mathematics

Article

SiCaSMA: An Alternative Stochastic Description via Concatenation of Markov Processes for a Class of Catalytic Systems

Vincent Wagner and Nicole Erika Radde *

Institute for Systems Theory and Automatic Control, University of Stuttgart, Pfaffenwaldring 9,
70569 Stuttgart, BW, Germany; vincent.wagner@ist.uni-stuttgart.de
* Correspondence: nicole.radde@ist.uni-stuttgart.de

Abstract: The Chemical Master Equation is a standard approach to model biochemical reaction networks. It consists of a system of linear differential equations, in which each state corresponds to a possible configuration of the reaction system, and the solution describes a time-dependent probability distribution over all configurations. The Stochastic Simulation Algorithm (SSA) is a method to simulate sample paths from this stochastic process. Both approaches are only applicable for small systems, characterized by few reactions and small numbers of molecules. For larger systems, the CME is computationally intractable due to a large number of possible configurations, and the SSA suffers from large reaction propensities. In our study, we focus on catalytic reaction systems, in which substrates are converted by catalytic molecules. We present an alternative description of these systems, called SiCaSMA, in which the full system is subdivided into smaller subsystems with one catalyst molecule each. These single catalyst subsystems can be analyzed individually, and their solutions are concatenated to give the solution of the full system. We show the validity of our approach by applying it to two test-bed reaction systems, a reversible switch of a molecule and methyltransferase-mediated DNA methylation.

Keywords: chemical master equation; DNA methylation; catalytic systems; stochastic simulation algorithm; time-continuous Markov process; equivalence of stochastic processes

check for updates

Citation: Wagner, V.; Radde, N.E. SiCaSMA: An Alternative Stochastic Description via Concatenation of Markov Processes for a Class of Catalytic Systems. *Mathematics* **2021**, *9*, 1074. https://doi.org/10.3390/math9101074

Academic Editor: Xavier Bardina

Received: 26 March 2021
Accepted: 4 May 2021
Published: 11 May 2021

1. Introduction

Different approaches exist to model (bio)chemical reaction networks. On one end of the spectrum, quantum-theoretical approaches provide insights into the dynamics of few-atom systems using nearly no approximations of the underlying physics. Force-field driven particle simulations approximate the positions and velocities of atoms or molecules deterministically and are therefore capable of simulating significantly larger systems. By assuming a well-stirred, isothermal and isobaric system, one can eliminate the spatial component completely, thus further increasing the feasible system size.

Without the spatial dependency, the only time-dependent system state is characterized by the number of molecules of each chemical species at a given time. An established method to describe the behavior of such a system is the Chemical Master Equation (CME). It is a system of coupled linear ordinary differential equations whose solution describes a time-dependent probability distribution over the set of possible states of the system (see, e.g., Higham [1], Wilkinson [2], Schnoerr et al. [3] for reviews on stochastic modeling approaches for chemical reactions).

The CME suffers from the curse of dimensionality, i.e., computational complexity and memory requirement grow exponentially with the number of chemical species. Thus, exact solutions of the CME are rare and only possible for very small and simple systems. Moreover, depending on the system at hand, it can be tedious to even find a proper state definition and to list all possible states.

Different approaches are used to approximate the solution of a CME. Some of these date back to the 1960s, when the use of stochastic models for natural phenomena became quite popular [4,5]. For example, the Stochastic Simulation Algorithm (SSA) is a Monte Carlo method which generates sample paths that are drawn from the CME's true solution [6]. It rapidly becomes computationally expensive if reaction propensities are large, which is usually the case already with a moderate number of molecules. Furthermore, many samples are needed for an accurate estimation of system parameters, especially if the system operates near an unstable point. Different methods have been introduced to increase the efficiency of the SSA. τ-leaping, for example, allows for several reactions to happen in a given time interval $[t, t+\tau]$ by neglecting changes in the reaction propensities during $[t, t+\tau]$. Other approaches approximate the CME's solution by calculating its central moments up to a certain order. The evolution of these moments is described by a set of coupled ODEs. Since the ODE system is not closed if any reaction propensity is superlinear with respect to the state of the system, closure schemes have to be applied [7,8]. Similar to moment closure methods, linear noise approximations or the chemical Langevin equation approximate the discrete state space with continuous quantities. In addition to the already mentioned techniques, finite state projections [9] or sliding window methods [10] and similar approaches [11] truncate the state space by neglecting states with low probability masses.

However, although many methods have been introduced to amplify the efficiency of CME-based approaches, these approaches quickly come to their limits for complex systems, and, hence, stochastic modeling of intracellular processes is still computationally challenging.

In recent years, especially the experimental methodology for single cells has experienced rapid growth. This not only enables new insights into heterogeneous cellular responses caused by stochastic effects, it also spotlights the need for simulation models and procedures that are able to include and predict said stochasticity. Different studies have already investigated this ubiquitous phenomenon in the context of gene expression [12,13], intracellular transport processes [14] or signal transduction pathways [15–17].

In this paper, we introduce a new methodology, the Single Catalyst Stochastic Modeling Approach (SiCaSMA), to describe catalytic reaction systems. This class of systems consists of substrate molecules which can be converted by catalysts. Instead of simulating the system as one large process, we subdivide the full system into smaller parts in which the catalyst molecules are simulated one after another, solve the corresponding equations and recombine the solutions. This either enables the direct solution of the CME or at least eases the implementation of the SSA for the smaller subsystems. Our approach is based on the assumption that catalyst molecules react independently of each other. The advantages of our approach are manifold:

1. It is not necessary to define the state transition graph of the entire system. This can be a real advantage, since the state transition graph can become quite large. This holds in particular for catalytic systems in which the substrate exists in many conformations. A prominent example is post-translational modification of a protein, e.g., phosphorylation at different sites. The nodes of the transition graph correspond here to all possible configurations of substrate phosphorylation states and catalyst binding states. Due to the combinatorial complexity, this number grows rapidly even for small molecule numbers. Moreover, it is for most of those systems not possible to exploit the underlying structure of the graph.
2. Intractable state transition graphs are replaced by a concatenation of much simpler graphs, resulting in lower dimensional differential equations for the CME approach. Instead of solving the conventional CME defined on the full state transition graph, one can solve the CME on these simpler graphs and concatenate the obtained solutions. This effectively enables the solution of the CME for complex systems, where it was formerly necessary to resort to simulation methods.
3. The implementation of the SSA is considerably simplified.

We show on different example systems, including a simple conversion reaction and a model for methyltransferase-mediated DNA methylation, that SiCaSMA is equivalent to the standard CME description of the full system.

2. Materials and Methods

The CME is a set of linear differential equations that describe the evolution of probability distribution $P(\mathbf{X}, t)$ over all possible system states $\mathbf{X}$ and time t:

$$\frac{dP(\mathbf{X},t)}{dt} = \sum_{j=1}^{M} \Big[\mathbf{X} - \boldsymbol{\nu}_j^- \geq 0\Big] a_j(\mathbf{X} - \boldsymbol{\nu}_j) P(\mathbf{X} - \boldsymbol{\nu}_j, t) - \sum_{j=1}^{M} \Big[\mathbf{X} + \boldsymbol{\nu}_j^+ \geq 0\Big] a_j(\mathbf{X}) P(\mathbf{X}, t) \tag{1}$$

Here, $j \in \{1, ..., M\}$ is the reaction index and $\boldsymbol{\nu}_j$ the state change vector of reaction j. Reaction propensities are denoted $a_j(\mathbf{X})$. Conditions in square brackets ensure that the summation is over feasible reactions only, where $\boldsymbol{\nu}_j = \boldsymbol{\nu}_j^+ + \boldsymbol{\nu}_j^-$ is decomposed into positive and negative parts. Equation (1) can be cast into the form

$$\dot{P}(\mathbf{X}, t) = SP(\mathbf{X}, t) \tag{2}$$

with system matrix S and solution

$$P(\mathbf{X}, t) = e^{St} P(\mathbf{X}, t = 0). \tag{3}$$

Hence, solving the CME requires the calculation of the exponential of the product of the system matrix S and time t.

The SSA simulates sample paths from the CME. For each reaction step, this includes the realization of two random variables, an exponentially distributed waiting time $\tilde{t}$ for the next reaction to happen, and a reaction index j to determine the type of reaction, as depicted in Algorithm 1.

Algorithm 1 SSA general scheme ($\mathbf{X}_{init}, t_f$, kinetic params)

```
Initialize $\mathbf{X}(0) = \mathbf{X}_{init}$ and set $t = 0$
while $t < t_f$ do
    Calculate $a_j(\mathbf{X}(t))$ and $a_{sum}(\mathbf{X}(t)) = \sum_j a_j(\mathbf{X}(t))$ for $j = 1, \ldots, m$
    Draw waiting time $\tilde{t}$ until next reaction event from $\tilde{T} \sim \mathrm{Exp}(a_{sum}(\mathbf{X}(t)))$
    Draw reaction index $j$ from discrete distribution $J \sim {a_j(\mathbf{X}(t))}/{a_{sum}(\mathbf{X}(t))}$
    Update $t = t + \tilde{t}$
    if $t < t_f$ then
        Update $\mathbf{X}(t) = \mathbf{X}(t) + \boldsymbol{\nu}_j$
    end if
end while
return $\mathbf{X}(t)$
```

3. Results

3.1. An Intuitive Example: Linear Conversion Process

As an intuitive example to illustrate the idea behind our approach, we consider the reversible conversion of molecules between two different forms with simple first order kinetics (Figure 1A), i.e., reactions of the form

$$\mathrm{A} \underset{\mathrm{k}_{-1}}{\overset{\mathrm{k}_1}{\rightleftarrows}} \mathrm{B} \cdot$$

This reaction system constitutes a special and very simple case of our system class. It represents, for example, reversible binding of a catalyst to a substrate, without modification of the substrate. The forms A and B represent free enzyme and enzyme substrate complex, respectively. When substrate is in excess, reaction rates only linearly depend on the amount of catalyst molecules and are of order zero with respect to the substrate, as assumed here. Especially for this simple system, it is intuitively clear that we can either describe the full system or consider each molecule independently and obtain equivalent solutions. We define the state $X(t)$ of the underlying Markov process to be the number of molecules in state A. Thus, for n molecules, the range of $X(t)$ is given by

$$X(t) \in \{0, 1, \dots, n\}. \tag{4}$$

The propensities and state change vectors of the two reactions are given by

$$a_1(X) = k_1 X, \quad a_{-1}(X) = k_{-1}(n - X), \quad \nu_1 = -1, \quad \nu_{-1} = 1. \tag{5}$$

The state transition graph of this reaction system for $n = 2$ is shown in Figure 1B (left). This system can alternatively be modeled by concatenating n single molecule systems with state spaces $Y^i(t) \in \{0, 1\}, i = 1, \dots, n$ (Figure 1B (right)). In this case, the concatenation is equivalent to a summation $Y(t) = \sum_{i=1}^{n} Y^i(t)$.

Next, we show that $X(t)$ and $Y(t)$ describe equivalent stochastic processes. We therefore solve the CME corresponding to each system analytically and compare the resulting probability distributions $P(X, t)$ and $P(Y, t)$. Regarding the original system, $P(X, t)$ corresponds to the solution of the respective CME, which can be derived from the state transition graph (Figure 1B) and reads

$$\dot{P}(X, t) = \begin{pmatrix} \dot{P}(X = 2, t) \\ \dot{P}(X = 1, t) \\ \dot{P}(X = 0, t) \end{pmatrix} = \begin{pmatrix} -2k_1 & k_{-1} & 0 \\ 2k_1 & -(k_1 + k_{-1}) & 2k_{-1} \\ 0 & k_1 & -2k_{-1} \end{pmatrix} \begin{pmatrix} P(X = 2, t) \\ P(X = 1, t) \\ P(X = 0, t) \end{pmatrix}. \tag{6}$$

Without loss of generality and for the sake of brevity, we choose $k_1 = k_{-1} = 1$ for the following calculations. Then, the solution of this set of linear differential equations is given by

$$P(X, t) = \frac{1}{4} \begin{pmatrix} 1 + e^{-4t} + 2e^{-2t} & 1 - e^{-4t} & 1 + e^{-4t} - 2e^{-2t} \\ 2 - 2e^{-4t} & 2 + 2e^{-4t} & 2 - 2e^{-4t} \\ 1 + e^{-4t} - 2e^{-2t} & 1 - e^{-4t} & 1 + e^{-4t} + 2e^{-2t} \end{pmatrix} P(X, t = 0), \tag{7}$$

and for $X(t = 0) = 2$

$$P(X, t) = \frac{1}{4} \begin{pmatrix} 1 + e^{-4t} + 2e^{-2t} \\ 2 - 2e^{-4t} \\ 1 + e^{-4t} - 2e^{-2t} \end{pmatrix}. \tag{8}$$

Applying SiCaSMA, for the equivalent process, we solve each subsystem to obtain $P(Y^i, t), i = 1, 2$ in a first step. The respective CME reads

$$\dot{P}(Y^i, t) = \begin{pmatrix} \dot{P}(Y^i = 1, t) \\ \dot{P}(Y^i = 0, t) \end{pmatrix} = \begin{pmatrix} -k_1 & k_{-1} \\ k_1 & -k_{-1} \end{pmatrix} \begin{pmatrix} P(Y^i = 1, t) \\ P(Y^i = 0, t) \end{pmatrix}, \tag{9}$$

with solution

$$P(Y^i, t) = \frac{1}{2} \begin{pmatrix} 1 + e^{-2t} & 1 - e^{-2t} \\ 1 - e^{-2t} & 1 + e^{-2t} \end{pmatrix} P(Y^i, t = 0). \tag{10}$$

This gives for $Y^i(t = 0) = 1$

$$P(Y^i, t) = \frac{1}{2} \begin{pmatrix} 1 + e^{-2t} \\ 1 - e^{-2t} \end{pmatrix}. \tag{11}$$

Since $Y(t) = \sum_{i=1}^{n} Y^i(t)$, the probability distribution $P(Y,t)$ is formally defined as a convolution of the probability distributions $P(Y^i,t)$. For $n = 2$, this results in $P(Y,t) = P(Y^1,t) * P(Y^2,t)$ and can also be formulated in a verbose fashion

$$\begin{aligned} P(X=2,t) &= P(Y^1=1,t)P(Y^2=1,t) \\ P(X=1,t) &= P(Y^1=1,t)P(Y^2=0,t) + P(Y^1=0,t)P(Y^2=1,t) \\ P(X=0,t) &= P(Y^1=0,t)P(Y^2=0,t). \end{aligned} \tag{12}$$

$P(X,t)$ and $P(Y,t)$ are equivalent, as can be derived from the Equations (11) and (12) in comparison to the solution of the full system (Equation (8)) and seen in the courses in Figure 1C.

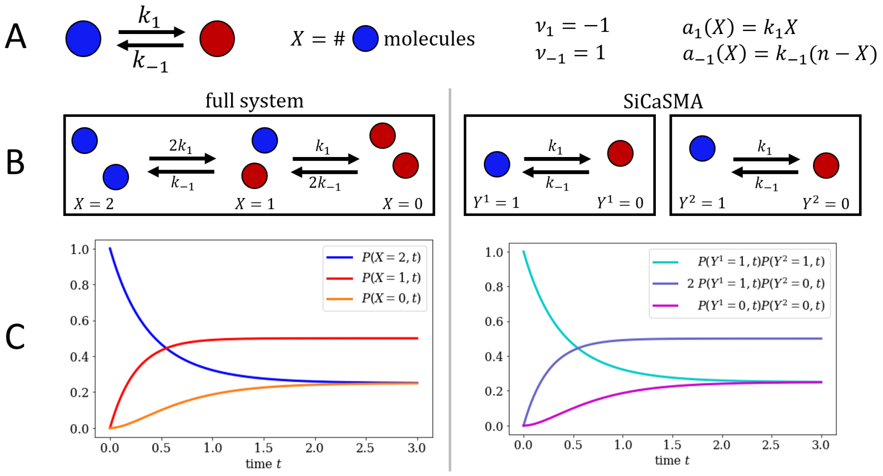


Figure 1. Application of SiCaSMA to a molecular conversion process. (**A**) We consider a reversible conversion reaction, in which the system state X is defined as the number of molecules in the first (blue) state. State change vectors and propensities are denoted ν_1, ν_{-1} and $a_1(X), a_{-1}(X)$, respectively. (**B**) State transition graph of the full system (**left**) and both single catalyst subsystems (**right**). (**C**) Visual representation of the solution $P(X,t)$ of the CME for the full system (**left**) and of the combined solution $P(Y,t)$ of the two single catalyst subsystems (**right**).

3.2. *Applying SiCaSMA to a Model for DNMT1-Mediated DNA Methylation*

While the just-described system is rather simple in nature, it provides an intuitive application of our approach to general catalytic systems, which are assumed to be of the following form: The chemical species can be subdivided into two categories. C denotes the catalyst species and S the substrate species, respectively. Catalyst-driven conversions of the substrate, such as conformational changes or post-transcriptional modifications, are treated as individual species in this framework. The key assumption for our approach is that the catalyst molecules function independently of each other. In our system, substrate molecules S are modified by catalysts C. Our specific system comprises reversible complex formation of catalyst and substrate molecules and catalyst-mediated modifications of the substrate. If multiple modifications are possible, this is implemented as a counter, and denoted by S^i for a single substrate molecule or CS^i in complex form, respectively,

$$\mathrm{C} + \mathrm{S}^i \underset{\mathrm{k}_{-1}}{\overset{\mathrm{k}_1}{\rightleftarrows}} \mathrm{CS}^i \quad i = 0, \ldots, n_S$$
$$\mathrm{CS}^i \xrightarrow{\mathrm{k}_m} \mathrm{CS}^{i+1} \quad i = 0, \ldots, n_S - 1.$$

While all catalyst and substrate molecules are simulated simultaneously in the classical approaches, we replace this procedure with consecutive simulations including all substrate molecules and only one catalyst molecule at a time. Both approaches differ in one major aspect: In the classical approach, each catalyst molecule faces the same substrate state at a given time, and if one catalyst molecule changes this substrate state, this change is directly visible for all other catalysts. Using SiCaSMA, however, a catalyst molecule initially faces the substrate state that results from all modifications of the previously simulated catalyst molecules. The equivalence of both procedures is therefore non-trivial. Moreover, unlike in the previous example, the subsystems cannot be simulated in parallel, since they indirectly depend on each other through the initial condition in the substrate state.

On this basis, we apply our method to a model for epigenetic regulation via DNA methylation, as described in Adam et al. [18]. In this system, the protein DNA methyltransferase 1 (DNMT1) can methylate DNA molecules at different sites and therefore serves as biocatalyst. The DNA strands serve as substrate molecules. Each DNA molecule contains 44 methylation sites in the original model. Analogously to Adam et al. [18], we model the binding reaction of DNMT1 and DNA with a propensity that increases linearly with the number of unbound DNMT1 molecules and is independent of the number of DNA molecules. This assumption comes from the fact that a DNA strand contains many (unspecific) sites at which DNMT1 can bind. Furthermore, each DNA molecule can accommodate arbitrarily many DNMT1 proteins, such that the DNMT1 species is the limiting factor here.

For an analytic CME solution, we have to simplify this model in order to obtain a tractable number of states. We do this by neglecting different DNMT1 conformations and consider a system consisting of $n_D = 1$ DNA molecule with $n_S = 1$ methylation site and $n_P = 2$ DNMT1 molecules. This keeps the system size tractable and enables a clear illustration. The reaction scheme is shown in Figure 2A. DNMT1 can bind to and dissolve from the DNA. In the bound state, it can methylate the DNA processively. The state of the system is described by a two-dimensional vector, in which the first and second component correspond to the number of enzymes bound to the DNA molecule and the number of methylated sites of this DNA molecule, respectively. The first entry, therefore, serves as catalyst species C, while the second one represents the substrate species S. Methylation reactions are modeled as irreversible processes in our system. We obtain the following reaction propensities and state transition vectors:

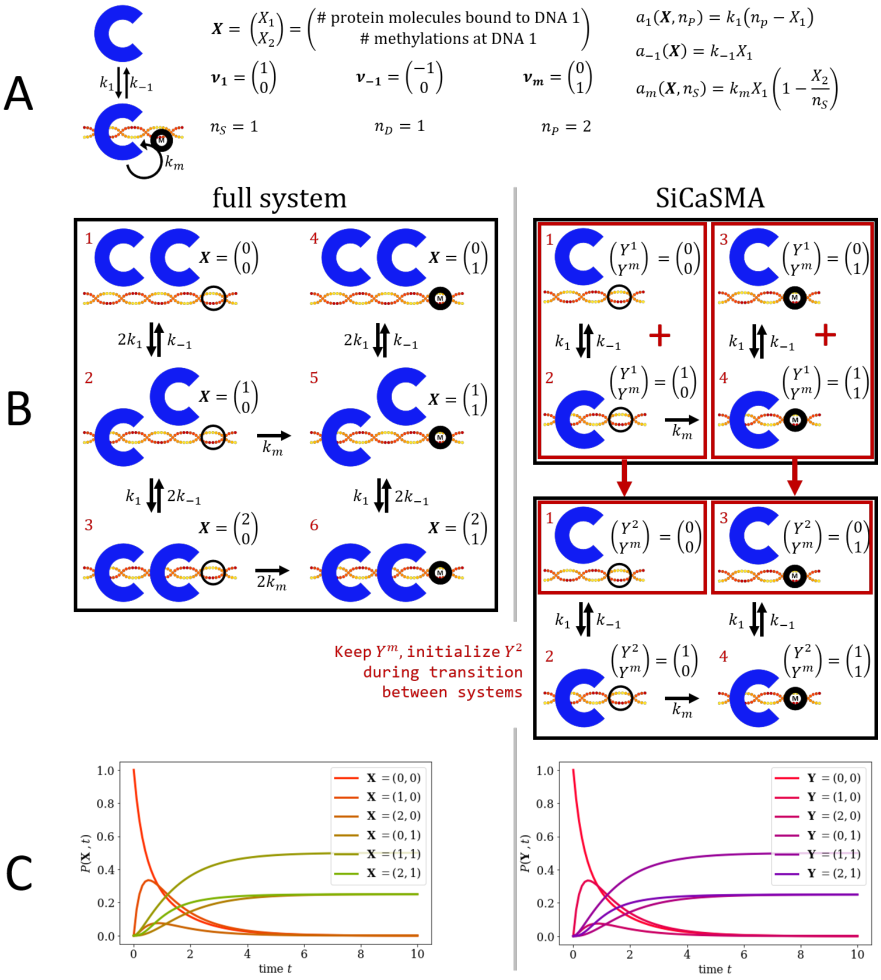
$$\begin{aligned} &a_1(\mathbf{X}, n_P) = k_1(n_P - X_1) && a_{-1}(\mathbf{X}) = k_{-1}X_1 && a_m(\mathbf{X}, n_S) = k_m X_1\left(1 - \frac{X_2}{n_S}\right) \\ &\boldsymbol{\nu}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} && \boldsymbol{\nu}_{-1} = \begin{pmatrix} -1 \\ 0 \end{pmatrix} && \boldsymbol{\nu}_m = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \end{aligned} \tag{13}$$

Here, the factor $(1 - X_2/n_S)$ in $a_m(\mathbf{X}, n_S)$ indicates the fraction of yet unmethylated sites. The state transition graph of the CME for the full system is depicted in Figure 2B (left). It comprises six possible states; hence, the CME is a six-dimensional linear differential equation system, which is given by

$$\dot{P}(\mathbf{X}, t) = \begin{pmatrix} -2k_1 & k_{-1} & 0 & 0 & 0 & 0 \\ 2k_1 & -(k_{-1} + k_1 + k_m) & 2k_{-1} & 0 & 0 & 0 \\ 0 & k_1 & -(2k_{-1} + 2k_m) & 0 & 0 & 0 \\ 0 & 0 & 0 & -2k_1 & k_{-1} & 0 \\ 0 & k_m & 0 & 2k_1 & -(k_{-1} + k_1) & 2k_{-1} \\ 0 & 0 & 2k_m & 0 & k_1 & -2k_{-1} \end{pmatrix} P(\mathbf{X}, t). \tag{14}$$

For the initial state $\mathbf{X} = (0, 0)^T$ and $k_1 = k_{-1} = k_m = 1$, the solution of Equation (14) is given by

$$P(\mathbf{X},t) = \begin{pmatrix} \frac{2e^{-3t}}{5} + \frac{2e^{-\sqrt{5}t-3t}}{5(3+\sqrt{5})} - \frac{2e^{\sqrt{5}t-3t}}{5(\sqrt{5}-3)} \\ -\frac{2e^{-3t}}{5} - \frac{2e^{-\sqrt{5}t-3t}}{5(3+\sqrt{5})} - \frac{2e^{-\sqrt{5}t-3t}}{\sqrt{5}(\sqrt{5}+3)} + \frac{2e^{\sqrt{5}t-3t}}{5(\sqrt{5}-3)} - \frac{2e^{\sqrt{5}t-3t}}{\sqrt{5}(\sqrt{5}-3)} \\ -\frac{2e^{-3t}}{5} + \frac{e^{-\sqrt{5}t-3t}}{5} + \frac{e^{\sqrt{5}t-3t}}{5} \\ \frac{1}{4} + \frac{e^{-4t}}{4} - \frac{2e^{-3t}}{5} + \frac{e^{-2t}}{2} - \frac{2e^{-\sqrt{5}t-3t}}{5(3+\sqrt{5})} + \frac{2e^{\sqrt{5}t-3t}}{5(\sqrt{5}-3)} \\ \frac{1}{2} - \frac{e^{-4t}}{2} + \frac{2e^{-3t}}{5} + \frac{2e^{-\sqrt{5}t-3t}}{5(3+\sqrt{5})} + \frac{2e^{-\sqrt{5}t-3t}}{\sqrt{5}(\sqrt{5}+3)} - \frac{2e^{\sqrt{5}t-3t}}{5(\sqrt{5}-3)} + \frac{2e^{\sqrt{5}t-3t}}{\sqrt{5}(\sqrt{5}-3)} \\ \frac{1}{4} + \frac{e^{-4t}}{4} + \frac{2e^{-3t}}{5} - \frac{e^{-2t}}{2} - \frac{e^{-\sqrt{5}t-3t}}{5} - \frac{e^{\sqrt{5}t-3t}}{5} \end{pmatrix}. \tag{15}$$



A
k_1 k_{-1}
k_m
M
$\boldsymbol{X} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} \text{\# protein molecules bound to DNA 1} \\ \text{\# methylations at DNA 1} \end{pmatrix}$
$\boldsymbol{\nu}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$
$\boldsymbol{\nu}_{-1} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$
$\boldsymbol{\nu}_m = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$
$n_S = 1$
$n_D = 1$
$n_P = 2$
$a_1(\boldsymbol{X}, n_P) = k_1(n_p - X_1)$
$a_{-1}(\boldsymbol{X}) = k_{-1}X_1$
$a_m(\boldsymbol{X}, n_S) = k_m X_1 \left(1 - \frac{X_2}{n_S}\right)$
B
full system
SiCaSMA
1 $\boldsymbol{X} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$
4 $\boldsymbol{X} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$
$2k_1$ k_{-1}
2 $\boldsymbol{X} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$
5 $\boldsymbol{X} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$
k_m
k_1 $2k_{-1}$
3 $\boldsymbol{X} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$
6 $\boldsymbol{X} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$
$2k_m$
1 $\begin{pmatrix} Y^1 \\ Y^m \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$
3 $\begin{pmatrix} Y^1 \\ Y^m \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$
+
2 $\begin{pmatrix} Y^1 \\ Y^m \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$
4 $\begin{pmatrix} Y^1 \\ Y^m \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$
1 $\begin{pmatrix} Y^2 \\ Y^m \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$
3 $\begin{pmatrix} Y^2 \\ Y^m \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$
2 $\begin{pmatrix} Y^2 \\ Y^m \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$
4 $\begin{pmatrix} Y^2 \\ Y^m \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$
Keep Y^m, initialize Y^2 during transition between systems
C
$P(\mathbf{X}, t)$
$\mathbf{X} = (0,0)$
$\mathbf{X} = (1,0)$
$\mathbf{X} = (2,0)$
$\mathbf{X} = (0,1)$
$\mathbf{X} = (1,1)$
$\mathbf{X} = (2,1)$
time t
$P(\mathbf{Y}, t)$
$\mathbf{Y} = (0,0)$
$\mathbf{Y} = (1,0)$
$\mathbf{Y} = (2,0)$
$\mathbf{Y} = (0,1)$
$\mathbf{Y} = (1,1)$
$\mathbf{Y} = (2,1)$
time t
0.0
0.2
0.4
0.6
0.8
1.0
0
2
4
6
8
10

Figure 2. Application of SiCaSMA to a model for DNMT1-mediated DNA methylation. (**A**) We consider a system of DNMT1 molecules (blue) and DNA molecules in which DNMT1 can reversibly bind to DNA and processively methylate different sites. The state $\mathbf{X}$, the state-transition vectors $\boldsymbol{\nu}_r$ and the propensities $a_r(\mathbf{X})$ for all reactions $r \in \{1, -1, m\}$ are indicated. The numbers n_S, n_D and n_P specify the number of methylation sites per DNA strand, the number of simulated DNA strands and the number of DNMT1 proteins, respectively. (**B**) State space of the full system (**left**) and the two SiCaSMA single catalyst subsystems including only one DNMT1 each (right). In the CME, we have used the ordering of states as indicated with red numbers on the top left of each system state. (**C**) Analytic solution of the full CME (left) and SiCaSMA as provided in Equation (A3) (**right**) for parameters $k_1 = k_{-1} = k_m = 1$.

The stochastic process, which results from applying SiCaSMA to the system, is illustrated in Figure 2B (right). Here, the system is modeled by simulating the two DNMT1 molecules one after another while accumulating the methylation state. For each single catalyst subsystem indicated by $p = 1, 2$, we obtain a CME that comprises four states,

$$\dot{P}\left(\begin{pmatrix} Y^p \\ Y^m \end{pmatrix}, t\right) = \underbrace{\begin{pmatrix} -k_1 & k_{-1} & 0 & 0 \\ k_1 & -(k_{-1} + k_m) & 0 & 0 \\ 0 & 0 & -k_1 & k_{-1} \\ 0 & k_m & k_1 & -k_{-1} \end{pmatrix}}_{S} P\left(\begin{pmatrix} Y^p \\ Y^m \end{pmatrix}, t\right). \tag{16}$$

The general solution of this equation is given by

$$P\left(\begin{pmatrix} Y^p \\ Y^m \end{pmatrix}, t\right) = e^{tS} P\left(\begin{pmatrix} Y^p \\ Y^m \end{pmatrix}, t = 0\right). \tag{17}$$

Using

$$c_- = exp\left(\frac{-\sqrt{5}t - 3t}{2}\right) \quad \text{and} \quad c_+ = exp\left(\frac{\sqrt{5}t - 3t}{2}\right), \tag{18}$$

the matrix exponent e^{tS} is given by

$$e^{tS} = \begin{pmatrix} \frac{2c_-}{5+\sqrt{5}} - \frac{(5-\sqrt{5})c_+}{5\sqrt{5}-15} & -\frac{(1+\sqrt{5})c_-}{5+\sqrt{5}} + \frac{c_+}{\sqrt{5}} & 0 & 0 \\ \frac{(\sqrt{5}-1)c_-}{\sqrt{5}-5} + \frac{\sqrt{5}c_+}{5} & \frac{(\sqrt{5}+3)c_-}{\sqrt{5}+5} + \frac{(\sqrt{5}+5)c_+}{5\sqrt{5}+15} & 0 & 0 \\ \frac{(1-\sqrt{5})c_-}{2\sqrt{5}} - \frac{2\sqrt{5}c_+}{5\sqrt{5}-5} + \frac{1}{2} + \frac{e^{-2t}}{2} & \frac{(\sqrt{5}+3)c_-}{5+3\sqrt{5}} - \frac{(5+\sqrt{5})c_+}{5+5\sqrt{5}} + \frac{1}{2} - \frac{e^{-2t}}{2} & \frac{1}{2} + \frac{e^{-2t}}{2} & \frac{1}{2} - \frac{e^{-2t}}{2} \\ \frac{(1-\sqrt{5})c_-}{\sqrt{5}-5} - \frac{\sqrt{5}c_+}{5} + \frac{1}{2} - \frac{e^{-2t}}{2} & -\frac{(\sqrt{5}+3)c_-}{\sqrt{5}+5} - \frac{(\sqrt{5}+5)c_+}{5\sqrt{5}+15} + \frac{1}{2} + \frac{e^{-2t}}{2} & \frac{1}{2} - \frac{e^{-2t}}{2} & \frac{1}{2} + \frac{e^{-2t}}{2} \end{pmatrix}. \tag{19}$$

All above equations hold for both single catalyst subsystems. However, the initial state of the second system depends on the solution of the first one. The state transition graph of the single catalyst subsystems (Figure 2B (right)) consists of two classes. The transient class, which contains all states with unmethylated DNA, is enclosed by the left red box. Similarly, the absorbing class, which contains the states with methylated DNA, is depicted inside the right red box. Exchanging the first catalyst molecule by the second one corresponds to keeping the methylation state from the first subsystem and starting in the respective class in the DNMT1 unbound state, i.e., the DNMT1 protein is re-initialized. The probabilities within each class are therefore added up, which is indicated by the red plus-sign. The distribution of the initial state of the second subsystem reads

$$\begin{aligned} P\left(\begin{pmatrix} Y^2 \\ Y^m \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, t = 0\right) &= P\left(\begin{pmatrix} Y^1 \\ Y^m \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, t = t_f\right) + P\left(\begin{pmatrix} Y^1 \\ Y^m \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, t = t_f\right) \\ P\left(\begin{pmatrix} Y^2 \\ Y^m \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, t = 0\right) &= P\left(\begin{pmatrix} Y^1 \\ Y^m \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, t = t_f\right) + P\left(\begin{pmatrix} Y^1 \\ Y^m \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, t = t_f\right), \end{aligned} \tag{20}$$

where t_f is the simulated time. The two states in which the catalyst molecule is bound have probability 0 for $t = 0$. From Y^1, Y^2 and Y^m, we construct the overall system state $\mathbf{Y}$, which will be shown to be equivalent to $\mathbf{X}$. We therefore define:

$$\mathbf{Y} = \begin{pmatrix} Y^1 + Y^2 \\ Y^m \end{pmatrix} \tag{21}$$

As for the first example system, the probability distribution $P(Y^1 + Y^2, t)$ is defined as the convolution

$$P(Y^1 + Y^2, t) = P(Y^1, t) * P(Y^2, t). \tag{22}$$

Using Equation (20), we construct all possible outcomes after two successive single catalyst subsystem simulations and compare the resulting probability $P(\mathbf{Y}, t)$ with

$P(\mathbf{X}, t)$ of the full system in Equation (A3). An exemplary derivation of the equality for $P\big(\mathbf{X} = (0,1)^T, t\big)$ and the initial state $\mathbf{X} = (0,0)^T$ is detailed in Appendix A. Similar derivations can be undertaken for all six entries of the full CME solution. Figure 2C visualizes the equivalent analytic solutions $P(\mathbf{X}, t)$ and $P(\mathbf{Y}, t)$ obtained for the full system (left) and SiCaSMA (right), respectively.

3.3. Applying SiCaSMA to Larger Networks via the SSA

Next, we apply SiCaSMA to a larger network by implementing it into an SSA approach. Therefore, we consider the methylation system of Figure 2 and arbitrary n_S, n_D, n_P. Pseudocode of the conventional SSA for the full system is depicted in Algorithm 2. Applying SiCaSMA (Algorithm 3), we initialize the number of methylated sites $Y^m = X_{init,2}$. For the n_P single catalyst systems, the state of the catalyst Y^p is set to $Y^p = X_{init,1}$, and sample paths are generated via the conventional SSA with a single catalyst molecule, $n_P = 1$. Finally, the state vector $\mathbf{X}(t)$, which is a two-dimensional vector with $X_1(t) = \sum_{i=1}^{n_P} Y^p(t)$ and $X_2(t) = Y^m(t)$, is returned. In summary, SiCaSMA SSA runs the conventional version of the SSA for each single catalyst subsystem and eventually combines the obtained solutions by treating the methylation state as a global variable throughout all single catalyst simulations and concatenating the catalyst states using a sum.

Algorithm 2 SSA DNA methylation system ($\mathbf{X}_{init}, t_f, n_P, n_S$)

```
Initialize $\mathbf{X} = \mathbf{X}_{init}$ and set $t = 0$
while $t < t_f$ do
    Calculate $A = (a_1, a_{-1}, a_m)$ and $a_{sum} = \sum_{j=1,2,3} A_j$
    Draw independent random numbers $\xi_1$ and $\xi_2$ uniformly from $(0,1)$
    Set $i$ to be the smallest integer satisfying $\sum_{j=1,\ldots,i} a_j > \xi_1 a_{sum}$
    Update $t = t + \frac{\ln(\xi_2^{-1})}{a_{sum}}$
    if $t < t_f$ then
        Update $\mathbf{X} = \mathbf{X} + \boldsymbol{\nu}_i$
    end if
end while
return $\mathbf{X}$
```

Algorithm 3 Single Catalyst SSA DNA methylation system ($\mathbf{X}_{init}, t_f, n_P, n_S$)

```
Initialize $Y^m = X_{init,2}$
for $p = 1, \ldots, n_P$ do
    Initialize $\begin{pmatrix} Y^p \\ Y^m \end{pmatrix} = \begin{pmatrix} X_{init,1} \\ Y^m \end{pmatrix}$
    Update $\begin{pmatrix} Y^p \\ Y^m \end{pmatrix} = \mathrm{SSA}\left(\begin{pmatrix} Y^p \\ Y^m \end{pmatrix}, t_f, 1, n_S\right)$
end for
return $\mathbf{Y} = \begin{pmatrix} \sum_{p=1}^{n_P} Y^p \\ Y^m \end{pmatrix}$
```

A visual comparison of the outcome of the two algorithms is shown in Figure 3. Results in Figure 3A are obtained with $n_S = 10$, $n_D = 2$, $n_P = 10$ and reaction rates $k_1 = 2, k_{-1} = 1, k_m = \frac{1}{2}$. The estimated distributions are obtained using 10^5 sample paths for each of the simulation algorithms evaluated for $t = 3.0$. Similarly, Figure 3B presents the estimated distributions for an even larger system defined by $n_S = 100$, $n_D = 5$, $n_P = 100$ and the same reaction rates as in Figure 3A. It can be seen that the probability distributions estimated with the SSA of the full system (Figure 3 (left)) and of SiCaSMA (Figure 3 (right)) are nearly identical. This becomes especially remarkable when recapitulating the dimension of the corresponding CME, which defines a probability for every single possible state of the system. Only considering the different possibilities of binding 100 DNMT1 molecules to five DNA strands renders a CME solution nearly impossible, not to mention the second, independent combinatorial explosion introduced via the different possible methylation states of the DNA. A convergence of SSA results such as the one observed in Figure 3 can therefore not be taken for granted.

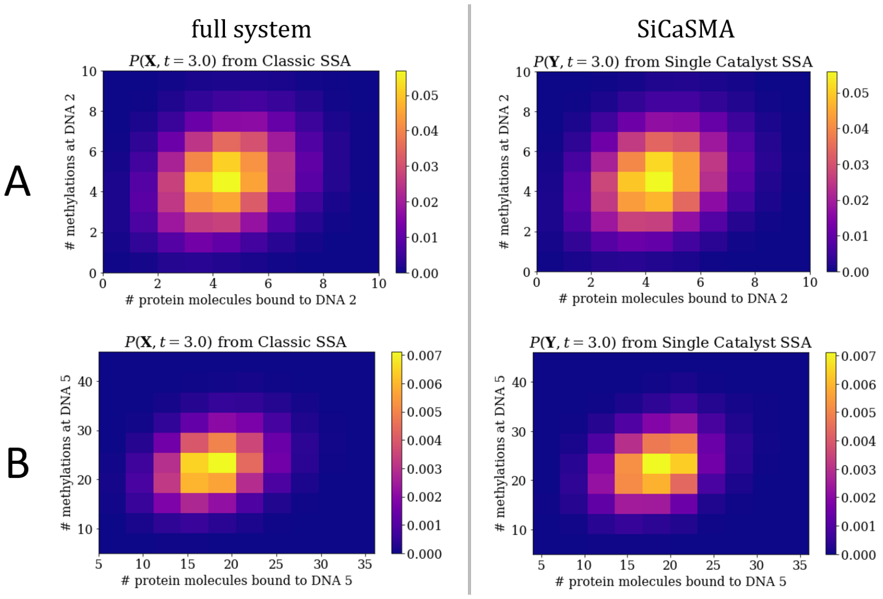


Figure 3. Application of SiCaSMA to a larger system via SSA. All distributions are obtained by 10^5 sample paths of the respective version of the SSA evaluated for $t_f = 3.0$ and parameters $k_1 = 2, k_{-1} = 1, k_m = \frac{1}{2}$. (**A**) Estimated probability distributions for the state of the second DNA molecule, which is defined by the number of DNMT1 molecules bound to it and the number of methylations. Results were obtained with $n_S = 10, n_D = 2$ and $n_P = 10$. (**B**) Analog results for DNA molecule 5 and an even larger system ($n_S = 100$, $n_D = 5$, $n_P = 100$).

4. Discussion

In this paper, we have introduced SiCaSMA, an alternative stochastic description for a class of catalytic systems in which catalysts act independently of each other on substrate molecules. Instead of considering the CME for the full system, SiCaSMA concatenates smaller subsystems which consist of all substrate molecules, but only one catalyst molecule each. The single catalyst subsystems are analyzed one after another. Hereby, the substrate state is inherited from actions of former subsystems, which is reflected in the initial conditions of the subsystem under consideration. Thus, the substrate state acts as a global variable. The states of the catalyst molecules in the subsystems are local variables, i.e., describe the state in the subsystems, and are re-initialized for each subsystem. By applying SiCaSMA, it is not necessary to characterize the state space of the full system. It is sufficient to perform calculations or simulations on the single catalyst subsystems, which usually have a much smaller state space. In the end, the state of the full system is reconstructed from the states of the individual single catalyst subsystems.

We have applied SiCaSMA to different systems. First, a simple reversible conversion reaction, in which a molecule can reversibly switch between two different states, was used to illustrate the approach. This system was described with simple first-order kinetics and represents an intuitive example for our approach. The partition into n individual subsystems is trivial here, but the example is well-suited to demonstrate the general idea of SiCaSMA. Second, we have considered DNMT1-mediated DNA methylation, which consists of reversible binding of DNMT1 to the DNA and processive methylation at different sites. For both systems and low numbers of variables, we have shown that SiCaSMA provides indeed an equivalent description of the underlying stochastic process. Since the dimensions of the single catalyst subsystems are smaller than those of the full systems, SiCaSMA can also be applied to calculate a solution of the full system when the full system is not feasible. However, reconstruction of the state of the full system also becomes more complicated in these cases, and a practically applicable general reconstruction scheme is still missing.

Beyond the specific methylation system discussed in the paper, we think that SiCaSMA can be applied to a broad range of biochemical reaction systems. Several literature models can be adopted to this form. Examples include mRNA transcription [19] and post-transcriptional modifications of proteins such as phosphorylation [20]. Moreover, we believe that it is also possible to extend the model class in several directions, e.g., regarding the number of different catalyst or substrate types and their numbers of different configurations. A more formal definition of the system class, which generalizes our results, is a challenging future task.

Author Contributions: Conceptualization, V.W. and N.E.R.; Methodology, V.W. and N.E.R.; Software, V.W.; Validation, V.W. and N.E.R.; Formal Analysis,V.W. and N.E.R.; Investigation, V.W. and N.E.R.; Resources, N.E.R.; Data Curation, V.W.; Writing—Original Draft Preparation, V.W. and N.E.R.; Writing—Review and Editing, V.W. and N.E.R.; Visualization, V.W. and N.E.R.; Supervision, N.E.R.; Project Administration, N.E.R.; Funding Acquisition, N.E.R. All authors have read and agreed to the published version of the manuscript.

Funding: Funded by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy—EXC 2075—390740016. We acknowledge the support by the Stuttgart Center for Simulation Science (SimTech). Funded by by the Deutsche Forschungsgemeinschaft grant RA1840/2-1 to N.R.

Data Availability Statement: All code created for this paper is accessible on https://fairdomhub.org/models/784 (accessed on 7 May 2021). We used the *python* language (Version 3.7.4) [21] with three additional packages: the symbolic computing package *sympy* for analytic calculations (Version 1.4) [22], the *numpy* package for numeric calculations (Version 1.15) [23] and *matplotlib* to create all figures (Version 3.1.1) [24].

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CME	Chemical Master Equation
SSA	Stochastic Simulation Algorithm
SiCaSMA	Single Catalyst Stochastic Modeling Approach
DNMT1	DNA methyltransferase 1

Appendix A. Proof of Equivalence of X and Y for the DNA Methylation Model

We perform an exemplary derivation of equivalence for $P\big(\mathbf{X} = (0,1)^T, t\big)$ and the initial state $\mathbf{X} = (0,0)^T$. Four additional algebraic identities are used during the derivation:

$$0 = \frac{1-\sqrt{5}}{4\sqrt{5}} + \frac{1}{5+\sqrt{5}} = \frac{5-\sqrt{5}}{10\sqrt{5}-30} + \frac{1}{5-\sqrt{5}} \tag{A1}$$

as well as

$$\frac{1-\sqrt{5}}{5+5\sqrt{5}} = -\frac{2}{5(3+\sqrt{5})}, \qquad \frac{\sqrt{5}-1}{20-10\sqrt{5}} = \frac{2}{5(\sqrt{5}-3)}. \tag{A2}$$

The relation between the states $\mathbf{X}$ and Y^i is given by the convolution, i.e.,

$$
\begin{aligned}
P\Big(\mathbf{X}=(0,0)^T,t\Big) &= P\Big((Y^1,Y^m)^T=(0,0)^T,t\Big)\,P\Big((Y^2,Y^m)^T=(0,0)^T,t\Big)\\
P\Big(\mathbf{X}=(1,0)^T,t\Big) &= P\Big((Y^1,Y^m)^T=(0,0)^T,t\Big)\,P\Big((Y^2,Y^m)^T=(1,0)^T,t\Big)\\
&\quad+P\Big((Y^1,Y^m)^T=(1,0)^T,t\Big)\,P\Big((Y^2,Y^m)^T=(0,0)^T,t\Big)\\
P\Big(\mathbf{X}=(2,0)^T,t\Big) &= P\Big((Y^1,Y^m)^T=(1,0)^T,t\Big)\,P\Big((Y^2,Y^m)^T=(1,0)^T,t\Big)\\
P\Big(\mathbf{X}=(0,1)^T,t\Big) &= P\Big((Y^1,Y^m)^T=(0,0)^T,t\Big)\,P\Big((Y^2,Y^m)^T=(0,1)^T,t\Big)\\
&\quad+P\Big((Y^1,Y^m)^T=(0,1)^T,t\Big)\,P\Big((Y^2,Y^m)^T=(0,1)^T,t\Big)\\
P\Big(\mathbf{X}=(1,1)^T,t\Big) &= P\Big((Y^1,Y^m)^T=(0,0)^T,t\Big)\,P\Big((Y^2,Y^m)^T=(1,1)^T,t\Big)\\
&\quad+P\Big((Y^1,Y^m)^T=(1,1)^T,t\Big)\,P\Big((Y^2,Y^m)^T=(0,1)^T,t\Big)\\
&\quad+P\Big((Y^1,Y^m)^T=(1,0)^T,t\Big)\,P\Big((Y^2,Y^m)^T=(0,1)^T,t\Big)\\
&\quad+P\Big((Y^1,Y^m)^T=(0,1)^T,t\Big)\,P\Big((Y^2,Y^m)^T=(1,1)^T,t\Big)\\
P\Big(\mathbf{X}=(2,1)^T,t\Big) &= P\Big((Y^1,Y^m)^T=(1,0)^T,t\Big)\,P\Big((Y^2,Y^m)^T=(1,1)^T,t\Big)\\
&\quad+P\Big((Y^1,Y^m)^T=(1,1)^T,t\Big)\,P\Big((Y^2,Y^m)^T=(1,1)^T,t\Big)
\end{aligned}
\tag{A3}
$$

Thus, for $P\big(\mathbf{X}=(0,1)^T,t\big)$ we get

$$
\begin{aligned}
&P\big(\mathbf{X}=(0,1)^T,t\big)\\
&\overset{(A3)}{=} P\big((Y^1,Y^m)^T=(0,0)^T,t\big)\,P\big((Y^2,Y^m)^T=(0,1)^T,t\big)\\
&+P\big((Y^1,Y^m)^T=(0,1)^T,t\big)\,P\big((Y^2,Y^m)^T=(0,1)^T,t\big)\\
&\overset{(17)}{=} e^{tS}_{11}\cdot e^{tS}_{31}+e^{tS}_{31}\cdot e^{tS}_{33}\\
&\overset{(19)}{=} \left(\frac{1}{2}+\frac{e^{-2t}}{2}+\frac{2c_-}{5+\sqrt{5}}-\frac{(5-\sqrt{5})c_+}{5\sqrt{5}-15}\right)\cdot\left(\frac{(1-\sqrt{5})c_-}{2\sqrt{5}}-\frac{2\sqrt{5}c_+}{5\sqrt{5}-5}+\frac{1}{2}+\frac{e^{-2t}}{2}\right)\\
&=\frac{1}{4}+\frac{e^{-2t}}{2}+\frac{e^{-4t}}{4}+\frac{(1-\sqrt{5})c_-}{4\sqrt{5}}+\frac{e^{-2t}(1-\sqrt{5})c_-}{4\sqrt{5}}+\frac{c_-}{5+\sqrt{5}}+\frac{e^{-2t}c_-}{5+\sqrt{5}}-\frac{2e^{-2t}c_+}{5-\sqrt{5}}\\
&-\frac{c_+}{5-\sqrt{5}}-\frac{(5-\sqrt{5})c_+}{10\sqrt{5}-30}-\frac{(5-\sqrt{5})e^{-2t}c_+}{10\sqrt{5}-30}+\frac{(1-\sqrt{5})c_-^2}{5+5\sqrt{5}}-\frac{2c_+c_-}{5}+\frac{(\sqrt{5}-1)c_+^2}{20+5\sqrt{5}}\\
&\overset{(A1)}{=} \frac{1}{4}+\frac{e^{-2t}}{2}+\frac{e^{-4t}}{4}+\frac{(1-\sqrt{5})c_-^2}{5+5\sqrt{5}}-\frac{2c_+c_-}{5}+\frac{(\sqrt{5}-1)c_+^2}{20+5\sqrt{5}}\\
&\overset{(18)}{=} \frac{1}{4}+\frac{e^{-2t}}{2}+\frac{e^{-4t}}{4}+\frac{(1-\sqrt{5})e^{-\sqrt{5}t-3t}}{5+5\sqrt{5}}-\frac{2e^{-3t}}{5}+\frac{(\sqrt{5}-1)e^{\sqrt{5}t-3t}}{20+5\sqrt{5}}\\
&\overset{(A2)}{=} \frac{1}{4}+\frac{e^{-2t}}{2}-\frac{2e^{-3t}}{5}+\frac{e^{-4t}}{4}-\frac{2e^{-\sqrt{5}t-3t}}{5(3+\sqrt{5})}+\frac{2e^{\sqrt{5}t-3t}}{5(\sqrt{5}-3)}
\end{aligned}
\tag{A4}
$$

This expression is indeed equivalent to the fourth entry of Equation (15).

References

1. Higham, D.J. Modeling and simulating chemical reactions. *SIAM Rev.* **2008**, *50*, 347–368. [CrossRef]
2. Wilkinson, D.J. *Stochastic Modelling for Systems Biology*; CRC Press: Boca Raton, FL, USA, 2018.
3. Schnoerr, D.; Sanguinetti, G.; Grima, R. Approximation and inference methods for stochastic biochemical kinetics—A tutorial review. *J. Phys. A Math. Theor.* **2017**, *50*, 093001. [CrossRef]
4. Van Kampen, N. A power series expansion of the Master equation. *Can. J. Phys.* **1961**, *39*, 551–567. [CrossRef]
5. Kampen, N.V. *Stochastic Processes in Physics and Chemistry*; Elsevier: North Holland, The Netherlands, 2007.
6. Gillespie, D.T. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.* **1976**, *22*, 403–434. [CrossRef]
7. Smadbeck, P.; Kaznessis, Y.N. A closure scheme for chemical master equations. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 14261–14265. [CrossRef]
8. Engblom, S. Computing the moments of high dimensional solutions of the master equation. *Appl. Math. Comput.* **2006**, *180*, 498–515. [CrossRef]

9. Munsky, B.; Khammash, M. The finite state projection algorithm for the solution of the chemical master equation. *J. Chem. Phys.* **2006**, *124*, 044104. [CrossRef]
10. Wolf, V.; Goel, R.; Mateescu, M.; Henzinger, T.A. Solving the chemical master equation using sliding windows. *BMC Syst. Biol.* **2010**, *4*, 1–19. [CrossRef]
11. Kazeev, V.; Khammash, M.; Nip, M.; Schwab, C. Direct solution of the chemical master equation using quantized tensor trains. *PLoS Comput. Biol.* **2014**, *10*, e1003359. [CrossRef]
12. Elowitz, M.B.; Levine, A.J.; Siggia, E.D.; Swain, P.S. Stochastic Gene Expression in a Single Cell. *Science* **2007**, *297*, 1183–1186. [CrossRef]
13. Paulsson, J. Models of stochastic gene expression. *Phys. Life Rev.* **2005**, *2*, 157–175. [CrossRef]
14. Bressloff, P.C.; Newby, J.M. Stochastic models of intracellular transport. *Rev. Mod. Phys.* **2013**, *85*, 135–196. [CrossRef]
15. Birtwistle, M.R.; Rauch, J.; Kiyatkin, A.; Aksamitiene, E.; Dobrzyński, M.; Hoek, J.B.; Kolch, W.; Ogunnaike, B.A.; Kholodenko, B.N. Emergence of bimodal cell population responses from the interplay between analog single-cell signaling and protein expression noise. *BMC Syst. Biol.* **2012**, *6*. [CrossRef]
16. Smolen, P.; Baxter, D.; Byrne, J. Bistable MAP kinase activity: A plausible mechanism contributing to maintenance of late long-term potentiation. *Am. J. Physiol. Cell. Physiol.* **2008**, *294*, C503–C515. [CrossRef] [PubMed]
17. Ueda, M.; Shibata, T. Stochastic Signal Processing and Transduction in Chemotactic Response of Eukaryotic Cells. *Biophys. J.* **2007**, *93*, 11–20. [CrossRef]
18. Adam, S.; Anteneh, H.; Hornisch, M.; Wagner, V.; Lu, J.; Radde, N.E.; Bashtrykov, P.; Song, J.; Jeltsch, A. DNA sequence-dependent activity and base flipping mechanisms of DNMT1 regulate genome-wide DNA methylation. *Nat. Commun.* **2020**, *11*, 3723. [CrossRef] [PubMed]
19. Voliotis, M.; Cohen, N.; Molina-París, C.; Liverpool, T.B. Fluctuations, pauses, and backtracking in DNA transcription. *Biophys. J.* **2008**, *94*, 334–348. [CrossRef] [PubMed]
20. Salazar, C.; Höfer, T. Multisite protein phosphorylation–from molecular mechanisms to kinetic models. *FEBS J.* **2009**, *276*, 3177–3198. [CrossRef]
21. Van Rossum, G.; Drake, F.L., Jr. *Python Reference Manual*; Centrum voor Wiskunde en Informatica Amsterdam: Amsterdam, The Netherlands, 1995.
22. Meurer, A.; Smith, C.P.; Paprocki, M.; Čertík, O.; Kirpichev, S.B.; Rocklin, M.; Kumar, A.; Ivanov, S.; Moore, J.K.; Singh, S.; et al. SymPy: Symbolic computing in Python. *PeerJ Comput. Sci.* **2017**, *3*, e103. [CrossRef]
23. Harris, C.R.; Millman, K.J.; van der Walt, S.J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N.J.; et al. Array programming with NumPy. *Nature* **2020**, *585*, 357–362. [CrossRef] [PubMed]
24. Hunter, J.D. Matplotlib: A 2D graphics environment. *IEEE Ann. Hist. Comput.* **2007**, *9*, 90–95. [CrossRef]