On the Properties of the Reaction Counts Chemical Master Equation

Vikram Sunkara [1,2]

[1] Computational Medicine, Zuse Institute Berlin, 14195 Berlin, Germany; sunkara@mi.fu-berlin.de
[2] Department of Mathematics and Computer Science, Freie Universität Berlin, 14195 Berlin, Germany

check for
updates

Abstract: The reaction counts chemical master equation (CME) is a high-dimensional variant of the classical population counts CME. In the reaction counts CME setting, we count the reactions which have fired over time rather than monitoring the population state over time. Since a reaction either fires or not, the reaction counts CME transitions are only forward stepping. Typically there are more reactions in a system than species, this results in the reaction counts CME being higher in dimension, but simpler in dynamics. In this work, we revisit the reaction counts CME framework and its key theoretical results. Then we will extend the theory by exploiting the reactions counts' forward stepping feature, by decomposing the state space into independent continuous-time Markov chains (CTMC). We extend the reaction counts CME theory to derive analytical forms and estimates for the CTMC decomposition of the CME. This new theory gives new insights into solving hitting times-, rare events-, and a priori domain construction problems.

Keywords: chemical master equation; jump continuous-time Markov chains; reaction counts

1. Introduction

Continuous-time Markov chains (CTMC) accurately capture the dynamics of a broad range of biochemical reaction systems. The CTMCs come in two flavours: The discrete state space, the chain transitions from one state to another, or the continuous state space, the state transitions are smooth yet non-differentiable. Each type has an intuitive interpretation. Discrete state spaces describe the population or counts of chemicals in a system, while continuous states spaces describe concentrations of chemicals in a system [1,2]. It is tempting to generalise concentrations to be simply scaled populations, which would mean that all systems could be mapped into the continuous setting and solved there. However, this is not the case mathematically. It was shown that for large populations, the discrete CTMC and continuous CTMC were equivalent [3,4]. If the large population assumption is not satisfied, mathematically, the discrete state space system has to be studied in its own right.

The probability distribution of a CTMC over the discrete state space in the biochemical context is found by solving the chemical master equation (CME). We first present the CME and then describe its components. The formula of the CME is written as follows,

$$\frac{\partial p(Z(t) = \mathbf{z})}{\partial t} = \sum_{n=1}^{N_r} a_n(\mathbf{z} - \nu_n)\, p(Z(t) = \mathbf{z} - \nu_n) - \sum_{n=1}^{N_r} a_n(\mathbf{z})\, p(Z(t) = \mathbf{z}). \tag{1}$$

In the equation above, $Z(t)$ is a CTMC over the state space $\Omega \subset \mathbb{N}_0^{N_s}$. Each state is a vector of $N_s \in \mathbb{N}$ non-negative integers describing the population of the species at time $t \geq 0$. Starting from the left-hand side, the equation states that the change in probability of being in a state $z \in \Omega$ at time t is given by the sum of two terms. The first term describes the probability flowing into the state $\mathbf{z}$ and the second term describes the probability flowing out of the state $\mathbf{z}$. There are $N_r \in \mathbb{N}$ reactions perturbing

the population; the terms v_n are referred to as stoichiometric vectors, they capture the net change in population if the nth reaction fires. Hence, subtracting the stoichiometric vectors v_n from **z** gives us the set of all previous states which might transition into **z** and "contribute" probability. The last piece of information in the equation are the rates at which reactions fire, these are encapsulated in the propensity functions $\{a_n : \Omega \rightarrow \mathbb{R}_+, \text{ for } n = 1, \ldots, N_r\}$. In summary, the CME describes the change in probability of observing a CTMC in a certain state, which is equal to the sum of probability transitioning into the state, minus the sum of probability leaving the state.

Even though the dynamics of the derivative are fairly simple, solving the CME is a hard problem [5–12]. A special case of the CME which is truly unyielding to approximation is when a system is close to its population boundary (for example, close to zero). In this scenario, none of the state transitions around a state near the boundary can be summarised by the average transitions out of the state. This impedes most approximation methods, and even if a method was successful, its computational effort would be equivalent to that of solving the Finite State Projection for the same error. An approach which was proposed in various different contexts to simplify dynamics, was to study the number of reactions fired rather than the population counts [6,13–17]. Because reactions cannot "un-fire", the state transitions are only moving forward. Furthermore, given that the system's starting population is known, the species count in any state can be reconstructed from the reaction counts.

In this work, we will reintroduce the reaction counts variant of the CME. We will then explore the theoretical and structural results which emerge from its forward stepping process. Then, we will show how the reaction counts CME can be used to construct solutions to the classical CME. Lastly, we decompose the state space of the CME into independent CTMCs and give analytical forms and estimates to calculate their probabilities. The purpose of this work is to gain intuition and explore structural properties of the CME, hence, the motivation is more theory rather than the application of the CME. In light of that, we omit a discussion of numerical methods in this paper and envisage this aspect in future research.

2. Reaction Counts CME

2.1. Formulation

We denote the state space of the reaction counts CME by $\Lambda \subset \mathbb{N}_0^{N_r}$. Each state is a vector of non-negative integers of length N_r. Each element of the vector represents the number of times its corresponding reaction has fired. In the CME setting, the change in state after a reaction fires is given by the stoichiometric vector, which we denoted earlier by v_n, for $n = \{1, \ldots, N_r\}$. This vector quantifies the net change in populations after the reaction has fired. Analogously, in the reaction counts CME setting, a change in state indicates that a reaction has fired, and its corresponding reaction count is incremented by one. Therefore, the stoichiometric vector for the nth reaction is simply the identity vector $\mathbf{1}_n$; the vector is zero in all but the nth position, where it is one.

We bridge the species counts CME and the reaction counts CME with the mapping $\Gamma : \Omega \times \Lambda \rightarrow \Omega$, that is, for $(x_0, r) \in \Omega \times \Lambda$,

$$
\begin{aligned}
\Gamma(x_0, r) \quad &:= \quad x_0 + [v_1, v_2, \ldots, v_{N_r}]\, r, \\
&= \quad x_0 + \sum_{n=1}^{N_r} v_n\, r_n.
\end{aligned}
\tag{2}
$$

The mapping above links the reaction counts CME to the species counts CME by stating that, given the starting state x_0 and the reactions which have fired are known, then the current state in the species counts is the starting species state plus the sum of stochiometries of all the reactions which have fired. It is important to note that the map Γ is injective into the species state space Ω. In most cases, Λ is of

higher dimension than Ω, hence, Γ is seldom bijective. For our purpose, we only need the pull back map of Γ. We denote the pull back map as $\Gamma^{-1} : \Omega \times \Omega \to \mathcal{P}(\Lambda)$, where for $(x_0, x) \in \Omega \times \Omega$,

$$\Gamma^{-1}(x_0, x) := \{r \in \Lambda : \Gamma(x_0, r) = x\}. \tag{3}$$

With the mapping between Λ and Ω established, we can now inherit the propensity function from the species counts over to the reaction counts. For $n = 1, \ldots, N_r$, the reaction counts propensity of the nth reaction is given by,

$$\alpha_n(x_0, r) := \begin{cases} a_n(\Gamma(x_0, r)) & \text{if } \Gamma(x_0, r) \in \Omega, \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

With the state space, stoichiometry, and the propensities established in the reaction counts setting, we are ready to formulate the reaction counts based Kurtz process. For $x_0 \in \Omega$, let $(R_{x_0}(t))_{t \geq 0}$ be the reaction counts Kurtz process,

$$R_{x_0}(t) := \sum_{n=1}^{N_r} \mathbf{1}_n \, \mathcal{P} \left(\int_0^t \alpha_n(x_0, R_{x_0}(s)) ds \right). \tag{5}$$

The corresponding reaction counts CME for the process above is given by,

$$\frac{dp(R_{x_0}(t) = r)}{dt} = \sum_{n=1}^{N_r} \alpha_n(x_0, r - \mathbf{1}_n) \, p(R_{x_0}(t) = r - \mathbf{1}_n) - \sum_{n=1}^{N_r} \alpha_n(x_0, r) \, p(R_{x_0}(t) = r), \tag{6}$$

for all $r \in \Lambda$. The initial condition for the reaction counts CME is a points mass on the origin,

$$p(R_{x_0}(0) = r) = \begin{cases} 1 & \text{if } r = \mathbf{0}, \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

Upon first glance, the CME of the reaction counts is simpler in its complexity compared to its species counter part. In the reaction case, the processes only move forward, which gives rise to simple forward propagating dynamics. Furthermore, given solutions to the reaction counts CME exist, with the simple application of the push forward measure, we would obtain:

$$p(X(t) = x) = \sum_{r \in \Gamma^{-1}(X(0), x)} p(R_{X(0)}(t) = r). \tag{8}$$

The relationship given above is the critical motivation for studying the reaction counts CME. In principle, if structures and results could be attained in the reaction counts setting, then by using Γ, these results can be mapped into species setting. The simplest example of this is the proof for the existence of solutions of the CME for finite time. We now show that the reaction counts CME has analytical solutions, then using the push forward measure in Equation (8), we prove that the solutions of the CME exist for a broad range of propensity functions.

2.2. Analytical Solutions of the Reaction Counts CME

Firstly, we will establish the notations and assumptions needed to present the results. For brevity, we denote the sum of all propensities α_n as:

$$\alpha(\cdot, r) := \sum_{n=1}^{N_r} \alpha_n(\cdot, r). \tag{9}$$

The reaction counts CME in Equation (6) is a difference differential equation, hence, by arranging the probability states in a vector $u(t) := (p(R_{x_0}(t) = r))_{r \in \Lambda}$, we reduce solving the reaction counts CME to solving the ODE,

$$\frac{du(t)}{dt} = A^* u(t).$$

For $r, l \in \Lambda$, the matrix A^* has the properties: $[A^*]_{r,r} \leq 0$, $[A^*]_{r,l} \geq 0$, with $\sum_{l \in \Lambda}[A^*]_{l,r} = 0$. We will prove that A^* is a generator of a one-parameter semigroup. Before this, let us consider a simple two reaction system to gain some intuition into the different components of the reaction counts setting.

Example 1. *Let us consider the birth-death process in the context of the reaction counts CME. The birth-death process in the species count is given by,*

$$Z(t) = x_0 + \mathcal{P}\left(\int_0^t c_b \, ds\right) - \mathcal{P}\left(\int_0^t c_d \, Z(s) \, ds\right),$$

where $x_0 \in \mathbb{N}_0$ is the starting population, and the birth and death rates are denoted by c_b and c_d, respectively. We can translate this process into the reaction counts setting using the mapping,

$$\Gamma_{x_0}(r_1, r_2) = x_0 + r_1 - r_2.$$

We restrict the reaction counts state space $\Lambda \subset \mathbb{N}_0 \times \mathbb{N}_0$ to only contain states which yield non-negative values by applying Γ_{x_0}. Substituting the mapping Γ_{x_0} and the process $Z(t)$ into Equations (2) to (5) leads to the reaction counts birth-death process formulation:

$$R_{x_0}(t) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}^T \mathcal{P}\left(\int_0^t c_b \, ds\right) + \begin{pmatrix} 0 \\ 1 \end{pmatrix}^T \mathcal{P}\left(\int_0^t c_d \, \Gamma_{x_0}\left((R_{x_0}(s))_1, (R_{x_0}(s))_2\right) \, ds\right). \quad (10)$$

The relationship between the reaction count state space and the species count state space is visualised in Figure 1A. The colours in the figure show how the species count state space partitions the reaction count state space. To demonstrate the forward moving structure of the reaction counts CME, we derive the generator for the first nine states of its state space:

$$\frac{du(t)}{dt} = A^* u(t),$$

where,

$$u(t) = \left(\frac{dp(R_{x_0}(t) = (0,0))}{dt}, \frac{dp(R_{x_0}(t) = (1,0))}{dt}, \frac{dp(R_{x_0}(t) = (0,1))}{dt}, \cdots\right)^T,$$

$$A^* = \begin{bmatrix} \overset{(0,0)}{-c_b - c_d x_0} & \overset{(1,0)}{0} & \overset{(0,1)}{0} & \overset{(2,0)}{0} & \overset{(1,1)}{0} & \overset{(0,2)}{0} & \cdots \\ c_b & -c_b - c_d(x_0+1) & 0 & 0 & 0 & 0 & \cdots \\ c_d x_0 & 0 & -c_b - c_d(x_0-1) & 0 & 0 & 0 & \cdots \\ 0 & c_b & 0 & -c_b - c_d(x_0+2) & 0 & 0 & \cdots \\ 0 & c_d(x_0+1) & c_b & 0 & -c_b - c_d x_0 & 0 & \cdots \\ 0 & 0 & c_d(x_0-1) & 0 & 0 & -c_b - c_d(x_0-2) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

The initial condition is the identity vector with the total probability mass on state $(0,0)$. We can observe that due to the forward moving nature of the reaction counts setting, A^* is a lower diagonal matrix. In Figure 1B,C, we present a birth-death process with initial state $x_0 = 0$, birth rate $c_b = 1.0$, and death rate $c_d = 0.1$. Specifically, we show the probability distributions of the first three states of that process which correspond to state 1 in the species count setting. In Figure 1B, it is shown that the probability distribution of being in state 1 in the species count CME setting bounds the corresponding reaction counts CME distributions. In Figure 1C, we

progressively add the first three reaction counts states corresponding to the state 1, showing that the reaction counts CME approximates the species counts CME from below.
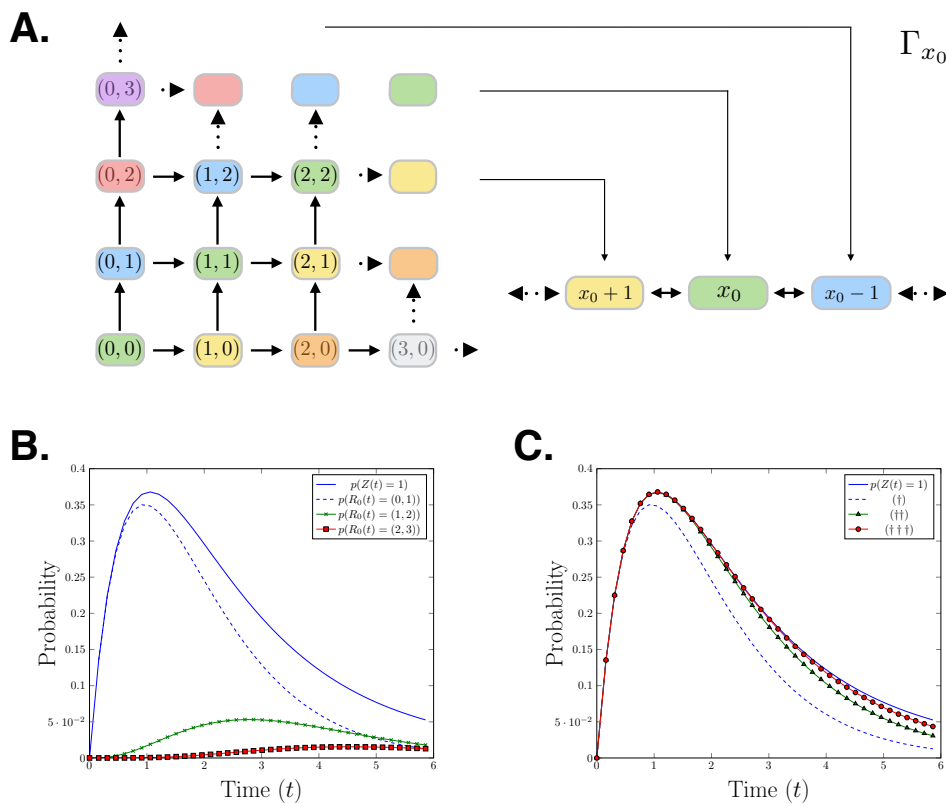

Figure 1. (**A**) Cartoon showing the mapping of the reactions counts birth-death process to the species count birth-death process. (**B**) An evaluation of the birth-death process for parameters $(x_0 = 0, c_b = 1.0, c_d = 0.1)$. The plot shows: The distribution of the species having population one over time $p(Z(t) = 1)$; and the distributions of reaction $(1,0), (1,2)$, and $(2,3)$ firing at time t, in the reaction counts setting. (**C**) (†) $= p(R_0(t) = (0,1))$, (††) $= p(R_0(t) = (0,1)) + p(R_0(t) = (1,2))$, (†††) $= p(R_0(t) = (0,1)) + p(R_0(t) = (1,2)) + p(R_0(t) = (2,3))$. The distribution of the species count chemical master equation (CME) in state 1 at t is approached from below by the sum of the probabilities of the reactions firing which result in being in state 1.

Theorem 1 (Sunkara 09 ([15])). *Given a system with N_s species and N_r reactions at a starting state $x_0 \in \Omega$. If A_r is the generator for the reaction counts CME (6), then:*

1. *There exists a permutation matrix P such that $P^T A^* P$ is lower triangular;*
2. *The spectrum of A^* is the set:*

$$spec(A^*) = \left\{ -\sum_{n=1}^{N_r} \alpha_n(x_0, r) : r \in \Lambda \right\}. \tag{11}$$

The property of the reaction counts setting, that the process only moves forward, aids in showing that the generator could be rewritten as a lower triangular matrix. We know that the spectrum of a lower triangular matrix are the diagonal elements of the matrix. In our case, the diagonal elements are simply the negative sum of the outgoing propensities. This then reveals that the spectrum of the generator A^* is in the negative real numbers. Then by the spectral mapping theorem, solutions to

the reaction counts CME exist. We can further exploit this "forward stepping" structure to write the analytical solutions for the probability at each state in Λ.

Proposition 1 (Sunkara 09 ([15])). *Given a system with N_s species and N_r reactions at a starting state $x_0 \in \Omega$. Let $C \in (0,1]$. If $(R_{x_0}(t))_{t \geq 0}$ is the reaction counts Kurtz process (5) with $\Lambda \neq \varnothing$, then the solution to (6) at $t > 0$ is given by:*

$$p(R_{x_0}(t) = \mathbf{0}) = e^{-\alpha(x_0, r) \, t} \, C, \tag{12}$$

for $r = \mathbf{0}$, and,

$$p(R_{x_0}(t) = r) = \sum_{n=1}^{N_r} \left(\alpha_n(x_0, r - \mathbf{1}_n) \int_0^t e^{-\alpha(x_0, r)\,(t-s)} \, p(R_{x_0}(s) = r - \mathbf{1}_n) ds \right), \tag{13}$$

for $r \in \Lambda \backslash \mathbf{0}$.

Combining Proposition 1 with the push forward measure in Equation (8), we see that at finite time t the solutions to the CME (1) exist. Unfortunately, since the reaction counts CME cannot reach a stationary state (in cases where the state space Λ is not finite), the proof for the existence of stationary solutions for a generalised CME is still an open problem. Now that the solutions of the reaction counts CME have been established, we can probe further into its properties. Firstly, we explore how the state space of the reaction count CME has a naturally hierarchical partitioning, and how this partitioning can be used to build a sequence of approximate sub-processes.

3. Partitioning the State Space

Definition 1. *For $m \in \mathbb{N}$, we define,*

$$\nabla_m := \{r \in \Lambda : \|r\|_1 = m\},$$

to be the set of all states which are reachable by m steps from the origin $\mathbf{0}$. The state space Λ naturally partitions into non-intersecting subsets,

$$\Lambda = \nabla_0 \cup \nabla_1 \cup \nabla_2 \cup \cdots.$$

If we consider truncating the state space progressively, we find that this is equivalent to the Finite State Projection with an N-step domain expander [9,18].

Lemma 1. *For $x_0 \in \Omega$, let $(R_{x_0}(t))_{t \geq 0}$ be the reaction counts Kurtz process (5). Then for any $m \in \mathbb{N}$,*

$$\sum_{r_1 \in \nabla_m} \sum_{n=1}^{N_r} \alpha_n(x_0, r_1) \int_0^t p(R_{x_0}(s) = r_1) \, ds = \sum_{\tilde{m}=m+1}^{\infty} \sum_{r_2 \in \nabla_{\tilde{m}}} p(R_{x_0}(t)) = r_2). \tag{14}$$

Proof. Fix $m \in \mathbb{N}$. Let $(R_{x_0, \nabla_m}(t))_{t \geq 0}$ be the same process as $(R_{x_0}(t))_{t \geq 0}$, but with the restriction that all propensities of states $r \in \cup_{\tilde{m}=m+1}^{\infty} \nabla_{\tilde{m}}$ are zero. Since in the reaction counts case, states only transition forward, we have that,

$$p(R_{x_0}(t) = r) = p(R_{x_0, \nabla_m}(t) = r), \text{ for all } r \in \cup_{\tilde{m}=0}^{m} \nabla_{\tilde{m}}. \tag{15}$$

That is, the probability of both processes are the same for states leading up to the states in ∇_m. However, since the process $R_{x_0, \nabla_m}(t)$ does not evolve past ∇_m, by the conservation of probability we have that,

$$\sum_{\tilde{m}=0}^{m} \sum_{r \in \nabla_{\tilde{m}}} p(R_{x_0, \nabla_m}(t) = r) + \sum_{r_1 \in \nabla_m} \sum_{n=1}^{N_r} \alpha_n(x_0, r_1) \int_0^t p(R_{x_0, \nabla_m}(s) = r_1)\, ds = 1. \tag{16}$$

Substituting in Equation (15) for states which appear before ∇_{m+1} reduces the above expression to,

$$\sum_{\tilde{m}=0}^{m} \sum_{r \in \nabla_{\tilde{m}}} p(R_{x_0}(t) = r) + \sum_{r_1 \in \nabla_m} \sum_{n=1}^{N_r} \alpha_n(x_0, r_1) \int_0^t p(R_{x_0}(s) = r_1)\, ds = 1. \tag{17}$$

By the conservation of probability,

$$\sum_{\tilde{m}=0}^{\infty} \sum_{r_2 \in \nabla_{\tilde{m}}} p(R_{x_0}(t) = r_2) = 1,$$

hence, substituting this into the right-hand side of Equation (17) and subtracting the like terms gives,

$$\sum_{r_1 \in \nabla_m} \sum_{n=1}^{N_r} \alpha_n(x_0, r_1) \int_0^t p(R_{x_0}(s) = r_1)\, ds = \sum_{\tilde{m}=m+1}^{\infty} \sum_{r_2 \in \nabla_{\tilde{m}}} p(R_{x_0}(t) = r_2).$$

$\square$

Lemma 1 is an alternative formulation of the principle behind the Finite State Projection method. We extend on this result and can show that for a desired error $\varepsilon \in (0, 1)$, there exists a subset of the state space Λ which will produce an approximation with the desired error.

Theorem 2. *For $x_0 \in \Omega$, let $(R_{x_0}(t))_{t \geq 0}$ be the reaction counts Kurtz process (5). For $\varepsilon > 0$ there exists $m \in \mathbb{N}$ such that,*

$$\sum_{\tilde{m}=m+1}^{\infty} \sum_{r \in \nabla_{\tilde{m}}} p(R_{x_0}(t) = r) < \varepsilon.$$

Proof. The proof is an extension of Lemma 1. We define,

$$\phi_m := \sum_{r_1 \in \nabla_m} \sum_{n=1}^{N_r} \alpha_n(x_0, r_1) \int_0^t p(R_{x_0}(s) = r_1)\, ds,$$

and,

$$\psi_{m+1} := \sum_{\tilde{m}=m+1}^{\infty} \sum_{r_2 \in \nabla_{\tilde{m}}} p(R_{x_0}(t) = r_2),$$

the left- and right-hand terms of Equation (14). Then for $m \in \mathbb{N}$, Lemma 1 can be reformulated to state,

$$\phi_m = \psi_{m+1};$$

separating the ∇_{m+1} states from ψ_{m+1} gives us,

$$= \sum_{r \in \nabla_{m+1}} p(R_{x_0}(t) = r) + \underbrace{\sum_{\tilde{m}=m+2}^{\infty} \sum_{r_2 \in \nabla_{\tilde{m}}} p(R_{x_0}(t) = r_2)}_{:=\psi_{m+2}},$$

$$= \sum_{r \in \nabla_{m+1}} p(R_{x_0}(t) = r) + \psi_{m+2}.$$

Combining the last and first term in the equalities gives us,

$$\psi_{m+1} = \sum_{r \in \nabla_{m+1}} p(R_{x_0}(t) = r) + \psi_{m+2}.$$

Rearranging the equality reduces to, $\psi_{m+2} - \psi_{m+1} < 0$. We have shown that $\psi_\bullet$ is monotonically decreasing. By definition, $\psi_\bullet$ has the properties that $\psi_1 = 1.0 - p(R_{x_0}(t) = \mathbf{0})$ and $\lim_{m \to \infty} \psi_m = 0$. Hence, there exists an m such that $\psi_m < \varepsilon$. $\square$

Theorem 2 is an alternative proof for the fact that, when studying transient dynamics (finite time) of the CME for an arbitrary precision, one can always find a finite state space to project the CME onto. The key structure that we used to prove this is that the state space of reaction counts has a natural partitioning and that the probability only flows forward. This gave us the monotonicity needed to prove that state space truncation of the CME is well-posed. In the next section, we further decompose each state in the reaction counts state space into paths over the state space.

4. Paths

Definition 2. *For $m \in \{2, 3, \ldots\}$, a vector $g = (g_1, \ldots, g_m) \in \Lambda^m$ is said to be an admissible path if for every index $i \in \{2, \ldots, m\}$ there exists an $n \in \{1, \ldots, N_r\}$ such that $g_i = g_{i-1} + \mathbf{1}_n$, where $\mathbf{1}_n$ is the identity vector with one in the nth position.*

Definition 3. *For $m \in \{2, 3, \ldots\}$, we denote the set of all admissible paths of length m by,*

$$\gamma_m := \{g \in \Lambda^m : g \text{ is admissible }\}$$

Definition 4. *For a point $r \in \Lambda$, we denote the set of all admissible paths of length $\|r\|_1 + 1$ which start at the origin $\mathbf{0}$ and end at the state r by,*

$$G(r) := \{g \in \gamma_{\|r\|_1 + 1} : g_1 = \mathbf{0} \text{ and } g_{\|r\|_1 + 1} = r\}.$$

4.1. Path Chains

We define a Markov chain over a path in the reaction counts state space. The chain must be such that it mimics a "realisation" over the reaction state space. Using Proposition 1, we define a path chain over an admissible path in the reaction state space. It is important to note that a path chain is a continuous-time Markov chain (CTMC) [19], where all reactions which transition the state off the path of the chain are accrued into a sink state. For brevity, we simply remove this component and only describe the states of interest (see Figure 2). Also, given all chains are from the same reaction counts setting, we omit explicitly stating x_0 in the propensity functions. We now define a path chain.
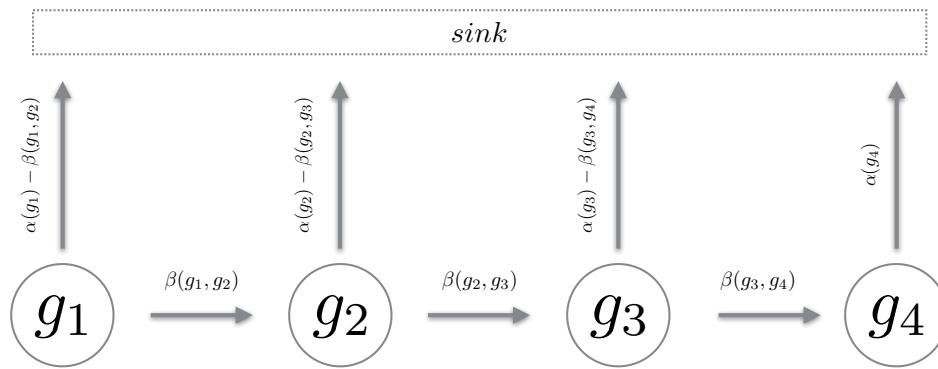
Figure 2. The cartoon above depicts the generator of a path chain. With $\alpha(\cdot)$ being the total outward propensity and $\beta(\cdot, \cdot)$ the propensity to transition to the next state in the chain. All reactions leading away from the chain are directed into the sink state.

Definition 5. *For $m \in \mathbb{N}$, $g \in \gamma_m$, and $C \in (0,1]$, we define a path chain to be the stochastic process $(X_{g,C}(t))_{t \geq 0}$, with state space $\{g \cup \{\text{sink}\}\}$ and the probability distribution given by,*

$$p(X_{g,C}(t) = g_1) := e^{-\alpha(g_1)\,t} C, \tag{18}$$

$$p(X_{g,C}(t) = g_k) := \beta(g_{k-1}, g_k) \int_0^t e^{-\alpha(g_k)\,(t-s)} \, p(X_{g,C}(s) = g_{k-1}) \, ds, \tag{19}$$

for $k \in \{2, \ldots, m\}$. Then to conserve probability,

$$p(X_{g,C}(t) = g_k) = \text{sink}) := 1.0 - \sum_{g_k \in g} p(X_{g,C}(t) = g_k).$$

The propensity function α is inherited from Equation (9) and $\beta(g_k, g_{k+1}) := \alpha_n(g_k)$, where n is the reaction index which transitions the state g_i to g_{i+1}.

Before stating the new proposition, let us recall the decompositions performed so far. Firstly, we showed that every state in the species state space can be decomposed into multiple reaction states in the reaction count setting. That is, a population state can be decomposed into all the different ways in which it can be visited. We then ventured further and showed that a path on the species counts state space corresponds to a path with only forward stepping transitions in the reaction counts state space. Now, we extend this further by showing that any state in the reaction counts state space can be decomposed into the sum of all independent path chains which start at the origin and end in that state.

Proposition 2. *For $x_0 \in \Omega$, let $(R_{x_0}(t))_{t \geq 0}$ be the Reaction counts Kurtz process (5).*
For $r \in \Lambda \backslash \mathbf{0}$, let
$$G(r) := \{g \in \gamma_{\|r\|_1 + 1} \; : \; g_1 = \mathbf{0} \text{ and } g_{\|r\|_1 + 1} = r\},$$
be the set of all admissible paths of length $\|r\|_1 + 1$ which start at $\mathbf{0}$ and ending at r (Definition 4).
Then,
$$p(R_{x_0}(t) = r) = \sum_{g \in G(r)} p(X_{g,1}(t) = r). \tag{20}$$

Proof. The Proposition will be proved using mathematical induction. Firstly, we show the base case; that every state which is reachable by one step from the origin satisfies Equation (20). Let $r = I_n$, with $n \in \{1, \ldots, N_r\}$. Applying Proposition 1 we know that,

$$p(R_{x_0}(t) = I_n) = \alpha_n(\mathbf{0}) \int_0^t e^{-a(I_n)(t-s)} \, p(R_{x_0}(s) = \mathbf{0}) \, ds. \tag{21}$$

Now we consider the path chains leading to I_n. The state can only be reached in one way, hence, we have that $G(I_n) = \{g = (\mathbf{0}, I_n)\}$. Now calculating the probability of the evolution of a path chain (Definition 5) over $g \in G(I_n)$ gives us,

$$p(X_{g,1,0}(t) = I_n) := \beta(\mathbf{0}, I_n) \int_0^t e^{-\alpha(I_n)(t-s)} p(X_{g,1}(s) = \mathbf{0}) \, ds. \tag{22}$$

Given the two equations above match, we can conclude that for $r \in \nabla_1$,

$$p(R(t) = r) = \sum_{g \in G(r)} p(X_{g,1}(t) = r).$$

Now we perform the inductive step. Fix $m > 1$. Assume that for $\check{r} \in \nabla_m$,

$$p(R_{x_0}(t) = \check{r}) = \sum_{g \in G(\check{r})} p(X_{g,1}(t) = \check{r}). \tag{23}$$

Then for $r \in \nabla_{m+1}$, Proposition 1 states that,

$$p(R(t) = r) = \sum_{n=1}^{N_r} \alpha_n(r - I_n) \int_0^t e^{-\alpha(r)(t-s)} p(R_{x_0}(s) = r - I_n) \, ds,$$

using the inductive assumption Equation (23) on the right-hand side inside the integral,

$$= \sum_{n=1}^{N_r} \alpha_n(r - I_n) \int_0^t e^{-\alpha(r)(t-s)} \sum_{g \in G(r-I_n)} p(X_{g,1}(s) = r - I_n) \, ds.$$

Rewriting the propensities in the path propensity notation gives,

$$= \sum_{n=1}^{N_r} \beta(r - I_n, r) \int_0^t e^{-\alpha(r)(t-s)} \sum_{g \in G(r-I_n)} p(X_{g,1}(s) = r - I_n) \, ds,$$

$$= \sum_{n=1}^{N_r} \sum_{g \in G(r-I_n)} \beta(r - I_n, r) \int_0^t e^{-\alpha(r)(t-s)} p(X_{g,1}(s) = r - I_n) \, ds.$$

Since all paths in $G(r)$ are paths from the set $\cup_{n=1}^{N_r} G(r - I_n)$, with r as the last state, the two summations reduce to give,

$$= \sum_{\hat{g} \in G(r)} \beta(\hat{g}_m, r) \int_0^t e^{-\alpha(r)(t-s)} p(X_{\hat{g},1}(s) = \hat{g}_m) \, ds.$$

By Definition 5, the term above captures all the paths in $G(r)$, therefore,

$$= \sum_{\hat{g} \in G(r)} p(X_{\hat{g},1}(s) = r) \, ds.$$

We have shown that given Equation (23) holds on ∇_m, then Equation (23) also holds for ∇_{m+1}. Hence, by mathematical induction, for any $r \in \Lambda$,

$$p(R_{x_0}(t) = r) = \sum_{g \in G(r)} p(X_{g,1}(t) = r) \, ds.$$

□

We have shown that each state in the reaction counts CME can be decomposed into the framework of "path chains" that we have defined. In essence, we have shown that the path chains are trajectories of the reaction counts Kurtz process. Even though at first glance the result is not a startling revelation, the real novelty lies in how we derived the probability distributions for these trajectories. By carefully decomposing the state space and concurrently deriving the corresponding probability distributions, we were able to give an analytical form for the probability of individual trajectories of the Kurtz process. Using the probability distribution given over the path chains in Definition 5, we can derive the classical stochastic simulation algorithm [20]. Using the path chains framework we have introduced so far, we can build realisations for applications such as: Hitting time problems, rare event problems, observable problems, etc. In the following section, we further build on the properties of path chains by proposing a method for estimating the path probabilities of trajectories.

4.2. Gated- and Un-Gated Path Chains

Given a path or trajectory, we define the path probability to be the probability distribution of the last state in the chain. We can use Definition 5 to analytically compute the probability of any path. However, we can exploit the properties of the path chains further and derive some simple estimates for the path probability. We borrow the notion of gating and un-gating introduced by Sunkara (2013) [21], and show how these can be used to calculate estimates for the path probabilities.

The concept of gating is mathematically tedious to prove, however, the concept is fairly simple. Given a chain g of length m, gating this chain at position $j < m$ means that we set the propensity of leaving state g_j to zero. Hence, the probability will simply flow into state g_j and remain there (hence, the term "gating"). Then we can "un-gate" the state g_j, which means that we reset the time to zero, set the initial probability at state g_j to the accumulated probability from the gating, and then allow the process to continue. The reason we are interested in this notion of gating and un-gating is that for path chains, it happens so that gating and un-gating the chains gives an upper bound for the original path probability. We will now prove that this is indeed the case.

Definition 6. *For an admissible path $g \in \gamma_m$, we define a truncated path $g^j := (g_j, g_{j+1}, \ldots, g_m)$, where the first $j-1$ terms of the path are ignored.*

Definition 7. *Fix $\Delta t < \infty$ and $x_0 \in \Omega$. Let $g \in \gamma_m$ be a path of length $m+1$. We define $(X_{g,C}(t))_{t \in [0,\Delta t]}$ to be a path chain over g in the time interval $[0, \Delta t]$. Then for $j \in \{2, \ldots, m\}$ we define an un-gated path chain over g^j to be given by:*

$$p(X_{g^j,C}(t) = g_k) := \beta(g_{k-1}, g_k) \int_0^t e^{-\alpha(g_k)\,(t-s)}\, p(X_{g,C_j}(s) = g_{k-1})\, ds, \tag{24}$$

for $k \in \{j+1, \ldots, m\}$ and,

$$p(X_{g^j,C}(t) = g_j) := e^{-\alpha(g_j)\,t}\, C_j$$

where,

$$C_j := \beta(g_{j-1}, g_j) \int_0^{\Delta t} p(X_{g,C}(s) = g_{j-1}).$$

To conserve probability,

$$p(X_{g^j,C}(t) = g_k) = \text{sink}) := 1.0 - \sum_{g_k \in g^j} p(X_{g,C}(t) = g_k).$$

We see that the un-gated path chain is a normal path chain with the initial probability reset to the total probability which has flown into the gated state (see Figure 3). We now show that gating and un-gating over a path gives an upper bound for the path probability. First we consider the case of a single gating and un-gating at the position g_2.
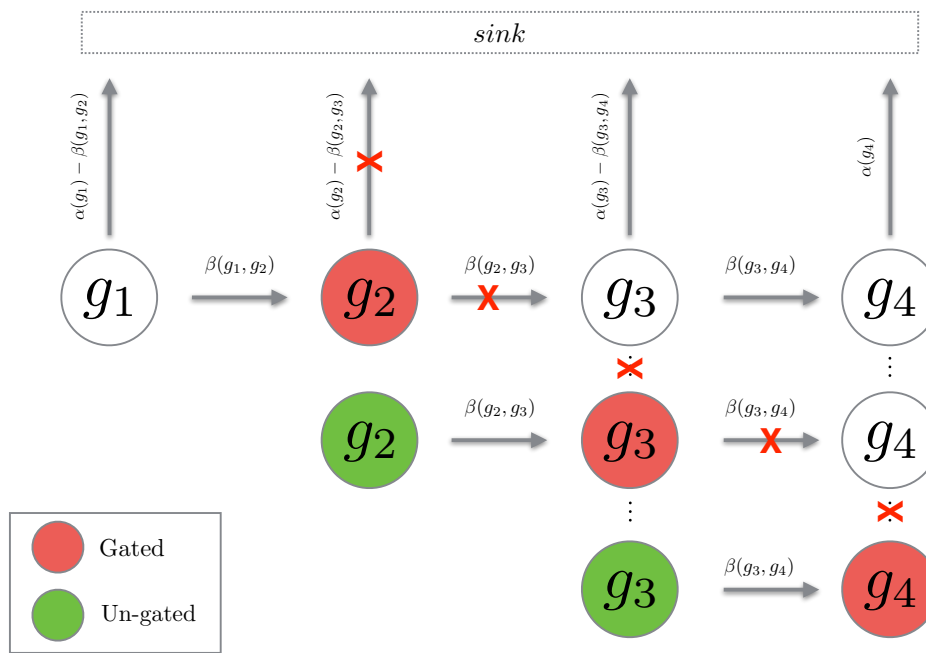
Figure 3. The cartoon above depicts the cascade of gating being performed on a path chain. When a state is gated, the propensities leaving the state are set to zero (depicted with a red cross). When the state is un-gated, the propensities are reintroduced.

Lemma 2. *Fix $g \in \gamma_m$ and $\Delta t > 0$. Given two path chains $(X_{g,C}(t))_{t \in [0,\Delta t]}$ and $(X_{g^2,C_2}(t))_{t \in [0,\Delta t]}$, it follows that for all $t \in [0, \Delta t]$,*

$$\int_0^t p(X_{g,C}(s) = 2) \, dt \leq \int_0^t p(X_{g^2,C_2}(s) = 2) \, dt,$$

and for $j \in \{3, \ldots, m\}$

$$p(X_{g,C}(t) = g_j) \leq p(X_{g^2,C_2}(t) = g_j).$$

Proof. We break down the proof into two steps. First we investigate the bounds formed by gating at the second step, and then we prove the bound over the remaining states in the chain. We begin by showing that for $t \in [0, \Delta t]$,

$$\int_0^t p(X_{g,C}(s) = 2) \, dt \leq \int_0^t p(X_{g^2,C_2}(s) = 2) \, dt. \tag{25}$$

From the definition of the path chain (Definition 5), the solution for the probability of being in state g_2 at time t is given by,

$$p(X_{g,C}(t) = g_2) = \frac{C \, \beta(g_1, g_2)}{\alpha(g_1) - \alpha(g_2)} \left(e^{-\alpha(g_2)t} - e^{-\alpha(g_1)t} \right). \tag{26}$$

Using the solution above, we can calculate the probability which would accumulate in g_2 if the chain was gated at g_2,

$$C_2 := \frac{C \, \beta(g_1, g_2)}{\alpha(g_1)} \left(1 - e^{-\alpha(g_1)\Delta t} \right). \tag{27}$$

Substituting the above terms into the definition of the probability of being is state g_2 at time t in the un-gated path chain $(X_{g^2,C_2}(t))_{t \in [0,\Delta t]}$ gives us,

$$p(X_{g^2,C_2}(t) = g_2) = \frac{C \, \beta(g_1, g_2)}{\alpha(g_1)} \left(1 - e^{-\alpha(g_1)\Delta t} \right) e^{-\alpha(g_2)t}. \tag{28}$$

Taking the integral of Equations (26) and (28) over the interval $[0, t] \subseteq [0, \Delta t]$ gives us,

$$\int_0^t p(X_{g,C}(s) = g_2)\,ds = \frac{C\,\beta(g_1, g_2)}{(\alpha(g_1) - \alpha(g_2))\,\alpha(g_1)\,\alpha(g_1)}\left(\alpha(g_2)(e^{-\alpha(g_1)\,t} - 1) - \alpha(g_1)(e^{-\alpha(g_2)\,t} - 1)\right), \quad (29)$$

$$\int_0^t p(X_{g^2,C_2}(s) = g_2)\,ds = \frac{C\,\beta(g_1, g_2)}{\alpha(g_1)\,\alpha(g_2)}\left(1 - e^{-\alpha(g_1)\Delta t}\right)\left(1 - e^{-\alpha(g_2)t}\right). \quad (30)$$

To simplify notation, let us define,

$$\hat{\phi}(t) := \int_0^t p(X_{g^2,C_2}(s) = g_2)\,ds \text{ and } \phi(t) := \int_0^t p(X_{g,C}(s) = g_2)\,ds.$$

Then the difference between $\hat{\phi}(t)$ and $\phi(t)$ reduces to,

$$\hat{\phi}(t) - \phi(t) = \frac{C\,\beta(g_1, g_2)}{\alpha(g_1)\,\alpha(g_2)}\left(1 - e^{-\alpha(g_1)\Delta t}\right)\left(1 - e^{-\alpha(g_2)t}\right)$$
$$- \frac{C\,\beta(g_1, g_2)}{(\alpha(g_1) - \alpha(g_2))\,\alpha(g_1)\,\alpha(g_1)}\left(\alpha(g_2)(e^{-\alpha(g_1)\,t} - 1) - \alpha(g_1)(e^{-\alpha(g_2)\,t} - 1)\right). \quad (31)$$

Since $e^{-\alpha(g_1)\Delta t} \leq e^{-\alpha(g_1)t}$ for all $t \in [0, \Delta t]$, we attain the lower bound,

$$\geq \frac{C\,\beta(g_1, g_2)}{\alpha(g_1)\,\alpha(g_2)}\left(1 - e^{-\alpha(g_1)t}\right)\left(1 - e^{-\alpha(g_2)t}\right)$$
$$- \frac{C\,\beta(g_1, g_2)}{(\alpha(g_1) - \alpha(g_2))\,\alpha(g_1)\,\alpha(g_1)}\left(\alpha(g_2)(e^{-\alpha(g_1)\,t} - 1) - \alpha(g_1)(e^{-\alpha(g_2)\,t} - 1)\right). \quad (32)$$

We show that the right-hand side of the equation above is always positive. We define the right hand-side of Equation (32) as a function,

$$\Psi(g_1, g_2, t) := \frac{C\,\beta(g_1, g_2)\left(e^{-\alpha(g_2)t}\alpha(g_2) - e^{-\alpha(g_1)t}\alpha(g_1) + e^{-(\alpha(g_1)+\alpha(g_2))t}(\alpha(g_1) - \alpha(g_2))\right)}{(\alpha(g_1) - \alpha(g_2))\,\alpha(g_1)\,\alpha(g_2)}. \quad (33)$$

We notice that the function Ψ has the property,

$$\Psi(g_1, g_2, t) = \Psi(g_2, g_1, t).$$

Firstly, if we consider the case $\alpha(g_1) = \alpha(g_2)$, we can see that $\psi(g_1, g_2, t) = 0$. Using the symmetry of Ψ, we only need to consider the case $\alpha(g_1) < \alpha(g_2)$. Then it follows that,

$$\Psi(g_1, g_2, t) = \frac{C\,\beta(g_1, g_2)\left(e^{-\alpha(g_2)\Delta t}\alpha(g_2) - e^{-\alpha(g_1)t}\alpha(g_1) + e^{-(\alpha(g_1)+\alpha(g_2))t}(\alpha(g_1) - \alpha(g_2))\right)}{(\alpha(g_1) - \alpha(g_2))\,\alpha(g_1)\,\alpha(g_2)},$$

replacing $e^{-\alpha(g_1)t}\alpha(g_1)$ with $e^{-\alpha(g_1)t}\alpha(g_2)$ gives us,

$$\geq \frac{C\,\beta(g_1, g_2)\left((e^{-\alpha(g_2)t} - e^{-\alpha(g_1)t})\alpha(g_2) + e^{-(\alpha(g_1)+\alpha(g_2))t}(\alpha(g_1) - \alpha(g_2))\right)}{(\alpha(g_1) - \alpha(g_2))\,\alpha(g_1)\,\alpha(g_2)},$$

$$= C\,\beta(g_1, g_2)\left(\underbrace{\frac{(e^{-\alpha(g_2)t} - e^{-\alpha(g_1)t})}{(\alpha(g_1) - \alpha(g_2))\,\alpha(g_1)}}_{>0} + \underbrace{\frac{e^{-(\alpha(g_1)+\alpha(g_2))\Delta t}}{\alpha(g_1)\,\alpha(g_2)}}_{>0}\right),$$

$$\geq 0.$$

Hence, the right-hand side of Equation (33) is positive. So it follows that for $t \in [0, \Delta t]$,

$$\int_0^t p(X_{g,C}(s) = 2)\, ds \leq \int_0^t p(X_{g^2,C_2}(s) = 2)\, ds.$$

We have shown that the outflow of probability in the un-gated chain is more than in the original path chain. We now show that this phenomenon continues all the way down the chain. We prove this using mathematical induction. First, let us consider the base case $j = 3$. By the definition of path chains,

$$p(X_{g,C}(t) = g_3) - p(X_{g^2,C_2}(t) = g_3) = \beta(g_2, g_3) \int_0^t e^{-\alpha(g_3)(t-s)} \left(p(X_{g,C}(s) = g_2) - p(X_{g^2,C_2}(s) = g_2) \right) ds,$$

$$\leq \beta(g_2, g_3) \int_0^t \left(p(X_{g,C}(s) = g_2) - p(X_{g^2,C_2}(s) = g_2) \right) ds,$$

applying condition (25) gives us that,

$$\leq 0.$$

Hence, for $t \in [0, \Delta]$,

$$p(X_{g,C}(t) = g_3) \leq p(X_{g^2,C_2}(t) = g_3).$$

We now build the inductive step. Fix $k \in \{4, \ldots, m\}$, assume that, for all $t \in [0, \Delta t]$,

$$p(X_{g,C}(t) = g_k) \leq p(X_{g^2,C_2}(t) = g_k). \tag{34}$$

Then by the definition of path chains,

$$p(X_{g,C}(t) = g_{k+1}) - p(X_{g^2,C_2}(t) = g_{k+1}) = \beta(g_k, g_{k+1}) \int_0^t e^{-\alpha(g_{k+1})(t-s)} \left(p(X_{g,C}(s) = g_k) - p(X_{g^2,C_2}(s) = g_k) \right) ds,$$

$$\leq \beta(g_k, g_{k+1}) \int_0^t \left(p(X_{g,C}(s) = g_k) - p(X_{g^2,C_2}(s) = g_k) \right) ds,$$

applying the inductive assumption Equation (34) gives us that,

$$\leq 0.$$

Hence, by mathematical induction, for $j \in \{3, \ldots, m\}$ and $t \in [0, \Delta t]$, the probabilities of the un-gated chain bound the probability on the original path chain, that is,

$$p(X_{g,C}(t) = g_j) \leq p(X_{g^2,C_2}(t) = g_j).$$

$\square$

The intuition behind Lemma 2 is that once we gate and un-gate a state, the un-gated chain has more probability in it than the regular path chain beyond the gated state. Since the dynamics of the path chain beyond the gated state are not changed, we would expect that the un-gated chain (having more probability) would be an upper bound for the original path chain. The proof, as seen above, is sadly tedious. We now consider a cascade of gating and un-gating, that is, we gate and un-gate as we transition through the path chain. We prove that this gives an upper bound for the path probabilities.

4.3. Cascade of Gating and Un-Gating

We now set up a cascade of gating and un-gating at ever state of the path chain. We achieve this by defining the starting probability of the un-gating. Previously, an un-gated path chain at position j had a initial probability of,

$$C_j := \beta(g_{j-1}, g_j) \int_0^{\Delta t} p(X_{g,C}(s) = g_{j-1}) \, ds.$$

We now introduce a recurrent initial probability,

$$\tilde{C}_j := \beta(g_{j-1}, g_j) \int_0^{\Delta t} p(X_{g^{j-1}, \tilde{C}_{j-1}}(s) = g_{j-1}) \, ds, \tag{35}$$

for $j \in \{2, \ldots, m-1\}$. Using this new recursive initial probability, we will build a sequence of path chains and using the idea of gating and un-gating at each transition of the path chain, we can construct an upper bound for the path probability.

Theorem 3. *Fix* $g \in \gamma_m$ *and* $\Delta t > 0$. *Let* $(X_{g,C}(t))_{t \in [0, \Delta t]}$ *be a path chain over* g. *We recursively define a set of un-gated chains,* $\{(X_{g^j, \tilde{C}_j}(t))_{t \in [0, \Delta t]} : j \in \{2, \ldots, m-2\}\}$. *Then for* $k \in \{3, \ldots, m-1\}$,

$$\int_0^{\Delta t} p(X_{g,C}(s) = g_k) \, ds \le \int_0^{\Delta t} p(X_{g^{k-1}, \tilde{C}_{k-1}}(s) = g_k) \, ds.$$

Proof. We prove this result by repeatedly applying Lemma 2. Fix $k \in \{3, \ldots, m-1\}$, then by the first clause of Lemma 2 we have that,

$$\int_0^{\Delta t} p(X_{g^{k-1}, \tilde{C}_{k-1}}(s) = g_k) \, ds \ge \int_0^{\Delta t} p(X_{g^{k-2}, \tilde{C}_{k-2}}(s) = g_k) \, ds,$$

since g_k is the third state of the chain g^{k-2}, we can apply the second clause of Lemma 2 to get,

$$\ge \int_0^{\Delta t} p(X_{g^{k-3}, \tilde{C}_{k-3}}(s) = g_k) \, ds,$$

stepping backward in the chain using Lemma 2, we reduce down to,

$$\vdots$$

$$\ge \int_0^{\Delta t} p(X_{g,C}(s) = g_k) \, ds.$$

□

In summary, by gating and un-gating along a path chain we can easily construct estimates for the probability of the trajectories of interest. We now consider an example where we build a path chain of the birth-death process introduced in Example 1, to show how to compute the gating and un-gating steps.

Example 2. *Let us consider the birth-death process introduced in Example 1, with the parameters* $x_0 = 2, c_b = 1.0$, *and* $c_d = 0.1$. *We will now construct the path probability of the chain:*

$$g = ((0,0), (1,0), (1,1), (1,2), (1,3)).$$

The generator for this path chain is given by,

$$
A^* =
\begin{array}{c}
\begin{array}{cccccc}
(0,0) & (1,0) & (1,1) & (1,2) & (1,3) & \underline{sink}
\end{array} \\
\begin{bmatrix}
-c_b - 2\,c_d & 0 & 0 & 0 & 0 & 0 \\
c_b & -c_b - 3\,c_d & 0 & 0 & 0 & 0 \\
0 & -3\,c_d & -c_b - 2\,c_d & 0 & 0 & 0 \\
0 & 0 & 2\,c_d & -c_b - c_d & 0 & 0 \\
0 & 0 & 0 & c_d & 0 & 0 \\
2\,c_d & c_b & c_b & c_b & 0 & 0
\end{bmatrix}
\end{array}.
$$

We can solve for $p(X_g, 1(t) = (1,3))$ by simply taking the matrix exponential at time t and applying it to the respective initial probability. Now we will construct the gating and un-gating approximation. This approximation is calculated using the following recursive equation: Let m be the number of states in the chain g, then,

$$
u_n(t) = \frac{\beta(g_{n-1}, g_n)}{\alpha(g_{n-1})} \left(1.0 - e^{-\alpha(g_{n-1})\,t} \right) u_{n-1}(t),
$$

for $n = 2, \ldots, m$ with $u_1 = 1.0$. The functions α and β are as given in Definition 5. Hence, $u_5(t)$ is the gated and un-gated estimate for the probability $p(X_{g,1}(t) = (1,3))$. In Figure 4A, we plot the true path probability, $p(X_{g,1}(t) = (1,3))$, and the upper bound for the path probability, $u_5(t)$, for time interval $t \in [0,10]$ seconds. The graph shows that $u_5(t)$ is an upper bound for $p(X_g, 1(t) = (1,3))$ at any time point. It is interesting to observe that the error between the two functions decreases over time, however it is not clear whether the difference will converge to zero as time goes to infinity. To test whether this behaviour can be observed in other parameter settings, we plot the analytical path probability and its upper bound for three different parameter settings in Figure 4B. We observe the same effect in all three presented cases. This raises the question whether this observation is a structural property or only an artefact of the specific model which is considered here.
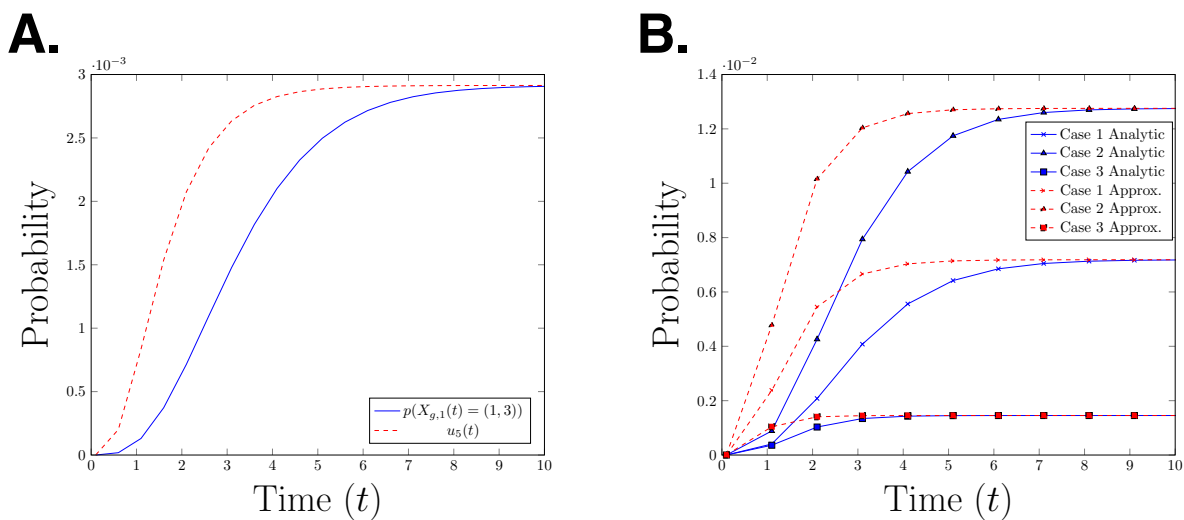
A.

B.

Figure 4. (**A**) Graph of the path probability $p(X_{g,1}(t) = (1,3))$ and the upper bound of the path probability $u_5(t)$ for the time interval $t \in [0,10]$. (**B**) Case 1 : $(c_b = 1.0, c_d = 0.15)$, Case 2 : $(c_b = 1.0, c_d = 0.2)$, Case 3 : $(c_b = 2.0, c_d = 0.15)$. "analytical" refers to the path probability and "approximation" refers to the upper bound of the path probability.

5. Conclusions

We began by formulating the reaction counts CME and showing that its forward stepping characteristic yields analytical solutions. Then, with a simple application of the push forward measure, we could prove the existence of solutions of the CME for finite time. We further decomposed the

reaction counts state space into independent CTMCs and gave analytical forms and estimates for their path probabilities. Hence, we have derived analytical theory for trajectories which arise from the solutions of the CME. This theory can be used for computing observable estimates of the underlying CTMCs. We can also use the paths to estimate the hitting times and rare event probabilities. A natural future direction would be to investigate if the current framework of the time independent propensities can be translated to the time dependent propensities. An even more challenging task would be, in the context of parameter inference, to reverse time on the trajectories and to cascade probability backwards.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Wilkinson, D.J. Mathematical and Computational Biology Series. In *Stochastic Modelling for Systems Biology*; Chapman & Hall/CRC: Boca Raton, FL, USA, 2006.
2. Van Kampen, N.G. *Stochastic Processes in Physics and Chemistry*, 3rd ed.; North Holland: Amsterdam, The Netherlands, 2007.
3. Kurtz, T.G. Strong approximation theorems for density dependent Markov chains. *Stoch. Process. Their Appl.* **1978**, *6*, 223–240, [CrossRef]
4. Grima, R.; Thomas, P.; Straube, A.V. How accurate are the nonlinear chemical Fokker-Planck and chemical Langevin equations? *J. Chem. Phys.* **2011**, *135*, 084103, [CrossRef] [PubMed]
5. Higham, D.J. Modeling and Simulating Chemical Reactions. *SIAM Rev.* **2008**, *50*, 347–368, [CrossRef]
6. Hegland, M.; Hellander, A.; Lötstedt, P. Sparse grids and hybrid methods for the chemical master equation. *BIT Numer. Math.* **2008**, *48*, 265–283, [CrossRef]
7. Engblom, S. Spectral approximation of solutions to the chemical master equation. *J. Comput. Appl. Math.* **2009**, *229*, 208–221, [CrossRef]
8. Jahnke, T.; Udrescu, T. Solving chemical master equations by adaptive wavelet compression. *J. Comput. Phys.* **2010**, *229*, 5724–5741, [CrossRef]
9. Sunkara, V.; Hegland, M. An Optimal Finite State Projection Method. *Procedia Comput. Sci.* **2010**, *1*, 1579–1586, [CrossRef]
10. 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] [PubMed]
11. 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]
12. Vlysidis, M.; Kaznessis, Y. Solving Stochastic Reaction Networks with Maximum Entropy Lagrange Multipliers. *Entropy* **2018**, *20*, 700, [CrossRef]
13. Haseltine, E.L.; Rawlings, J.B. Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics. *J. Chem. Phys.* **2002**, *117*, 6959–6969, [CrossRef]
14. Goutsias, J. Quasiequilibrium approximation of fast reaction kinetics in stochastic biochemical systems. *J. Chem. Phys.* **2005**, *122*, 184102, [CrossRef] [PubMed]
15. Sunkara, V. The chemical master equation with respect to reaction counts. In Proceedings of the 18th World IMACS Congress and (MODSIM09) International Congress on Modelling and Simulation, Cairns, Australia, 13–17 July 2009; Volume 1, pp. 2377–2383.
16. Menz, S.; Latorre, J.; Schütte, C.; Huisinga, W. Hybrid Stochastic–Deterministic Solution of the Chemical Master Equation. *Multiscale Model. Simul.* **2012**, *10*, 1232–1262, [CrossRef]
17. Black, A.J.; Ross, J.V. Computation of epidemic final size distributions. *J. Theor. Biol.* **2015**, *367*, 159–165, [CrossRef] [PubMed]

18. Khammash, M.; Munsky, B. The finite state projection algorithm for the solution of the chemical master equation. *J. Chem. Phys.* **2006**, *124*, 1–12, [CrossRef]

19. Norris, J.R. *Markov Chains*; Cambridge University Press: Cambridge, UK, 1997; doi:10.1017/CBO9780511810633.

20. Gillespie, D.T. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **1977**, *81*, 2340–2361, [CrossRef]

21. Sunkara, V. Analysis and Numerics of the Chemical Master Equation. Ph.D. Thesis, The Australian National University, Canberra, Australia, 2013.