14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4-8, 2006, copyright by EURASIP

AN EQUIVALENT MARKOV MODEL FOR GILLESPIE'S STOCHASTIC SIMULATION ALGORITHM FOR BIOCHEMICAL SYSTEMS

Ronit Bustin, and Hagit Messer

School of Electrical Engineering
Tel-Aviv University 69978, ISRAEL
phone: 972-3-6408764, fax: 972-3-6407095, email: bustin,messer@eng.tau.ac.il

ABSTRACT

Mathematical/statistical modeling of biological systems is a desired goal for many years. It aims to be able to accurately predict the operation of such systems under various scenarios using computer simulations. In this paper we revisit Gillespie's Stochastic Simulation Algorithm for biochemical systems and we suggest an equivalent Markov Model for it. We show that under certain conditions it is a 1^{st} order homogenous Markov process and we analyze these conditions. Our suggested model can be used to simulate the probability density function of a biochemical processes which, in turn, can be used for applying statistical signal processing and information theory tools on them.

1. INTRODUCTION

A comprehensive understanding of biochemical processes is equivalent to having a mathematical model that can encounter and predict the experimental results of the biological system under different conditions. The stochastic approach [4, 5] is a model which describes by stochastic equations the dynamic of a biochemical system. The Stochastic Simulation Algorithm (SSA) is an algorithm that results in a single realization of the stochastic process, based on that model. As such, Monte Carlo approach is required in order to study the statistics of the process.

Statistical signal processing and information theory tools are based on availability of the statistics of the system or the process in hand. By approximating the stochastic approach model as a 1^{st} order Markov process, we provide a convenient formalism for the probability density function (pdf) of the stochastic approach, to be used for modeling and simulating biochemical processes. Moreover, while the stochastic approach is based on modeling the process in continuous time, discrete representation is more appropriate for simulation and signal processing. The equivalent Markov modeling describes the evolution of the process from one state to the other, over time samples which are design parameters. There choice is discussed in the sequel.

In [2] we study DNA repair processes in the *E. Coli* using an approximated Markov model [8]. The result of this work can be used to enable to simulate our suggested model, and to validate our suggested hypothesis regarding the optimization of the genetic information flow process.

The paper is organized as follows: In section 2, the stochastic approach and the SSA are reviewed. It is followed by a presentation and discussion of the equivalent Markov model (section 3). Section 4 provides summary and conclusions of the paper.

2. REVIEW OF THE STOCHASTIC APPROACH & SIMULATION

2.1 The time-evolution problem

The goal of many biologists at present time is to simulate the exact behavior of a biochemical system [1, 7]. Such simulations are used in order to predict actual results, understand the influence of specific chemical species in the system, and validating certain *in vivo* behaviors.

Definition 1 The time-evolution problem: *Assume a well-stirred mixture of N molecular species* $\{S_1, S_2, \ldots, S_N\}$ *in a fixed volume V at a constant temperature. These N species can chemically interact through M reactions* $\{R_1, R_2, \ldots, R_M\}$. *Given the initial number of molecules of each of the N species, what will be the molecular population at any time t?*

Define: $X_i(t) \quad \forall i \in \{1, 2, \ldots, N\}$ represents the number of molecules of the i^{th} species at time t and $x(t) = (X_1(t), X_2(t), \ldots, X_N(t))$, then, the question of "what will be the molecular population at any time t" is identical to asking "what is $x(t)$ for all t".
There are two formalisms for mathematically describing $x(t)$: The deterministic approach, and the stochastic approach.

2.2 The deterministic approach

The deterministic approach is the traditional way of solving $x(t)$ for all t. It begins by translating the chemical system into a set of differential equations:

$$\frac{dX_i(t)}{dt} = f_i(X_1(t), \ldots, X_N(t)) \quad \forall i \in \{1, \ldots, N\} \quad (1)$$

This set of equations is called the "reaction-rate equations" (RRE). Analytic solutions for the RRE can be found only for simple systems, thus the RRE set is usually solved numerically.
The deterministic approach is based on the assumptions that a chemical system is continuous in time and is deterministic. As explained by Gillespie in [5], both assumptions are not correct for a biochemical system. In particular, while there are cases where the deterministic formalism can be used, this is never the case for molecular systems, where even small fluctuations in the molecular populations can have large effects on the process outcomes.

2.3 The stochastic approach

2.3.1 The fundamental hypothesis of the stochastic formulation

In [5] Gillespie has shown that one can rigorously calculate the probability of a collision occurring in the volume V in any infinitesimal time interval. Thus, he asserts the following:

Definition 2 The fundamental hypothesis of the stochastic formulation: *$c_\mu dt$ = the average probability that a particular combination of the reactant molecules of reaction μ (R_μ) will react accordingly in the next infinitesimal time interval dt.*

That is, by multiplying $c_\mu dt$ by the total number of distinct combinations of R_μ reactant molecules in volume V at time t, one obtains the probability that the reaction R_μ will occur in the time interval $(t, t+dt)$.
As intuition suggests, the stochastic reaction constant c_μ is proportional to the reaction rate k_μ:

$$c_\mu = \frac{k_\mu \cdot l!}{V^{n-1}} \tag{2}$$

where l is the number of identical reactants in reaction R_μ, V is the volume of the system and n is the order of the reaction [9].

2.3.2 The master equation

In the center of the stochastic approach stands the "chemical master equation" (CME). The CME is the traditional approach of calculating the stochastic time evolution of a chemically reacting system. The key element in the CME is $P(X_1, X_2, \ldots, X_N; t)$ - the probability that at time t there will be in V: X_1 molecules of the molecular species S_1, X_2 molecules of the molecular species S_2, ..., and X_N molecules of the molecular species S_N.
We will use the following definitions [6]:

1. $a_\mu(x)dt$ is $c_\mu dt$ times the number of distinct R_μ reactants molecular combinations in a specific state $x = (X_1, X_2, \ldots, X_N)$. This is exactly the probability that the reaction R_μ will occur in V at time $(t, t+dt)$ given that the system is in state $x = (X_1, X_2, \ldots, X_N)$ at time t.
2. v_μ is the state-change vector for reaction R_μ. Its entries are defined by $v_{\mu,j}$, the change in the number of S_j molecules as a result of a single R_μ reaction.

The time evolution of the probability of x can therefore described by:

$$\begin{aligned} &P(x; t+dt|x_0; t_0) = \\ &P(x; t|x_0; t_0) \cdot [1 - \sum_{\mu=1}^{M} a_\mu(x)dt] + \\ &\sum_{\mu=1}^{M} a_\mu(x - v_\mu)dt \cdot P(x - v_\mu; t|x_0; t_0) \end{aligned} \tag{3}$$

The probability is constructed of two parts:

1. $P(x; t|x_0; t_0) \cdot [1 - \sum_{\mu=1}^{M} a_\mu(x)dt]$ is the probability that at time t the system is in the same state x times the probability that none of the M possible reactions occur.
2. $\sum_{\mu=1}^{M} a_\mu(x - v_\mu)dt \cdot P(x - v_\mu; t|x_0; t_0)$ is a sum of probabilities. Each element in the sum is the probability that at time t the system is in a state that is one reaction away from state x, times the probability that, given that the system is one state away from state x the desired reaction, that will bring us to state x, will occur.

It is simple to see that (3) leads to the following relation, which is the canonical form of the "master equation":

$$\begin{aligned} \frac{\partial}{\partial t} P(x; t|x_0; t_0) = &\sum_{\mu=1}^{M} a_\mu(x - v_\mu) \cdot P(x - v_\mu; t|x_0; t_0) \\ &- a_\mu(x) \cdot P(x; t|x_0; t_0) \end{aligned} \tag{4}$$

Although the "master equation" is exact, there are very few cases in which it can be solved analytically.

2.3.3 The reaction probability density function

In [5] Gillespie considered the question of simulating the stochastic time evolution. In order to do so, two critical questions need to be answered: *Given that the system is in state x at time t, when will the next reaction occur, and what kind of reaction will it be?* Since the system is of a stochastic nature, the answer to these two questions is stated by a probability function, $P(\tau, \mu)$, the probability that the next reaction in V will occur in the infinitesimal time interval $(t, t+\tau)$, and it will be the reaction R_μ, given the state x at time t. In [5] Gillespie claims the following:

1. Define $P_0(\tau)$ as the probability that no reaction will occur in the time interval $(t, t+\tau)$, given that at time t the system is in state x. Then, $P_0(\tau' + d\tau') = P_0(\tau') \cdot [1 - \sum_{\mu=1}^{M} a_\mu d\tau'] \Rightarrow P_0(\tau') = exp[-\sum_{\mu=1}^{M} a_\mu \tau']$.
2. $P(\tau, \mu)d\tau = P_0(\tau) \cdot a_\mu d\tau$, that is, the probability $P(\tau, \mu)d\tau$ can be calculated as the probability that nothing will happen in the time interval $(t, t+\tau)$, times $a_\mu d\tau$ - the probability that the reaction R_μ will occur in the time interval $(t+\tau, t+\tau+d\tau)$

From these two claims, Gillespie has derived the joint probability of $P(\tau, \mu)$:

$$P(\tau, \mu) = \begin{cases} a_\mu exp(-a_0\tau) & if \quad 0 \le \tau < \infty \quad and \\ & \quad \mu = 1, \ldots, M \\ 0 & otherwise \end{cases} \tag{5}$$

where $a_0 = \sum_{\mu=1}^{M} a_\mu$.
From (5) the marginal distributions can be evaluated [4]:

$$P(\mu) = \int_0^\infty a_\mu exp(-a_0\tau)d\tau = \frac{a_\mu}{a_0}, \quad \mu = 1, \ldots, M \tag{6}$$

$$P(\tau) = \sum_{\mu=1}^{M} a_\mu exp(-a_0\tau) \cdot u(\tau) = a_0 \cdot exp(-a_0\tau) \cdot u(\tau) \tag{7}$$

As such, the two random variables τ and μ are statistically independent, and one can write the joint provability function as a product of the two marginal distributions:

$$P(\tau, \mu) = P(\tau) \cdot P(\mu) = a_0 \cdot exp(-a_0\tau) \cdot u(\tau) \cdot \sum_{i=1}^{M} \delta(\mu - i) \cdot \frac{a_\mu}{a_0} \tag{8}$$

2.4 Implementation of the stochastic approach

In [4] and [5] Gillespie introduced two methods to simulate a stochastic process, described by the joint probability function $P(\tau,\mu)$: "The direct Method" and "The first-reaction Method". By carrying either one of these methods, one obtains a possible realization of the stochastic process defined by the joint probability function $P(\tau,\mu)$. As such, by using Monte Carlo techniques, $P(\tau,\mu)$ can be estimated. Both methods are exact and are equivalent to the "master equation" (4), in the sense that they are both rigorous consequences of the fundamental hypothesis (2). In 2.4.1 we will describe "The direct method". "The first-reaction Method" differs only slightly from "The direct method" as detailed in [4], and it was later improved by Gibson and Bruck [3].

2.4.1 The Direct Method

This method is based on the fact that the joint probability $P(\tau,\mu)$ can be written as a product of the two marginal distribution $P(\tau)$ and $P(\mu)$, as described in (8). The algorithm outputs the following results: a states vector that holds the sequence of states the system has gone through, and a time vector that holds the time each reaction occurred in. The algorithm is constructed of the following steps:

1. Initialization - Set the total time of the simulation as T. Set the time variable $t=0$ and update the time vector. Specify the initial state $x_0(t)$ (the initial amount of all the chemical reacting species $X_1,X_2,\ldots,X_N$) and update the states vector. Specify the reaction constants c_i for all M reactions as defined in (2).
2. Calculate the set of a_μ for all M possible rections, based on the current state $x(t)$ and the set of reaction constants c_μ (as defined in subsection 2.3.2).
3. Generate two random variables according to the joint distribution $P(\tau,\mu)$. Define these variables as τ_μ and μ (for the generation of these two random variables see appendix A).
4. Using the two random variables τ_μ and μ, one can tell which reaction occurred first and at what time:
   - According to the reaction μ, change the state: $x(t+\tau_\mu)=x(t)+v_\mu$ and update the states vector.
   - Update the time variable $t=t+\tau_\mu$ and update the time vector.
5. If the time variable $t\geq T$ - finish the simulation. Otherwise, go back to step 2.

The "direct method" is simple to implement and is exact. Its most visible drawback is the fact that one needs to go through each and every reaction, so its complexity is linear in the amount of reaction occurring in the time interval $[0,T]$.

3. THE STOCHASTIC APPROACH AS A MARKOV MODEL

Based on the fact that for a stochastic chemical system at state x at time t, the questions: *what will the next reaction be, and when will it happen?* are independent (subsection 2.3.3), we suggest the following:

Theorem 1 The stochastic chemical system is a first order Markov process: *given that the stochastic chemical system is currently in state x_i, the probability that the next state will be x_j depends only on the current state x_i, and not on past states.*

The proof of 1 lies in equation (6), which describes the probability that reaction R_μ will be the next reaction to occur. This probability, $P(\mu)$, does not dependent on the question: *when will the next reaction occur?*. It depends only on the values of a_μ, which is a function of the constant reaction rates c_μ and of the current state of the system x (through the number of reactant combinations for each of the M possible reactions). Therefore, as the next reaction to occur will determine the change in state, we can conclude that the process is a Markov order I process.

In order to describe the Markov model as a Markov chain, it must be a homogenous process, having a finite set of states and transfer probabilities that are time invariant. Given an initial state, x_0, the system has a finite set of states, limited by the initial amount of molecular species, and by the M possible reactions that can occur.

The transfer probabilities are given by:

$$\begin{aligned} &P(\text{reaction } \mu \text{ occurs before time } t \text{ from last reaction}) \\ &\quad = P(\mu)\cdot P(\text{next reaction will occur before time } t) \end{aligned} \tag{9}$$

To ensure homogenously, we need to grantee that the transfer probabilities, and − in particular the probability that the next reaction will occur before time t is time invariant. Assume we are currently in state x. We define the time segment Δt_x (depends on the current state x). The set of transition probabilities from state x is then given by:

$$\begin{aligned} P(\mu,\tau\leq\Delta t_x) &= P(\mu)\cdot P(\tau\leq\Delta t_x) \\ &= \frac{a_\mu}{a_0}\cdot\int_0^{\Delta t_x} a_0\cdot exp(-a_0\tau)d\tau \\ &= \frac{a_\mu}{a_0}\cdot(1-exp(-a_0\Delta t_x)) \end{aligned} \tag{10}$$

Thus, the probability to transfer to state x_i given that the system is currently in state x_j in the next time segment of Δt_{x_j}, is time invariant.

We have shown that the stochastic chemical reacting system can be described as an ergodic homogenous Markov order I process. However, the question of choosing the parameter Δt_x is still open.

Our goal is not only to find the transfer probabilities that describe the probability for each of the M possible reactions, but also to give an approximation of the time each reaction occurred.

The choice of the Δt_x parameters determines the quality of our approximation, and thus requires delicate handling. The following issues concern such an approximation:

1. Can we sample the biochemical random process, while keeping the ability of reconstructing the original process?
2. Assuming that no reaction occurred in a specific time segment, will the transfer probabilities change?
3. What is a "good approximation" of the continuous random process? What is the effects of the Δt_x parameters on it?
4. How should we define the event of an error, when analyzing an approximation model? Can we bound it?

3.1 Can we sample the biochemical random process, while keeping the ability of reconstructing the original process?

The biochemical process is a continuous random process. In order to present it as a Markov order I process, one must

sample the process. Assume for simplicity that all states x have equal values of the parameter a_0. We define the process t_i as the time in which reaction i took place.

As we saw in (7), τ, the random variable that represents the time from the last reaction to the next one, is exponentially distributed with the appropriate constant a_0. Thus, we can look at the set of random variables τ_i, each representing the time from the reaction $i-1$ to the reaction i, such that $t_i = \sum_{j=1}^{i} \tau_j$. Under our simplifying assumption, τ_i are i.i.d., exponentially distributed with parameter a_0. Thus, t_i is an *arrival process*, that describes the arrival times of a *Poisson* process (with parameter a_0). Putting all together, we can see that in order to describe the time in which reactions occurred, we need only to look at the derivative *Poisson* process. This process is a pure white process, with autocorrelation function:

$$R(l) = a_0 \cdot \delta(l) + (a_0)^2 \tag{11}$$

The Nyquist sampling theorem applies to wide-sense stationary (WSS) processes, with the spectrum of intrest being the power spectral density (PSD) function. Thus, a stationary process can be sampled without loss of information iff its PSD function is band-limited. We can conclude that for any set of Δt_x parameters, it is impossible to sample the random biochemical process without loss.

3.2 Assuming that no reaction occurred in a specific time segment, will the transfer probabilities change?

In (10) we derive the probabilities

$$P(\mu, t \leq \Delta t_x \text{ from the last reaction to occur})$$

Assuming that none of the reactions occurred, the state has not changed, so we use the same set of transition probabilities for the next time segment. We test if:

$$\begin{aligned} &P(\mu, t \leq 2 \cdot \Delta t_x \text{ from the last reaction} \mid \\ &\text{no reaction occurred } \Delta t_x \text{ since the last}) = \\ &P(\mu, t \leq \Delta t_x \text{ from the last reaction}) \end{aligned}$$

In order to prove this equality holds, we will use the fact that an exponentially distributed random variable is a *memoryless* random variable, thus:

$$P(X > t+s | X > t) = P(X > s) \tag{12}$$

which, in our case leads to the desired result:

$$\begin{aligned} &P(\mu, t \leq n \cdot \Delta t_x \quad | \quad t > (n-1) \cdot \Delta t_x) = \\ &P(\mu) \cdot P(t \leq n \cdot \Delta t_x \quad | \quad t > (n-1) \cdot \Delta t_x) = \\ &P(\mu) \cdot (1 - P(t > n \cdot \Delta t_x \quad | \quad t > (n-1) \cdot \Delta t_x)) = \\ &P(\mu) \cdot (1 - P(t > \Delta t_x)) = \\ &P(\mu) \cdot P(t \leq \Delta t_x) = \\ &P(\mu, t \leq \Delta t_x) \end{aligned} \tag{13}$$

3.3 What is a "good approximation" of the continuous random process? What is the effects of the Δt_x parameters on it?

The Markov chain we defined depends on the choice of the Δt_x parameters that divide the time line. Each realization will result in a different partitioning of the time line, since the Δt_x parameters are state dependent.

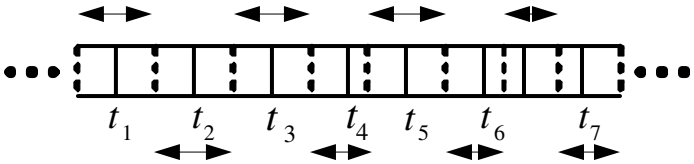

Figure 1: A "good approximation": every reaction time, t_i (the full lines) falls in a separate time segment, Δt_x (the dotted lines). If for *almost all* realization this is the case - the model is a "good approximation" of the process.

Our goal is to receive a "good approximation" of the continuous process using the Markov model, but first we must define *what is a "good approximation"?* We refer to a model as a "good approximation" of the continuous random process if for *almost all* realizations of the continuous random process the error for each reaction-time is limited to $\pm\Delta t_x$.

The choice of the Δt_x parameters will determine the quality of our model. Assume that Δt_x is a large time segment. In such a case, the probability $P(\tau \leq \Delta t_x)$ will be very close to 1, meaning that it is highly probable that the next reaction will occur before the end of the time segment Δt_x. For a very large Δt_x, the probability that more than a single reaction will occur in this time segment increases, and thus results in a model that, most probably, will not be a "good approximation". Sufficiently small Δt_x segments are more likely to result in a "good approximation", and to give an estimate of the reaction time that is about $\pm\Delta t_x$. Figure 1 illustrates a realization for which the model gives a "good approximation" of the reaction, since each reaction-time falls into a separate time segment, and thus the model approximation for such a realization will have an error of $\pm\Delta t_x$ for each reaction-time.

Now, lets assume a very small Δt_x. Such a choice results in $P(\tau \leq \Delta t_x) < 1$, so the $\sum_{\nu=1}^{M} P(\mu, \tau \leq \Delta t_x) < 1$, meaning there is a probability larger than zero to stay in the current state x. If we choose a Δt_x that is very small, $\sum_{\nu=1}^{M} P(\mu, \tau \leq \Delta t_x) \ll 1$, and the probability to stay in the current state will $\rightarrow 1$. Such a case is problematic since it will lead to multiple simulation steps before the next reaction takes place - unnecessary burden that will effect the efficiency of computations based on such a model. Moreover, choosing a small Δt_x, such that $\sum_{\nu=1}^{M} P(\mu, \tau \leq \Delta t_x) < 1$ (but not close to 1), ensures a non-periodic, ergodic Markov chain which converges to the stationary probability of the process.

3.4 How should we define the event of an error, when analyzing an approximation model? Can we bound it?

As we saw in subsection 2.3.3, the next reaction to occur and the time of the occurrence are independent events. Thus, the time of the next reaction to occur is independent of the specific reaction and is simply an exponentially distributed random variable:

$$P(\tau) = a_0 \cdot exp(-a_0 \tau) \cdot u(\tau) \tag{14}$$

As an exponential random variable, τ has mean $\frac{1}{a_0}$ and variance of $\frac{1}{(a_0)^2}$. Thus, taking into account all our knowledge of

τ and the effects of the different choices of Δt_x, our choice for Δt_x will be $\frac{1}{S \cdot a_0}$, where S is a positive constant larger than one. We would like to emphasize that the Δt_x parameters are state dependent since the value of a_0 is state dependent. Such a choice takes into account the fastest reaction possible in each state of the chemical system, since the fastest reaction possible in a specific state will have the largest a_μ that will have the largest effect in a_0. The fastest reaction will determine how small Δt_x must be to ensure a "good approximation", minimum simulation steps and an ergodic Markov chain.

In order to chose S we will look at a simplified case, where all states have equal values of a_0 and thus there is a single possible partition of the time line (all Δt_x are equal). We will define an error as a realization of the continuous random process that *can not* result from our model. In this simplified case, we will give an upper bound for the probability of an error. Note that the simplified case is realistic since the Markov model can be built with equal time segments based on the largest possible a_0 of the biochemical system.

$$\begin{aligned} Pr(error) < Pr(&\text{ there is at least a single case of} \\ &\text{consecutive reactions occurring} \\ &\text{within a time frame } \leq \frac{1}{S \cdot a_0}) \widehat{=} P \end{aligned} \quad (15)$$

This is an upper bound on the probability of error since not all cases of consecutive reactions occurring within a time frame smaller than or equal to $\frac{1}{S \cdot a_0}$ will lead to an error.

Let us assume now that we wish to simulate N reactions. We can write the bound (15) as:

$$\begin{aligned} P &= 1 - Pr(\text{the time difference between every two} \\ &\quad \text{consecutive reactions } > \frac{1}{S \cdot a_0}) \\ &= 1 - \prod_{i=2}^{N} Pr(\tau_i > \frac{1}{S \cdot a_0}) \\ &= 1 - exp(-\frac{N-1}{S}) \end{aligned} \quad (16)$$

The second transition makes use of the fact that the reaction-time of the next reaction depends only on the previous reaction. The last transition makes use of the fact that all reaction-times are exponentially distributed, and using our simplifying assumption, with the same parameter a_0.

As expected, the larger the number of reactions (N) is, the higher the possibility that at least one case of consecutive reactions will lead to an error. On the other hand, increasing S reduces Δt_x and reduces the probability of an error.

4. SUMMARY & CONCLUSIONS

Starting by presenting the stochastic approach for the time-evolution of biochemical systems, and the SSA by Gillespie as a tool for analyzing the statistics of a biochemical process using Monte Carlo simulations, we suggest an alternative road. Our analysis of the stochastic approach, allows us to approximate the time-evolution (or, the continuous random process) using a Markov Order I model. Such a model reduces the complexity of the pdf evaluation, and thus allows us the usage of a large variety of signal processing tools, that require the knowledge of the pdf. Using the Markov model in a straight forward manner, as an algorithm that replaces the SSA, is currently not feasible: the amount of states for even the simplest biochemical system is massive. There is a time-memory trade-off: the main problem in the SSA's is the time factor of the simulation, since they must go through each and every reaction. On the other hand, trying to use a Markov approximation, will theoretically reduce the time factor (since it can allow us any resolution we choose), but the amount of states requires a massive amount of memory.

A. THE GENERATION OF RANDOM VARIABLES IN THE SSA

Step 3 of the "Direct Method" algorithm requires the generation of two random variables according to the joint distribution $P(\tau, \mu)$. Since the two random variables are independent, one can simply generate two random variable, the first based on the exponential distribution $P(\tau)$, and the second based on the distribution $P(\mu)$. These two variables can be constracted using random variables, r_1 and r_2, independent and uniformly distributed in the range $[0,1]$, in the following manner:

$$\tau = \frac{1}{\sum_{i=1}^{M} a_i} \ln(\frac{1}{r_1}) \quad (17)$$

$$\sum_{i=1}^{\mu-1} a_i < r_2 \sum_{i=1}^{M} a_i \leq \sum_{i=1}^{\mu} a_i \quad (18)$$

REFERENCES

[1] A. Arkin, J. Ross and H. H. Mcadam "Stochastic Kinetic Analysis of Development Pathway Bifurcation in Phage λ-Infected *Escherichia coli* Cells", Genetics 149: 1633-1648, August 1998.

[2] R. Bustin and H. Messer, "Optimization Criteria for DNA Repair", In Proceedings of SSP'2005.

[3] M. A. Gibson and J. Bruck, "Efficient Exact Stochastic Simulation of Chenicl System with Many Species and Mant Channels", J. Phys. Chem. A 2000, 104, 1876-1889

[4] D. T. Gillespie, "A General Method for Numerically Simulating the Stochastic Time Evolution of Coupled Chemical Reactions", Journal of Computational Physics 22, 403-434, 1976.

[5] D. T. Gillespie, "Exact Stochastic Simulation of Coupled Chemical Reactions", The Journal of Physical Chemistry, Vol. 81 No. 25, 1977

[6] D. T. Gillespie, "Approximate accelerated stochastic simulation of chemically reacting systems", Journal of Chemical physics, Vol. 115, number 4.

[7] A. M. Kierzek, J. Zaim and P. Zielenkiewicz "The Effect of Transcription and Translation Initiation Frequencies on the Stochastic Fluctuations in Prokaryotic Gene Expression", The Journal of Biological Chemistry Vol. 276, No. 11, March 16, pp. 8165-8172.

[8] R. Sever and H. Messer, System approach to the DNA Repair Process in *E.coli*, In Proceedings of GENSIPS'2004.

[9] D. Voet and J. G. Voet, "Biochemistry, secon edition", John Wiley & Sons.