statistics of the random time interval between successive switchings of S_1. However, by choosing sufficiently small Δ and proper $A([\text{SigH}])$, one can approximate the random time interval probabilistically, as long as it has an exponential distribution.

The choice of the transition matrix in (9) should, nevertheless, satisfy the following inherent biochemical constraint. Denote by p_{rk} the probability of S_1 switching on in the equilibrium distribution. Then one should have [18]

$$p_{rk} = \frac{e^{-\Delta G_{rk}/RT}[\text{SigH}]}{1 + e^{-\Delta G_{rk}/RT}[\text{SigH}]}, \tag{10}$$

where ΔG_{rk} is the Gibbs free energy of the molecular configuration when the switch is on (the Gibbs free energy is 0 if the switch is off), T is the temperature in Kelvin (K), and $R = 1.99$ cal/mol/K is the gas constant. A simple calculation shows that for transition matrix (9), the equilibrium probability of S_1 switching on is

$$\frac{a_0([\text{SigH}])}{a_0([\text{SigH}]) + a_1([\text{SigH}])}.$$

Thus we must have

$$\frac{a_0([\text{SigH}])}{a_0([\text{SigH}]) + a_1([\text{SigH}])} = \frac{e^{-\Delta G_{rk}/RT}[\text{SigH}]}{1 + e^{-\Delta G_{rk}/RT}[\text{SigH}]}. \tag{11}$$

The set of $a_0([\text{SigH}])$ and $a_1([\text{SigH}])$ satisfying condition (11) is characterized exactly by the following family:

$$a_0([\text{SigH}]) = \mu, \quad a_1([\text{SigH}]) = \mu e^{\Delta G_{rk}/RT}[\text{SigH}]^{-1}, \tag{12}$$

for all $0 < \mu < \min\{1, e^{-\Delta G_{rk}/RT}[\text{SigH}]\}$. The difference between these choices is that they result in different frequencies of actual switchings for the Markov chain S_1. Indeed, starting from the on (respectively, off) state, it takes an expected time of $\Delta\mu^{-1}e^{-\Delta G_{rk}/RT}[\text{SigH}]$ (respectively, $\Delta\mu^{-1}$) for S_1 to switch to the other state, provided that [SigH] is kept constant. Thus μ (together with Δ) is a parameter controlling the switching frequency of S_1.

A particular choice in (12) is

$$\begin{aligned} a_0([\text{SigH}]) &= \frac{e^{-\Delta G_{rk}/RT}[\text{SigH}]}{1 + e^{-\Delta G_{rk}/RT}[\text{SigH}]}, \\ a_1([\text{SigH}]) &= \frac{1}{1 + e^{-\Delta G_{rk}/RT}[\text{SigH}]}. \end{aligned} \tag{13}$$

This results in a Markov chain (9) that can be verified to be *reversible* [5]. Reversible Markov chains (or random walks on graphs) find many applications in diverse fields of physics, such as the Ising model [4]. Thus they may be particularly relevant in modeling S_1 in the thermal equilibrium state.

The dynamics of the concentration of the third protein, SpaS, is similar to that of [SpaRK]. There is a switch S_2 with two states 1 (on) and 0 (off), corresponding to the cases where activated SpaR is bound and unbound to the promoter region of the gene *spaS*, respectively. SpaS is being produced at a constant rate k_5 whenever S_2 is on, and at a negligible rate whenever S_2 is off. In other words,

$$\frac{d[\text{SpaS}]}{dt} = \begin{cases} -\lambda_3[\text{SpaS}] & \text{if } S_2 \text{ is off,} \\ k_5 - \lambda_3[\text{SpaS}] & \text{if } S_2 \text{ is on,} \end{cases} \tag{14}$$

where λ_3 is the natural decaying rate of SpaS. Moreover, S_2 switches randomly at constant time interval Δ according to a Markov chain with a probability transition matrix dependent on [SpaRK]:

$$B([\text{SpaRK}]) = \begin{pmatrix} 1 - b_0([\text{SpaRK}]) & b_0([\text{SpaRK}]) \\ b_1([\text{SpaRK}]) & 1 - b_1([\text{SpaRK}]) \end{pmatrix}. \tag{15}$$

It is required that the equilibrium probability of S_2 switching on be ([18])

$$p_s = \frac{e^{-\Delta G_s/RT}[\text{SpaRK}]}{1 + e^{-\Delta G_s/RT}[\text{SpaRK}]}, \tag{16}$$

where ΔG_s is the Gibbs free energy of the molecular configuration when the switch S_2 is on.

A family of $b_0([\text{SpaRK}])$ and $b_1([\text{SpaRK}])$ satisfying constraint (16) is

$$b_0([\text{SpaRK}]) = \nu, \qquad b_1([\text{SpaRK}]) = \nu e^{\Delta G_s/RT}[\text{SpaRK}]^{-1},$$

where $0 < \nu < \min\{1, e^{-\Delta G_s/RT}[\text{SpaRK}]\}$ is a parameter controlling the actual switching frequency of S_2. In particular, if we choose

$$\begin{aligned} b_0([\text{SpaRK}]) &= \frac{e^{-\Delta G_s/RT}[\text{SpaRK}]}{1 + e^{-\Delta G_{rk}/RT}[\text{SpaRK}]}, \\ b_1([\text{SpaRK}]) &= \frac{1}{1 + e^{-\Delta G_s/RT}[\text{SpaRK}]}, \end{aligned} \tag{17}$$

then the corresponding Markov chain S_2 is reversible.

3.3 Stochastic Hybrid Systems Model

To sum up, each individual *B. subtilis* cell can be modeled as a stochastic hybrid system. Its discrete state is $(S_1, S_2) \in \{0, 1\} \times \{0, 1\}$, where $(S_1, S_2) = (0, 1)$ corresponds to the case when the switch S_1 is off and the switch S_2 is

on, *etc.* So there are four possible discrete states in total. Its continuous state is $([\text{SigH}], [\text{SpaRK}], [\text{SpaS}]) \in \mathbb{R}^3$ with dynamics given by equations (7), (8), and (14). The discrete state changes mode randomly every Δ time according to a Markov chain whose probability transition matrix can be obtained from $A([\text{SigH}])$ in (9) and $B([\text{SpaRK}])$ in (15) (in fact, it is the tensor product $A([\text{SigH}]) \otimes B([\text{SpaRK}])$), and depends on the values of [SigH] and [SpaRK] at the moment of transition. The food level X appearing in the continuous dynamics in (7) is a population level quantity and can be thought of as an external input to the stochastic hybrid system.

One way to study a population of *B. subtilis* is to model it as a collection of such stochastic hybrid systems evolving independently from one another. However, this is difficult to implement in practice, since the population size D is changing according to (3). Thus the number of individual stochastic hybrid systems is time varying. To overcome this difficulty, one can replace the differential equation (3) by a birth-and-death Markov chain D on $\mathbb{N}$ with properly chosen transition probabilities, with D being the (integer) number of cells in the population. One keeps track of the state of each cell, from its birth to its death, and $\overline{[\text{SpaS}]}$ is the average of [SpaS] of the currently living cells. This approach will be pursued in future work. In this paper we adopt a simplified version by assuming that $\overline{[\text{SpaS}]} = \xi[\text{SpaS}]$ for some constant ξ. Although simulation results under this assumption tend to exaggerate the random fluctuations of $\overline{[\text{SpaS}]}$, insights can still be gained on how individual *B. subtilis* cell modulates its production of subtilin at different growth stages of the population.

Intuitively, the *B. subtilis* population as a whole reacts to the food level signal X in a way similar to a feedback control system. For example, when X drops below the threshold ηD_{max}, SigH starts to be produced according to (7). An increased [SigH] then makes it more likely for S_1 to switch on, thus resulting in an increased [SpaRK]. In turn, [SpaS] will increase due to a higher probability of S_2 switching on. As this occurs for the cells in the population, more foods will be made available by equation (6), offsetting the decrease in X. Exactly the opposite happens when X is above the threshold ηD_{max}. So it is reasonable to expect that some equilibrium state will be reached eventually for the overall system, as will be illustrated in the next two sections by analysis and simulations, respectively.

4 Analysis

In this section, we focus on a single cell, and study how the concentrations of its various proteins evolve over time, i.e., the continuous dynamics of the stochastic hybrid system modeling the cell as described in Section 3.

We first observe that the dynamics of [SpaS] is affected indirectly through S_2 by [SpaRK], and in turn the dynamics of [SpaRK] is affected indirectly through S_1 by [SigH]. Moreover, [SigH] is relatively slow-varying compared with [SpaRK] and [SpaS]. Therefore, we shall focus on two sub-problems: the evolution of [SpaRK] under a fixed [SigH] and the evolution of [SpaS] under a fixed [SpaRK].

Suppose that [SigH] is fixed. Then it is easy to see that

Proposition 1. *([SpaRK], S_1) at the time $0, \Delta, 2\Delta, \ldots$ is a Markov process.*

Indeed, S_1 follows a Markov chain with probability transition matrix (9), and according to (8), [SpaRK] transits as follows:

$$[\text{SpaRK}]_{(n+1)\Delta} = \begin{cases} e^{-\lambda_2 \Delta}[\text{SpaRK}]_{n\Delta} & \text{if } S_1 \text{ is off,} \\ \frac{k_4}{\lambda_2} + e^{-\lambda_2 \Delta}\{[\text{SpaRK}]_{n\Delta} - \frac{k_4}{\lambda_2}\} & \text{if } S_1 \text{ is on,} \end{cases} \tag{18}$$

for all $n = 0, 1, \ldots$. Here $[\text{SpaRK}]_{n\Delta}$ denotes the value of [SpaRK] at the time epoch $n\Delta$. Note that $([\text{SpaRK}], S_1)_{n\Delta}$ is not a Markov chain in the conventional sense since [SpaRK] takes values in the uncountable set $\mathbb{R}$.

Note that S_1 evolves independently from [SpaRK], and its state distribution at time $n\Delta$ will converge to the stationary distribution as $n \to \infty$, namely, S_1 will open with probability p_{rk} given in (10), and close with probability $1 - p_{rk}$.

Proposition 2. *Suppose that S_1 has reached its stationary distribution. Then $[\text{SpaRK}]_{n\Delta}$, $n = 0, 1, \ldots$, is a Markov process with transition probabilities*

$$\begin{aligned} &P\{[\text{SpaRK}]_{(n+1)\Delta} = y | [\text{SpaRK}]_{n\Delta} = x\} \\ &= \begin{cases} 1 - p_{rk} & \text{if } y = e^{-\lambda_2 \Delta} x, \\ p_{rk} & \text{if } y = \frac{k_4}{\lambda_2} + e^{-\lambda_2 \Delta}\{x - \frac{k_4}{\lambda_2}\}. \end{cases} \end{aligned}$$

In other words, $[\text{SpaRK}]_{n\Delta}$ is a random walk on $\mathbb{R}$ that in one time step either jumps towards 0 or towards k_4/λ_2 by a fixed proportion $e^{-\lambda_2 \Delta}$ with constant probabilities. Thus one can expect that it will eventually achieve an equilibrium distribution on the interval $[0, k_4/\lambda_2]$. Depending on whether $p_{rk} < 0.5$ or $p_{rk} > 0.5$, the equilibrium distribution will concentrate more on the left (or right) half of the interval. Moreover, the smaller $e^{-\lambda_2 \Delta}$ is, the further $[\text{SpaRK}]_{n\Delta}$ will jump towards either 0 or k_4/λ_2, hence the larger the variance of $[\text{SpaRK}]_{n\Delta}$ (normalized by k_4/λ_2) in the equilibrium distribution.

Remark 2. The exact expression of the equilibrium distribution of $[\text{SpaRK}]_{n\Delta}$ can be complicated. For example, assuming $k_4/\lambda_2 = 1$, if $e^{-\lambda_2 \Delta}$ is rational and [SpaRK] starts from a rational number, then $[\text{SpaRK}]_{n\Delta}$ can only jump to rational numbers. In this case, $[\text{SpaRK}]_{n\Delta}$ is a random walk on $\mathbb{Q}$ instead of on $\mathbb{R}$. In particular, let us take the example of $e^{-\lambda_2 \Delta} = 1/m$ for some integer $m \geq 2$. If $[\text{SpaRK}]_{n\Delta}$ has an m-nary expression $0.\gamma_1\gamma_2\ldots$ where $\gamma_1, \gamma_2, \ldots \in \{0, \ldots, m-1\}$, then by Proposition 3, $[\text{SpaRK}]_{(n+1)\Delta}$ will have an m-nary expression $0.\gamma_0\gamma_1\gamma_2\ldots$, where $\gamma_0 = m - 1$ with probability p_{rk} and $\gamma_0 = 0$ with probability $1 - p_{rk}$. From this it is easy to see that the equilibrium distribution of $[\text{SpaRK}]_{n\Delta}$ as $n \to \infty$ is characterized by $0.\gamma_1\gamma_2\ldots$ where $\gamma_1, \gamma_2, \ldots$ are a sequence of i.i.d. random variables that take the value $m - 1$ with probability p_{rk} and the value 0 with probability $1 - p_{rk}$. Note that in this characterization $0.\gamma_1\gamma_2\ldots$ must represent a rational number, i.e., we want the resulting distribution to be restricted on $\mathbb{Q}$ since we have assumed that [SpaRK] starts from a rational number. Unless $m = 2$, this equilibrium distribution is concentrated only on a subset of $\mathbb{Q}$.

Suppose now that [SpaRK] is fixed. Then the dynamics of [SpaS] as given by (14) can be analyzed in a similar way to obtain

Proposition 3. *([SpaS], S_2) at the time $0, \Delta, 2\Delta, \ldots$ is a Markov process.*

Proposition 4. *Suppose that S_2 has reached its stationary distribution. Then $[\text{SpaS}]_{n\Delta}$, $n = 0, 1, \ldots$, is a Markov process with transition probabilities*

$$P\{[\text{SpaS}]_{(n+1)\Delta} = y | [\text{SpaS}]_{n\Delta} = x\} = \begin{cases} 1 - p_s & \text{if } y = e^{-\lambda_3 \Delta} x, \\ p_s & \text{if } y = \frac{k_5}{\lambda_3} + e^{-\lambda_3 \Delta}\{x - \frac{k_5}{\lambda_3}\}. \end{cases}$$

The equilibrium distribution of $[\text{SpaS}]_{n\Delta}$ concentrates on the interval $[0, k_5/\lambda_3]$, and has more weight on the left half interval or the right half interval depending on whether $p_s < 0.5$ or $p_s > 0.5$. In addition, the smaller $e^{-\lambda_3 \Delta}$ is, the larger the variance of $[\text{SpaS}]_{n\Delta}$ (normalized by k_5/λ_3) in the equilibrium distribution.

Instead of the above Markov process analysis, one can adopt the following deterministic approximation procedure. Suppose that [SigH] is kept constant, and that the Markov chain S_1 is in its equilibrium distribution. Then by averaging the dynamics of [SpaRK] in (8) in the two cases using the equilibrium probabilities of S_1, one obtains the "averaged dynamics" of [SpaRK]:

$$\frac{d}{dt}[\text{SpaRK}] = \frac{\alpha_1 [\text{SigH}]}{1 + \alpha_2 [\text{SigH}]} - \lambda_2 [\text{SpaRK}], \tag{19}$$

where $\alpha_1 = k_4 e^{-\Delta G_{rk}/RT}$ and $\alpha_2 = e^{-\Delta G_{rk}/RT}$. This practice of obtaining deterministic dynamics through averaging the random dynamics is often used in the biological literature. Similarly, the averaged dynamics of SpaS can be obtained from (14) as

$$\frac{d}{dt}[\text{SpaS}] = \frac{\beta_1 [\text{SpaRK}]}{1 + \beta_2 [\text{SpaRK}]} - \lambda_3 [\text{SpaS}], \tag{20}$$

where $\beta_1 = k_5 e^{-\Delta G_s/RT}$ and $\beta_2 = e^{-\Delta G_s/RT}$.

Equations (19) and (20), together with equations (3), (6) and (7), form a deterministic differential system whose solutions approximate the solutions to the original stochastic hybrid system in the average sense. Simulation results of both approaches will be compared in the next section.

5 Simulation Results

In this section, we present the simulation results for the stochastic hybrid system model introduced in Section 3, which consists of equations (3), (6), (7), (8), and (14). We also compare the results with those obtained using the deterministic model consisting of equations (3), (6), (7), (19), and (20). Unless otherwise stated, the parameters are chosen as follows: $r = 0.02$, $D_{max} = 1$, $k_1 = 0.1$,

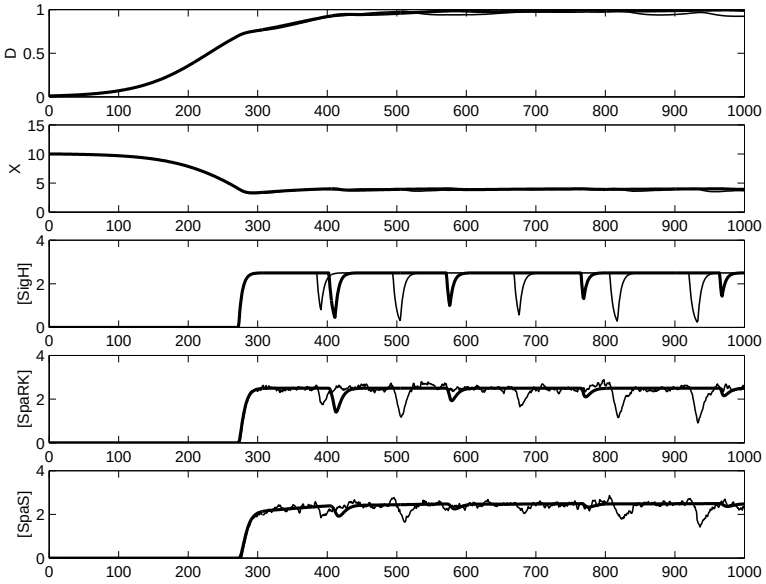

Fig. 2. A typical solution of the stochastic hybrid system model.

$k_2 = 0.4$, $k_3 = 0.5$, $k_4 = 1$, $k_5 = 1$, $\xi = 0.1$, $\lambda_1 = \lambda_2 = \lambda_3 = 0.2$, $\eta = 4$, $X_0 = 4$, $e^{-\Delta G_{rk}/RT} = 0.4$, and $e^{-\Delta G_s/RT} = 0.4$. The initial conditions are $D(0) = 0.01$, $X(0) = 10$, and $[\text{SigH}]_0 = [\text{SpaRK}]_0 = [\text{SpaS}]_0 = 0$.

Fig. 2 plots one typical realization of the continuous state trajectories of the stochastic hybrid system in thin solid lines. Simulation results of the deterministic model are plotted in the same figure in thick solid lines. As expected, the continuous state tends to some equilibrium position eventually. In particular, the population density D grows from 0.01 to some value slightly below the maximal density D_{max}, while the food level X drops from 10 to around $\eta D_{max} = 4$. The continuous states that exhibit the most random fluctuations are [SpaRK] and [SpaS], both of which fluctuate around the values predicted by the deterministic model. On the other hand, there is a visible discrepancy between the values of [SigH] in the two models.

In order to examine the population density-dependent behavior of subtilin production in the cell, we simulate our models by varying the maximum density D_{max} (carrying capacity of the incubating medium) at which a cell culture can grow to. Fig. 3 shows the results obtained by using the stochastic model (left) and the deterministic model (right), respectively. In both plots the vertical axis represents D_{max}, the horizontal axis represents the time, and the gray level at each point represents the concentration level of SpaS. Thus, each horizontal slice is one simulation run of [SpaS] with a particular D_{max}. The contours of [SpaS] are also plotted in both cases. It can be seen that a cell exhibits greater variations in the strength of subtilin production at lower cell densities in the stochastic case

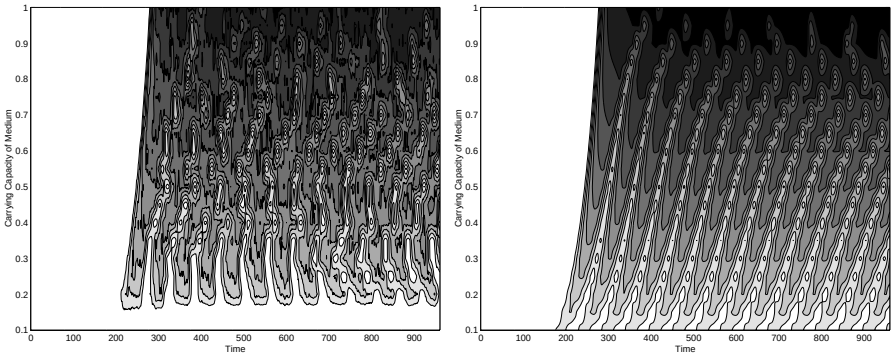

Fig. 3. Simulation results of the stochastic (left) and deterministic (right) models under different carrying capacity of medium D_{max}.

than in the deterministic case. From a biological perspective, these variations may be attributed to uncertainty in the signal transduction pathway (between the SpaK and SpaR proteins) under the condition of low level of external stimuli (*i.e.* extracellular subtilin) and are captured by the stochastic formalism of our model. In addition, as shown in Fig. 4 where the average subtilin production strength at equilibrium condition is plotted at different D_{max}, results of the stochastic model reveals that there is a threshold for D_{max} below which the cell can not be induced to produce subtilin (the concentration of subtilin in the medium is too dilute due to the sparse population to induce subtilin production in individual cells) and that the production level exhibits a linear response to the extracellular subtilin concentration if the maximum population density is above this value, as is observed experimentally in [12]. The deterministic model of subtilin production, in contrast, shows a linear response over the entire range of the carrying capacity of the medium.

We also examine the level of subtilin production in the cell in response to intra-cellular signals (*i.e.* SigH). We obtain different average concentrations of SigH at equilibrium condition by varying the parameter k_3 (larger k_3 implies larger average [SigH] at steady state) and simulate the corresponding subtilin production process. The results are shown in Fig. 5 for the stochastic model (left) and the deterministic model (right). Similar to the previous case, the "noise" in the SpaS protein synthesis due to low level of the activating signals inside the cell can be better reflected by using the Markov chain formalism. In addition, Fig. 6 suggests that the SpaS protein synthesis seems to exhibit a switching behavior in response to the increasing concentration of the activating SigH in the cell. Again, such a switching behavior is more obvious in the stochastic case than in the deterministic case. This finding, along with the simulation of varying D_{max}, supports the hypothesis that a threshold mechanism exists for the subtilin production in *B. subtilis*.

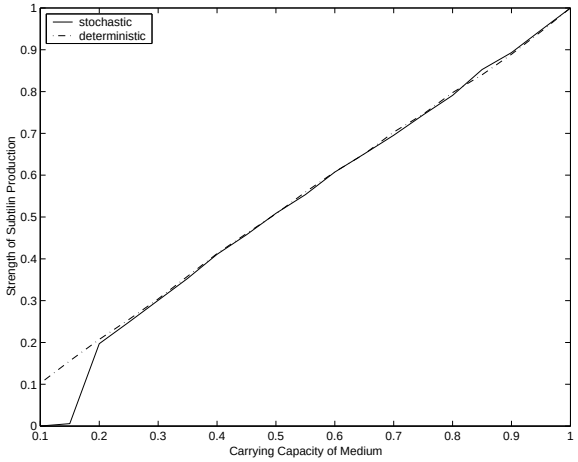

Fig. 4. Average subtilin production strength vs. carrying capacity of medium.

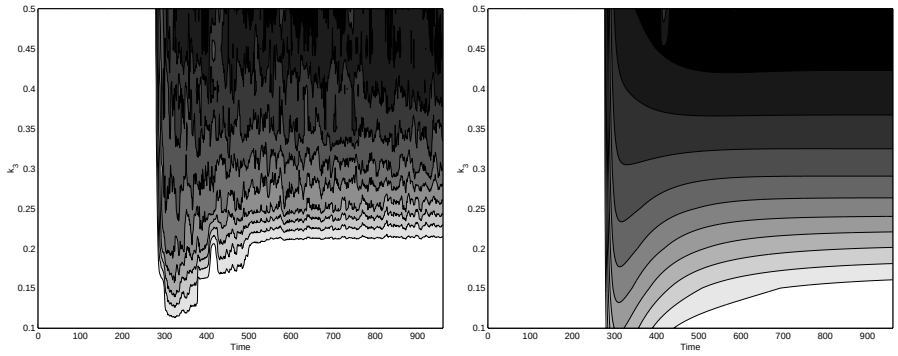

Fig. 5. Simulation results of the stochastic (left) and deterministic (right) models under different values of k_3.

6 Conclusion and Future Directions

In this paper a stochastic hybrid system model and a deterministic model of the subtilin synthesis process for *B. subtilis* cells are constructed. Although simulations of the two models generate results similar in the average sense, the stochastic model may be better suited to explain the intrinsic random fluctuations in gene expressions under the condition of a small number of participating players (*i.e.* low concentrations of regulatory proteins intra-cellularly and a small cell population extracellularly). The stochastic model also demonstrates the cell density-dependent, sigmoid-like switching behavior of subtilin production in *B. subtilis*.

This work can be extended in several directions. For example, more regulating proteins such as SpaB, SpaT, and SpaC could be incorporated into the models;

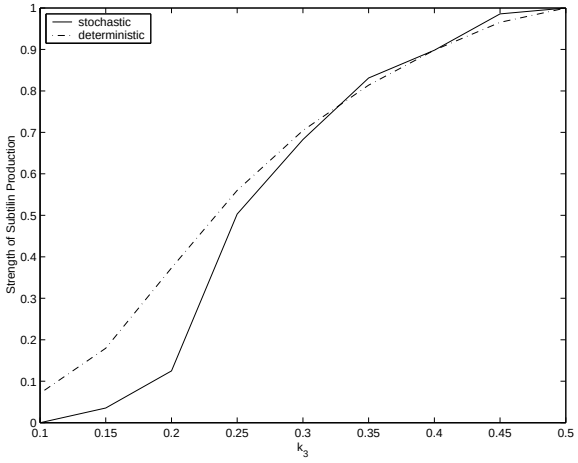

Fig. 6. Average subtilin production strength vs. k_3.

the population size D could be modeled as a birth-and-death Markov chain instead of as the solution to the logistic equation, as is pointed out in Section 3.3; and experiments on *B. subtilis* could be performed and the data collected to validate the proposed models and the simulation results.

Acknowledgment: The authors would like to thank D. M. Wolf for helpful discussions and pointers to the relevant literatures.

References

1. K. Amonlirdviman, R. Ghosh, J. D. Axelrod, and C. Tomlin. A hybrid systems approach to modeling and analyzing planar cell polarity. In *Proc. 3rd Int. Conf. on Systems Biology*, Stockholm, Sweden, December 2002.
2. S. Banerjee and J. N. Hansen. Structure and expression of a gene encoding the precursor of subtilin, a small protein antibiotic. *J. Biol. Chem.*, 263(19):9508–9514, 1988.
3. M. L. Bujorianu and J. Lygeros. Reachability questions in piecewise deterministic markov processes. In O. Maler and A. Pnueli, editors, *Hybrid Systems: Computation and Control, 6th International Workshop, Prague, Czech Republic*, volume 2623 of *Lecture Notes in Computer Science*, pages 126–140. Springer-Verlag, 2003.
4. B. A. Cipra. An introduction to the Ising model. *Amer. Math. Monthly*, 94:937–959, 1987.
5. R. Durrett. *Probability: Theory and Examples*. Duxbury Press, second edition, 1996.
6. M. B. Elowitz, A. J. Levine, E. D. Siggia, and P. S. Swain. Stochastic gene expression in a single cell. *Science*, 297:1183–1186, 2002.
7. K. D. Entian and W. M. de Vos. Genetics of subtilin and nisin biosyntheses: biosynthesis of lantibiotics. *Antonie van Leeuwenhoek*, 69:109–117, 1996.

8. R. Ghosh and C. Tomlin. Lateral inhibition through delta-notch signaling: A piecewise affine hybrid model. In M. D. di Benedetto and A. L. Sangiovanni-Vincentelli, editors, *Hybrid Systems: Computation and Control, 4th International Workshop, Rome, Italy*, volume 2034 of *Lecture Notes in Computer Science*, pages 232–246. Springer-Verlag, 2001.
9. D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *J. of Phys. Chem.*, 81(25):2340–2361, 1977.
10. J. Hu, J. Lygeros, and S. Sastry. Towards a theory of stochastic hybrid systems. In N. Lynch and B.H. Krogh, editors, *Hybrid Systems: Computation and Control, 3rd International Workshop, Pittsburgh, PA*, volume 1790 of *Lecture Notes in Computer Science*, pages 160–173. Springer-Verlag, 2000.
11. P. Kiesau, U. Eikmanns, Z. G. Eckel, S. Weber, M. Hammelmann, and K. D. Entian. Evidence for a multimeric subtilin synthetase complex. *J. of Bacteriology*, 179(5):1475–1481, 1997.
12. M. Kleerebezem and L. E. Quadri. Peptide pheromone-dependent regulation of antimicrobial peptide production in gram-positive bacteria: a case of multicellular behavior. *Peptides*, 22:1579–1596, 2001.
13. H. H. McAdams and A. Arkin. Stochastic mechanisms in gene expression. *Proc. Natl. Acad. Sci.*, 94:814–819, 1997.
14. H. H. McAdams and A. Arkin. It's a noisy bussiness! genetic regulation at the nanomolar scale. *TIG*, 15(2):65–69, 1999.
15. T. Msadek. When the going gets tough: survival strategies and environmental signaling networks in Bacillus subtilis. *Trends in Microbiology*, 7(5):201–207, 1999.
16. E. M. Ozbudak, M. Thattai, I. Kurtser, A. D. Grossman, and A. van Oudenaarden. Regulation of noise in the expression of a single gene. *Nature Genetics*, 31:69–73, 2002.
17. C. V. Rao, D. M. Wolf, and A. P. Arkin. Control, exploitation and tolerance of intracellular noise. *Nature*, 402:231–237, 2002.
18. M. A. Shea and G. K. Ackers. The O_R control system of bacteriophage lambda. A physical-chemical model for gene regulation. *J. Mol. Biol.*, 181:211–230, 1985.
19. T. Stein, S. Borchert, P. Kiesau, S. Heinzmann, S. Kloss, C. Klein, M. Helfrich, and K. D. Entian. Dual control of subtilin biosynthesis and immunity in Bacillus subtilis. *Molecular Microbiology*, 44(2):403–416, 2002.