Stochastic reaction networks within interacting compartments

by

Aidan S. Howells

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Mathematics

at the
University of Wisconsin-Madison
2024

Date of Final Oral Exam: 2/27/2024
The dissertation is approved by the following members of the Final Oral Committee:

David F. Anderson, Professor, Mathematics
Hyukpyo Hong, Assistant Professor, Mathematics
Sebastien Roch, Professor, Mathematics
Amy Cochran, Assistant Professor, Mathematics

Stochastic reaction networks within interacting compartments

Aidan S. Howells

Abstract

Stochastic reaction networks, which are usually modeled as continuous-time Markov chains on $\mathbb{Z}^d_{\geq 0}$, and simulated via a version of the "Gillespie algorithm," have proven to be a useful tool for the understanding of processes, chemical and otherwise, in homogeneous environments. There are multiple avenues for generalizing away from the assumption that the environment is homogeneous, with the proper modeling choice dependent upon the context of the problem being considered. One such generalization was recently introduced in [1], where the proposed model includes a varying number of interacting compartments, or cells, each of which contains an evolving copy of the stochastic reaction system. The novelty of the model is that these compartments also interact via the merging of two compartments (including their contents), the splitting of one compartment into two, and the appearance and destruction of compartments. This thesis begins the systematic exploration of the mathematical properties of this model. We (i) obtain basic/foundational results pertaining to explosivity, transience, recurrence, and positive recurrence of the model, (ii) explore a number of examples demonstrating some possible non-intuitive behaviors of the model, (iii) identify the limiting distribution of the model in a special case that generalizes three formulas from an example in [1], and (iv) examine the case where the splitting rate is content-dependent. This thesis is best viewed as an expanded version of [2], and will be familiar to readers of that paper. New items in the thesis include an expanded exposition (e.g. Remark 2.3.5), proofs of some results which were left as exercises to the reader (e.g. Proposition 4.2.3), and a brand-new chapter with a case study on content-dependent fragmentation (point (iv) above).

Acknowledgements

For helpful conversations about notation and/or other mathematical details, I would like to thank Otto Baier, Erik Bates, David Clancy, Kevin Dao, Ari Davidovsky, Jacob Denson, James Hanson, Logan Heath, Hyukpyo Hong, Alex Hof, Amin Idelhaj, David Keating, Sam McKeown, Omer Mermelstein, Caroline Nunn, Diego Rojas La Luz, Evan Sorenson, and John Spoerl. I am sure there are others who I am forgetting; to them I apologize.

For suggesting the idea I used to complete the proof of Theorem 6.1.1, I would like to thank Martina Bonelli and Daniele Cappelletti.

For thoughtful advice about the big-picture content and direction of the thesis, I would like to thank Daniele Cappelletti, Gheorghe Craciun, and Diego Rojas La Luz.

In her role as director of graduate studies, for taking the time to reach out and make sure I was on track as the pandemic raged, I would like to thank Betsy Stovall.

For years of support, I am indebted to my wife Samantha. She put up with my busy schedule and travel, and has moved cities to follow me.

Lastly, I would like to thank my advisor, Dave Anderson. For being more patient with my early flailing attempts then I felt I deserved, for making time to meet with me even when he had lots on his plate at home, for advice about math from big picture direction to all sorts of details, for advice not about math, and for much more. I cannot help but feel that had I been less lucky in my choice of advisors, this thesis would never have been written.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Stochastic reaction networks are commonly utilized to model various types of systems in the biological sciences. These mathematical models are often continuous-time Markov chains and are used when the counts of at least some of the underlying "species," which are most commonly different molecule types, are low. In this low copy-number case, the state of the model is a vector giving the integer counts of the different species and transitions are governed by the different possible "reactions" that can take place. These models are typically simulated via the Gillespie algorithm (see [3, 4]) or the next reaction method (see [5, 6]). See [7], and references therein, for more on this type of model.

One potential drawback to the standard model is that it assumes a homogeneous environment. There are multiple ways to generalize away from this assumption. One common generalization is to split the environment into different fixed pieces (often called "voxels"), assume that the chemistry is well mixed within each voxel, and then allow for transitions between adjacent voxels; see [8, 9]. Thinking of the size of the voxels going to zero leads naturally to a model with continuous space in which the state of the system is given by the type, position, velocity, etc., of each particle in the system. A reaction can then only take place when the necessary constituent molecules are near each other (with the precise mechanism for defining when they are "near enough" left to the modeler). One of the first examples of such a continuous space model was introduced by Doi in [10]. More

generally, there are a whole class of continuous space models known as reaction-diffusion models. For a brief overview of such models, see [11]. For a comparison of two specific such models, with an approachable introduction, see [12]; for a more general approach, see the introduction of [13].

A different approach to generalize from the homogeneous case is to imagine some fixed collection of compartments and model the dynamics within each compartment in the usual way (as a continuous-time Markov chain as described in the first paragraph above) while also allowing for interactions between adjacent compartments. This is the approach taken in [14] in an ecological context (their "patches" are our "compartments"). However, ideally one might like to also account for situations like in cellular biology, where reactions take place in cells that are not static but, for example, can appear, divide, possibly merge, or even be destroyed. That is the approach presented in a recent paper by Duso and Zechner, where they developed a Markov model for stochastic reaction networks within interacting compartments [1]. In particular, their model consists of two basic components:

1. a stochastic model of a chemical reaction network;

2. a dynamic model of compartments, or cells, which themselves undergo basic transitions such as (i) arrivals, (ii) departures, (iii) mergers, and (iv) divisions. In the context of [1], these four transition types are referred to as inflows, exits, coagulations, and fragmentations, respectively.

Each compartment, or cell, contains a copy of the (evolving) chemical reaction network. When two cells merge, their contents are combined. When a cell divides, its contents are randomly split among the two new daughter cells. Beyond the framework itself, their paper focuses on the framework's practical use, using moment closure methods to derive estimates for various population statistics which are then validated by simulation. They also derive stationary distributions for some special cases.

In this thesis, we attempt to lay the groundwork for exploration of mathematical questions about the Markov chain model developed in [1]. We focus on the special case where the compartments can only enter, leave, merge, and divide, all according to mass action

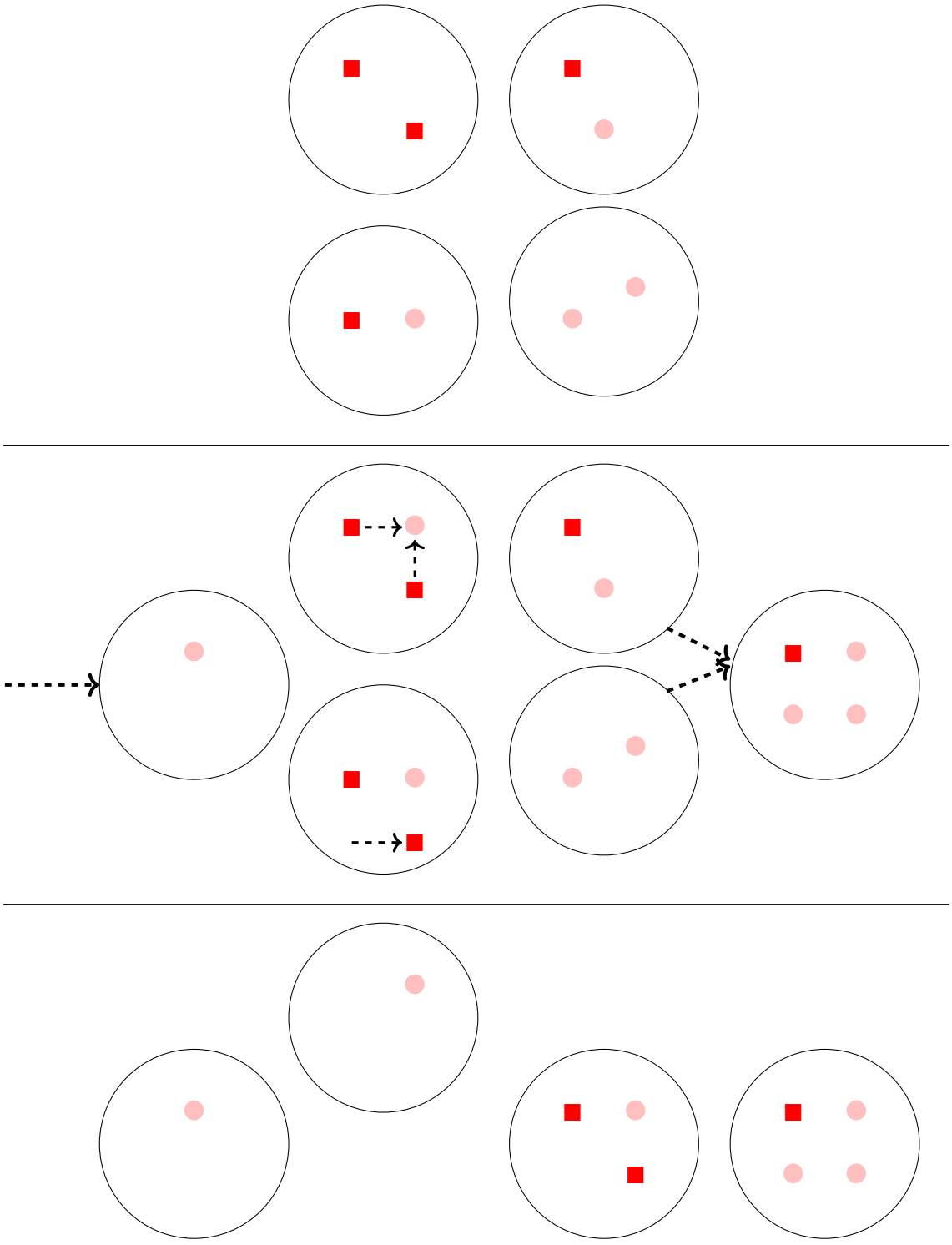


Figure 1.1: Imagine the reactions ■ + ■ → ● and 0 → ■, taking place inside compartments. The top shows a possible initial state of the system. The middle shows four possible changes that could take place in the system: one compartment is entering, two compartments are merging, and the other two compartments have chemistry taking place inside them. The bottom shows the state of the system after the changes.

kinetics and unaffected by their contents. Questions pertaining to recurrence, transience, and explosivity are all considered. We show that in most, but not all, parameter regimes the overall qualitative behavior of the model (i.e., recurrence, transience, or explosivity) is the same as that of one of the associated stochastic reaction networks. We also analyze myriad examples that, taken together, demonstrate some of the non-intuitive (and interesting) possible behaviors of the model. Moreover, we derive the stationary distribution for the model in the case where the chemistry inside the compartments is well understood in the sense that a formula for the distribution is known for all time (e.g., the DR models of [15]) and the compartments themselves are not allowed to interact (but are not totally static, being allowed to enter and leave the system). Two special cases of this stationary distribution are provided as illustration, both of which generalize formulas from an example in [1]. Lastly, we devote a chapter to the case where the rate of fragmentation *is* affected by the contents of the compartment. In that case we are able to obtains some results about explosivity, positive recurrence, and transience, some of which imply that a stationary distribution which was approximated numerically in [1] actually exists.

Before moving on, we warn the reader that in the field of epidemiology, the term "compartment model" has a different meaning. There the compartments are what we would call species. For example, they would speak of an SIR model as dividing individuals into a susceptible compartment, an infected compartment, and a recovered compartment. See e.g. [16].

A standard knowledge of continuous-time Markov chains is assumed. See for example Norris [17] for a detailed introduction to the topic. That said, we briefly pause to clarify the role of the generator of a process, which plays a key role in much of the analysis found in this thesis. Multiple different definitions of the generator of a of continuous-time Markov chain exist in the literature, all equivalent under suitable convergence assumptions (for example, the generator and full generator are defined in a more general context in chapter 1 of [18], whose proposition 1.5(c) relates them; see also chapter 5 of the same for details on the specialization of those definition to Markov processes). For the purpose of

this thesis, we work with the following:

Definition 1.0.1: Suppose X is a Markov chain with countable state space $\mathbb{S}$. For $x, y \in \mathbb{S}$, let q_{xy} denote the rate that X transitions to state y from state x (that is, if $p_{xy}(t) = \mathbb{P}_x(X(t) = y)$, then $q_{xy} = p'_{xy}(0)$). For $f : \mathbb{S} \times [0, \infty) \to \mathbb{R}$ such that $t \mapsto f(x, t)$ is differentiable for each x, define the *generator* $\mathcal{L}$ of X via

$$\mathcal{L}f(x,t) := \left(\sum_{y \in \mathbb{S}} f(y,t) q_{xy} \right) + \frac{\partial}{\partial t} f(x,t).$$

whenever the sum is absolutely convergent (in which case we will say that f is in the domain of the generator).

Note that most of the functions f which we take the generator of in this paper will not vary in time (that is, $f(x, t)$ will take the form $f(x)$, and so the time derivative term will vanish). The exception is in the proof of Theorem 6.1.1, which involves constructing a time-dependent Lyapunov function.

For notational convenience, we will use the following shorthand notations: for any two vectors $v, w \in \mathbb{R}^d_{\geq 0}$ and any vector $x, y \in \mathbb{Z}^d_{\geq 0}$ we denote

$$v^w = \prod_{i=1}^{d} (v_i)^{w_i} \quad \text{and} \quad x! = \prod_{i=1}^{d} (x_i)! \quad \text{and} \quad \binom{x}{y} = \prod_{i=1}^{d} \binom{x_i}{y_i},$$

with the conventions that $0^0 = 1$ and that $\binom{x}{y} = 0$ for $y < 0$ or $y > x$. Moreover, we will always use d to represent the number of species in the model. Finally, for $x \in \mathbb{Z}^d_{\geq 0}$ we define $e_x : \mathbb{Z}^d_{\geq 0} \to \mathbb{Z}$ to be the function taking the value of one at x and zero otherwise.

The remainder of the paper is outlined as follows. In chapter 2, we fully specify the model. Further, we give two different mathematical representations that are both useful and prove some first basic properties. In the brief section 3.1, we prove that the full model is explosive if and only if the associated reaction network is. In section 3.2, we give conditions for when the full model is recurrent, positive recurrent, or transient. In section 4, we provide the stationary distribution for a special class of models. Finally, in

chapter 5, we define and study a related model in which the compartment fragmentation rate depends on the contents of the compartment.

Chapter 2

The reaction network within interacting compartments (RNIC) model

As discussed in the introduction, the full model we consider here consists of two sub-models: (i) a stochastic reaction network and (ii) a dynamic model of compartments, or cells, each of which contains an evolving copy of the stochastic reaction network. We first describe these sub-models individually and then specify how they are combined to make the full model.

2.1 Stochastic reaction networks

Suppose we have a finite set $\mathcal{S}$, whose elements we shall call *species*, and a directed graph whose vertices are unique linear combinations of species with non-negative integer coefficients. The edges of the graph are called *reactions*; let $\mathcal{R}$ denote the set of reactions. The linear combinations which appear as vertices in the graph are called *complexes*; the set of complexes will be denoted $\mathcal{C}$. A *chemical reaction network* (or just *reaction network*; *CRN* for short) is the tuple $\mathcal{I} = (\mathcal{S}, \mathcal{C}, \mathcal{R})$, where $\mathcal{S}$, $\mathcal{C}$ and $\mathcal{R}$ are as above. See Figure 2.1

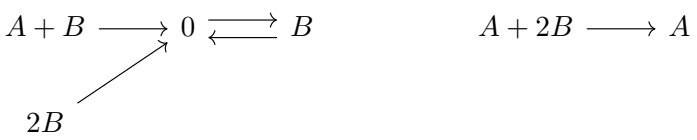


Figure 2.1: The CRN with species A and B and reactions $A + B \to 0$, $0 \to B$, $B \to 0$, $2B \to 0$, and $A + 2B \to A$. Note that 0 here denotes the linear combination $0A + 0B$.

for an example reaction network.

When talking about specific reaction networks, the species will usually be represented by capital Latin letters. When talking generally, there will be d species $S_1, \ldots, S_d$. In this case we will identify $\mathbb{Z}^d$ with the space of linear combinations of species with integer coefficients. That is, we naturally identify $\nu \in \mathcal{C}$ with the vector in $\mathbb{Z}^d$ whose ith element is the coefficient of S_i in ν. We will speak of reactions $\nu \to \nu' \in \mathcal{R}$, or sometimes, when we wish to enumerate the reactions as $\{\nu_r \to \nu'_r\}$, we will simply write $r \in \mathcal{R}$.

There are multiple ways to associate a mathematical model to a given reaction network, including the use of a deterministic ODE [19], a diffusion process [20, 21], and a continuous-time Markov chain [7]. The only one of concern to us here is the continuous-time Markov chain model with stochastic mass-action kinetics, in which the state of the system is a vector giving the number of each species present and transitions are determined by the reactions. To fully specify the model, positive (or sometimes, merely non-negative) numbers, called *rate constants*, are assigned to each reaction. If the reaction $\nu \to \nu'$ has rate constant κ, then in state x that particular reaction occurs with rate $\kappa\binom{x}{\nu}$ and when it occurs the chain transitions to state $x + \nu' - \nu$. So the reactions will happen with rate proportional to the number of ways the chemicals can combine to allow them to happen, and κ is the constant of proportionality. If $\mathcal{K}$ is a set of rate constants, one for each reaction, we denote by $\mathcal{I}_\mathcal{K} = (\mathcal{S}, \mathcal{C}, \mathcal{R}, \mathcal{K})$ the corresponding stochastic mass-action system. If we let $\kappa_{\nu \to \nu'}$ be the rate constant for the reaction $\nu \to \nu'$, then the Markov

chain transitions from state $x \in \mathbb{Z}^d_{\geq 0}$ to state $y \in \mathbb{Z}^d_{\geq 0}$ with rate

$$q(x,y) = \sum_{\substack{\nu \to \nu' \in \mathcal{R} \\ \nu' - \nu = y - x}} \kappa_{\nu \to \nu'} \binom{x}{\nu} = \sum_{\substack{\nu \to \nu' \in \mathcal{R} \\ \nu' - \nu = y - x}} \kappa_{\nu \to \nu'} \prod_{j=1}^{d} \binom{x_j}{\nu_j} \tag{2.1}$$

where the sum is over those reactions for which $\nu' - \nu = y - x$. For $r = \nu_r \to \nu'_r \in \mathcal{R}$, we denote the rate of the reaction r in state $x \in \mathbb{Z}^d_{\geq 0}$ by $\lambda_r(x)$:

$$\lambda_r(x) = \kappa_r \binom{x}{\nu_r} \tag{2.2}$$

Note that $\lambda_{\nu \to \nu'}(x) = 0$ if $x_i < \nu_i$ for some i, since $\binom{m}{k} = 0$ for $k > m$. Note also that not all authors take the same conventions as we do here. In fact, the convention we use here pertaining to our rate constants is more in line with the biology literature [22]. In the mathematical literature it is more common to use a falling factorial $\lambda_{\nu \to \nu'}(x) = \kappa_{\nu \to \nu'} \prod_j (x_j)(x_j - 1) \cdots (x_j - \nu_j + 1) = \kappa_{\nu \to \nu'} \frac{x!}{(x-\nu)!}$, at the cost that their rate constant κ is no longer the constant of proportionality when the reaction takes multiple inputs [7]. This choice plays no fundamental role in our results, but makes certain expressions cleaner in the present context.

We note here that many of the results found in this paper can be generalized to systems with kinetics, i.e., rate functions λ_r, other than mass-action. See Remark 3.2.7.

Put more succinctly, we have a Markov process on $\mathbb{Z}^d_{\geq 0}$ with infinitesimal generator

$$\mathcal{L}f(x) = \sum_{r \in \mathcal{R}} \lambda_r(x)(f(x + \nu'_r - \nu_r) - f(x)),$$

where λ_r is determined via (2.2), and the above is valid for all functions f that are compactly supported [18]. The Kolmogorov forward equation, often called the chemical master equation in the context of reaction networks, is then

$$\frac{d}{dt} P_\mu(x,t) = \sum_{r \in \mathcal{R}} \lambda_r(x - (\nu'_r - \nu_r)) P_\mu(x - (\nu'_r - \nu_r), t) - \sum_{r \in \mathcal{R}} \lambda_r(x) P_\mu(x,t),$$

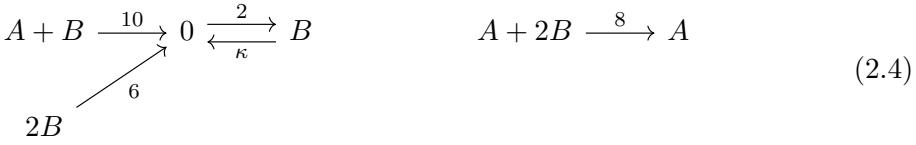
where $P_\mu(x,t) = P_\mu(X(t) = x)$ is the probability the process X is in state $x \in \mathbb{Z}^d_{\geq 0}$ at time t, given an initial distribution of μ. We take the convention that $P_\mu(x,t) = 0$ for $x \notin \mathbb{Z}^d_{\geq 0}$.

One way to represent the solution to the stochastic model described above is via a representation developed and popularized by Thomas Kurtz. Let $\{Y_r\}_{r\in\mathcal{R}}$ be a collection of independent, unit-rate Poisson processes, one for each possible reaction, and let $X(t), t \geq 0$, be the solution to

$$X(t) = X(0) + \sum_{r\in\mathcal{R}} Y_r \left(\int_0^t \lambda_r(X(s))ds \right) (\nu'_r - \nu_r), \tag{2.3}$$

then X is a continuous-time Markov chain that satisfies the conditions of the model specified above [7, 18, 23].

Example 2.1.1: Suppose we assign rate constants to the example CRN in Figure 2.1 as follows:



(2.4)

Let $x = (a,b) \in \mathbb{Z}^2_{\geq 0}$ denote an arbitrary state of the system. For the particular choice of rate constants given above the positive transition rates $q((a,b),\cdot)$, for $a, b \in \mathbb{Z}_{\geq 0}$, are

Reaction(s)	Transition	Rate
$A + B \to 0$	$(a,b) \mapsto (a-1, b-1)$	$10ab$
$0 \to B$	$(a,b) \mapsto (a, b+1)$	2
$2B \to 0$ and $A + 2B \to A$	$(a,b) \mapsto (a, b-2)$	$6\frac{b(b-1)}{2} + 8a\frac{b(b-1)}{2}$
$B \to 0$	$(a,b) \mapsto (a, b-1)$	κb

We chose to write $6\frac{b(b-1)}{2} + 8a\frac{b(b-1)}{2}$ instead of $3b(b-1) + 4ab(b-1)$ to emphasize our choice of intensity functions. Note that all other rates, such as $q((a,b),(a+1,b))$ or $q((a,b),(a+12,b-3))$, are zero. △

2.2 Compartment model

Having fully specified our CRN, $\mathcal{I}_{\mathcal{K}} = (\mathcal{S}, \mathcal{C}, \mathcal{R}, \mathcal{K})$, we turn to our next sub-model: the compartment model. As mentioned in the introduction, we will assume that compartments, or cells, can arrive, depart, merge, and divide. We can use the notation of chemical reaction networks to describe the four possibilities visually via a reaction network,

$$0 \underset{\longleftarrow}{\longrightarrow} C \underset{\longleftarrow}{\longrightarrow} 2C$$

with $0 \to C$ representing arrivals, $C \to 0$ representing departures, $C \to 2C$ representing division, or fragmentation, and $2C \to C$ representing mergers, or coagulations. Moreover, we assume that the stochastic model tracking the number of compartments behaves as a standard stochastic reaction network as already described in the previous section (however, see Remark 3.2.7 for an allowable generalization to the choice of kinetics). We will term this reaction network the *compartment network*, and denote it by $\mathcal{H} = (\mathcal{S}_{\text{comp}}, \mathcal{C}_{\text{comp}}, \mathcal{R}_{\text{comp}})$. Note that $\mathcal{S}_{\text{comp}} = \{C\}$ and $\mathcal{C}_{\text{comp}}$ is a subset of $\{0, C, 2C\}$ (depending on which rate constants are non-zero). If rate constants are added as follows,

$$0 \underset{\kappa_E}{\overset{\kappa_I}{\rightleftarrows}} C \underset{\kappa_C}{\overset{\kappa_F}{\rightleftarrows}} 2C$$

where each $\kappa_E, \kappa_I, \kappa_C, \kappa_F \geq 0$, then we will denote the corresponding stochastic mass-action system by $\mathcal{H}_{\mathcal{K}} = (\mathcal{S}_{\text{comp}}, \mathcal{C}_{\text{comp}}, \mathcal{R}_{\text{comp}}, \mathcal{K}_{\text{comp}})$. According to (2.3), if we denote by $M_C(t)$ the number of compartments at time t, then one way to represent this model is as the solution to

$$\begin{aligned} M_C(t) = M_C(0) + Y_I\left(\kappa_I t\right) - Y_E\left(\int_0^t \kappa_E M_C(s) ds\right) + Y_F\left(\int_0^t \kappa_F M_C(s) ds\right) \\ - Y_C\left(\int_0^t \kappa_C \frac{M_C(s)(M_C(s)-1)}{2}\, ds\right), \end{aligned}$$

where Y_I, Y_E, Y_F, and Y_C are independent unit-rate Poisson processes.

2.3 Specifying the full, combined model

Our full model, which we will term a *reaction network within interacting compartments* (RNIC), begins with two networks, one representing the dynamics of the compartments themselves and one representing the chemistry taking place inside the compartments.

- A CRN $\mathcal{H}_{\mathcal{K}}$ of the form $0 \leftrightarrows C \leftrightarrows 2C$, called the *compartment network*. The state of this CRN (in $\mathbb{Z}_{\geq 0}$) will be the number of compartments.
- An CRN $\mathcal{I}_{\mathcal{K}}$, called the *chemistry* (or *Internal network*), with d species.

The behavior of the model between transitions of the compartment model is straightforward: the CRN within each compartment evolves independently as a Markov chain with transition rates specified by (2.1). All that remains is to specify what happens to the full model at the transition times of the compartment model. Hence, there are four cases to consider.

- An arrival: $0 \to C$. We assume the existence of a probability measure μ on $\mathbb{Z}^d_{\geq 0}$. Each time an arrival event occurs, we add a new compartment whose initial state is chosen according to μ, independent of the past. (Note that μ is not necessary when $\kappa_I = 0$.)
- A departure: $C \to 0$. When a departure event occurs, we choose one of the compartments, uniformly at random, for deletion.
- A merger: $2C \to C$. When a merger event occurs, we select two compartments, uniformly at random. We replace the chosen compartments with a single compartment. The state of the new compartment is the sum of the states of the two it replaced.
- A division: $C \to 2C$. When a division event occurs, we select a compartment, uniformly at random. We replace the chosen compartment with two new compartments, whose initial states are determined by having each molecule from the chosen compartment select one of the two new compartments uniformly. For example, if

there are n_A type A species in the chosen compartment, then one of the new compartments will get a number of A molecules given by a binomial distribution with parameters n_A and $p = \frac{1}{2}$, and the other compartment will get n_A minus that value.

This whole system will be denoted $\mathcal{F} = (\mathcal{I}_\mathcal{K}, \mathcal{H}_\mathcal{K}, \mu)$.

Remark 2.3.1: Above, we assume that when divisions, i.e., compartment transitions of the form $C \to 2C$, happen, each molecule picks a new compartment uniformly at random. This assumption makes the constructions in this paper easier. However, our proofs only require that the total number of each species across compartments is preserved when each division happens. △

Similar to our network representations for reaction networks, we can specify the above model through a picture of the following form:

$$\mathcal{I}_\mathcal{K} \qquad 0 \underset{\kappa_E}{\overset{\kappa_I}{\rightleftarrows}} C \underset{\kappa_C}{\overset{\kappa_F}{\rightleftarrows}} 2C \qquad \mu \tag{2.5}$$

where "$\mathcal{I}_\mathcal{K}$" is a stand-in for a standard CRN diagram, such as the one in (2.4).

Example 2.3.2: If $\mathcal{I}_\mathcal{K}$ is exactly the network diagrammed in Example 2.1.1 and μ is the point mass with 3 molecules of A and 4 molecules of B, we would write

$$A + B \xrightarrow{10} 0 \underset{\kappa}{\overset{2}{\rightleftarrows}} B \qquad A + 2B \xrightarrow{3} A \qquad 2B \xrightarrow{5} A + B$$

$$0 \underset{\kappa_E}{\overset{\kappa_I}{\rightleftarrows}} C \underset{\kappa_C}{\overset{\kappa_F}{\rightleftarrows}} 2C \qquad \delta_{(3,4)}(a, b)$$

△

See also Example 2.3.8 for another specific example.

There are multiple avenues for generalizations. For example, when a merger occurs it could be that not all the molecules make it into the new compartment, or when a division occurs it could be that some molecules are lost, or there is a non-uniform mechanism for distributing the molecules. Moreover, it could be that the rate of compartment fragmentation or exit depends on the internal state of the compartment. These models all fall

under the more general framework given in [1] and could be studied mathematically in the future if there is a desire, but for the initial development of the mathematics we choose to keep things simpler.

2.3.1 Simulation representation

There are multiple ways to describe a Markov model satisfying the information given in the ingredients $\mathcal{F} = (\mathcal{I}_{\mathcal{K}}, \mathcal{H}_{\mathcal{K}}, \mu)$. The first we give is what we term a "simulation" representation in which we enumerate the compartments and track the counts of the species in each compartment.

The simulation representation will be a Markov chain F^{sim} whose state is a finite vector of elements of $\mathbb{Z}^d_{\geq 0}$, where d, as always, is the number of species. We first describe the model via an example. Afterwards we will provide the mathematical details.

Example 2.3.3: Consider again the model from Example 2.3.2. Suppose that at time T there are 4 compartments, where the first has two A and two B, the second has no A and one B, the third again has two of each, and the last has one A and twelve B. Then the state of the model F^{sim} would be the vector

$$\left(\begin{bmatrix} 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 12 \end{bmatrix} \right).$$

We now suppose that at time T a transition occurs. We first consider four possibilities if the transition is due to a reaction of the compartment model.

- Suppose first that the compartment transition is an inflow event. We will make the convention that the new compartment due to an inflow reaction will always be placed at the end of the vector of states. Hence, because the initial distribution for arriving compartments is a point mass at $(3, 4)$ the new state of the full system is

$$\left(\begin{bmatrix} 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 12 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \end{bmatrix} \right).$$

- Next suppose that the compartment transition is an exit event. In this case we must choose a compartment at random, delete it from the vector, and re-index the other components. Thus, we start by choosing from $\{1, 2, 3, 4\}$, each with probability $1/4$. Suppose that the value 3 is chosen so that the third compartment will be deleted. In this case, the new state of the full system is

$$\left(\begin{bmatrix} 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 12 \end{bmatrix} \right).$$

- Now suppose that the compartment transition is a merger, or coagulation. Now we must select two compartments at random and combine their contents. We will always choose that the combined contents of the compartments will be placed within the compartment with the lower index and will delete the compartment with the higher index. Thus, assuming we choose the compartments indexed 1 and 2, we then merge the first and second compartments and place their contents into compartment 1 (since it has the smaller index of the two chosen) and then delete the second compartment. The resulting state is

$$\left(\begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 12 \end{bmatrix} \right).$$

- Finally, we suppose that the compartment transition is a fragmentation. The procedure will be as follows. We will first choose the index of the compartment that fragments, we then create two new compartments and will then split the contents between these new compartments (with each particular molecule choosing between the new compartments with equal probability). The originally chosen compartment will be deleted and the two new compartments will be placed at the end of the vector of states.

  For example, suppose we choose compartment 3 for fragmentation (which occurs

with probability 1/4). We then split the contents of the original third compartment (four molecules total, 2 of A and 2 of B) uniformly at random between the two new compartments. Suppose for concreteness that we split as $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Then, after deleting the 3rd compartment and adding these two onto the end we have a new state for the full model of

$$\left(\begin{bmatrix} 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 12 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right).$$

It is also possible that the transition at time T was due to a reaction taking place within one of the compartments. For example, if the reaction $A + 2B \to A$ happens inside the fourth compartment, then the state of the whole system, F^{sim}, will become

$$\left(\begin{bmatrix} 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 10 \end{bmatrix} \right).$$

△

Now we give the formal mathematical description of F^{sim}. First, let $\{M_C(t)\}_{t \geq 0}$ be the Markov chain associated to the compartment network $\mathcal{H}_{\mathcal{K}}$. Then $M_C(t)$ will be the number of compartments at time t. Let $\{T_i\}_{i=0}^{\infty}$ be the jump times for this Markov chain, where $T_0 = 0$. For any $i \geq 0$ and any $j = 1, \ldots, M_C(T_i)$, let $\{X_j^i(t)\}_{t \in [T_i, T_{i+1}]}$ be realizations of the Markov chain associated to $\mathcal{I}_{\mathcal{K}}$ with initial distributions (at time T_i) specified below. Suppose that for any i_1, i_2 and j_1, j_2 with either $i_1 \neq i_2$ or $j_1 \neq j_2$, the chains $X_{j_1}^{i_1}$ and $X_{j_2}^{i_2}$ are independent conditional on their initial conditions, and suppose that the initial distributions are chosen in the following manner (which are just formal characterizations of the details provided in the example above):

- If the compartment transition at time T_{i+1} was an inflow event ($0 \to C$), then let $X_j^{i+1}(T_{i+1}) = X_j^i(T_{i+1})$ for $j = 1, \ldots, M_C(T_i)$, and for $j = M_C(T_{i+1}) = M_C(T_i) + 1$ let $X_j^{i+1}(T_{i+1})$ be distributed according to μ, independently of everything in the

past.

- If the compartment transition at time T_{i+1} was an exit event ($C \to 0$), then let J_i be chosen uniformly at random from $\{1, \cdots, M_C(T_i)\}$, independently of everything in the past. Let $X_j^{i+1}(T_{i+1}) = X_j^i(T_{i+1})$ for $j < J_i$, and let $X_j^{i+1}(T_{i+1}) = X_{j+1}^i(T_{i+1})$ for $j \geq J_i$.

- If the compartment transition at time T_{i+1} was a merger, or coagulation, event ($2C \to C$), then let J_i^1 and J_i^2 be chosen uniformly at random from $\{1, \cdots, M_C(T_i)\}$ and $\{1, \cdots, M_C(T_i)\} \setminus \{J_i^1\}$, respectively, independent of everything in the past. Let $X_j^{i+1}(T_{i+1}) = X_j^i(T_{i+1})$ for $j < \max\{J_i^1, J_i^2\}$ with $j \neq \min\{J_i^1, J_i^2\}$, let $X_j^{i+1}(T_{i+1}) = X_{j+1}^i(T_{i+1})$ for $j \geq \max\{J_i^1, J_i^2\}$, and let $X_j^{i+1}(T_{i+1}) = X_{J_i^1}^i(T_{i+1}) + X_{J_i^2}^i(T_{i+1})$ for $j = \min\{J_i^1, J_i^2\}$.

- If the compartment transition at time T_{i+1} was a fragmentation event ($C \to 2C$), then let J_i be chosen uniformly at random from $\{1, \cdots, M_C(T_i)\}$, independently of everything in the past. Let $\{Z_k^i(x) : x \in \mathbb{Z}^d, k = 1, \ldots, d\}$ be a collection of random variables, independent of each other and everything else, with $Z_k^i(x) \sim \text{Binom}\,(0.5, x_k)$. Let $Z^i(x)$ denote the vector $\left(Z_1^i(x), \cdots, Z_d^i(x)\right)$. Let $X_j^{i+1}(T_{i+1}) = X_j^i(T_{i+1})$ for $j < J_i$, let $X_j^{i+1}(T_{i+1}) = X_{j+1}^i(T_{i+1})$ for $j = J_i, \ldots, M_C(T_i) - 1$, and for $j = M_C(T_i)$ let $X_j^{i+1}(T_{i+1}) = Z^i(X_{J_i}^i(T_{i+1}))$ and $X_{j+1}^{i+1}(T_{i+1}) = X_{J_i}^i(T_{i+1}) - X_j^{i+1}(T_{i+1})$.

Let $F^{\text{sim}}(t)$ be the vector $\left(X_1^i(t), X_2^i(t), \cdots, X_{M_C(t)}^i(t)\right)$, where i is such that $T_i \leq t < T_{i+1}$.

Lemma 2.3.4: The process $\{F^{\text{sim}}(t)\}_{t \geq 0}$ is a continuous time Markov chain with state space $\bigcup_{m \geq 0} \left(\mathbb{Z}_{\geq 0}^d\right)^m$, the space of finite tuples of elements of $\mathbb{Z}_{\geq 0}^d$.

Proof. To show that this is a Markov process we have to show that the holding times are exponential and the updates are independent of the holding times. To see that the holding times are exponential, notice that since M_C is a Markov chain it has exponential holding

times, and similarly for each X_j^i. But the holding times for these processes are independent, and the minimum of independent exponential random variables is itself exponential.

Furthermore, the minimum of a (finite) collection of independent exponential random variables is independent of the index at which the minimum occurs, so the updates are indeed independent of the holding times.

The fact that F^{sim} takes values in the space of finite tuples is equivalent to M_C being finite for all time, which in turn is equivalent to the fact that M_C is not explosive, regardless of the choice of rate constants in $\mathcal{H}_{\mathcal{K}}$. This is a standard result in the theory of 1-d mass action stochastic reaction networks; see for instance [24]. □

2.3.2 An explicit construction of the simulation representation

We discuss one way of constructing the model described in Section 2.3.1, in the spirit of the Kurtz representation (2.3). Here, by "construction" we mean an explicit detailing of the random processes and random variables needed to generate a single realization of the process. The construction is of interest since it is amenable to analysis, coupling methods, simulation methods, etc. The construction will be used later in this paper to verify some behaviors of Example 3.2.21.

Let $\mathcal{F} = (\mathcal{I}_{\mathcal{K}}, \mathcal{H}_{\mathcal{K}}, \mu)$ be as above. Suppose that $M_C(0)$ is the initial number of compartments in the system and further suppose that M_C is given as the solution to

$$\begin{aligned} M_C(t) = M_C(0) + Y_I\left(\kappa_I t\right) - Y_E\left(\int_0^t \kappa_E M_C(s)ds\right) + Y_F\left(\int_0^t \kappa_F M_C(s)ds\right) \\ - Y_C\left(\int_0^t \kappa_C \frac{M_C(s)(M_C(s)-1)}{2}\, ds\right), \end{aligned} \tag{2.6}$$

where Y_I, Y_E, Y_F, and Y_C are independent unit-rate Poisson processes. Then M_C is the Markov chain on $\mathbb{Z}_{\geq 0}$ associated to $\mathcal{H}_{\mathcal{K}}$, so that $M_C(t)$ gives the number of compartments at any time $t \geq 0$.

The jump times of the counting processes $R_I(t) = Y_I\left(\kappa_I t\right)$, $R_E(t) = Y_E\left(\int_0^t \kappa_E M_C(s)ds\right)$, $R_F(t) = Y_F\left(\int_0^t \kappa_F M_C(s)ds\right)$, and $R_C(t) = Y_C\left(\int_0^t \kappa_C \frac{M_C(s)(M_C(s)-1)}{2}\, ds\right)$ determine when the RNIC model transitions due to changes in the count of the compartments. To each

such transition we will also require a collection of random variables needed to carry out the updates in the RNIC model. We detail these random variables below.

The construction below may be confusing; the following remark is intended to help clarify.

Remark 2.3.5: At several places in this construction (specifically, when compartment transitions happen), we need to sample from a random variable which is not uniform on $[0,1]$. For example, when a cell inflow happens we need to sample from μ, when a cell exit happens we need to sample uniformly from $1, 2, \cdots, M_C(t)$, etc.. But in each case we define uniform random variables u and use them in the construction. Why does this work? And why do we do it?

This works because a random variable with any distribution can be defined as a function of a uniform $[0,1]$ random variable. For example, considering μ again, if we order $\mathbb{Z}^d_{\geq 0}$ somehow as $x_1, x_2, x_3, \cdots$, let $p_k = \sum_{j=1}^{k} \mu(x_k)$, $k = 0, 1, 2, \cdots$, and let $f(t) = x_k \mathbb{I}_{(p_{k-1}, p_k]}(t)$, then $f(u)$ is distributed according to μ whenever u is uniform.

Why do we do it this way, instead of just asking directly for a collection of random variables with the required distribution? The answer is that we need to fully specify how all random variables depend on each other, and if (for example) we just used uniform random variables on $1, 2, \cdots, M_C(t)$ for compartment exit, then those random variables wouldn't be independent of Y_I, Y_E, Y_F, Y_C, because $M_C(t)$ depends on Y_I, Y_E, Y_F, Y_C. Our approach splits the construction of the random variable for the index of the exiting compartment into two steps: Choosing a distribution to sample from, and sampling from it. The first step is a function of Y_I, Y_E, Y_F, Y_C, and the latter is independent from them, so the dependence is fully described. △

In the description below all random variables are independent of each other and of the Poisson processes Y_I, Y_E, Y_F, Y_C. We require:

- A collection of independent uniform random variables $\{u_i^I\}$ on $[0,1]$, $i = 1, 2, \ldots$. When $R_I(T) - R_I(T-) = 1$, the random variable $u^I_{R_I(T)}$ is used to generate a sample from μ.

- A collection of independent uniform random variables $\{u_i^E\}$ on $[0,1]$, $i = 1, 2, \ldots$. When $R_E(T) - R_E(T-) = 1$, the random variable $u^E_{R_E(T)}$ is used to determine which compartment exits at that time.

- Two collections of independent uniform random variables on $[0,1]$: (i) $\{u_i^F\}$, $i = 1, 2, \ldots$, and (ii) an array $\{\hat{u}^F_{i,j}\}$, $i, j \in \{1, 2, \ldots\}$. When $R_F(T) - R_F(T-) = 1$, the random variable $u^F_{R_F(T)}$ is used to determine which compartment fragments. We then utilize the finite collection $\{\hat{u}^F_{R_F(T),j}\}$, $j = 1, \ldots, M$, where M is the total number of molecules in the chosen compartment, to divide the different molecules between the two new cells.

- A collection of independent uniform random variables $\{u_i^C\}$ on $[0,1]$, $i = 1, 2, \ldots$. When $R_C(T) - R_C(T-) = 1$, the random variable $u^C_{R_C(T)}$ is used to determine which two compartments are chosen to merge.

Note that the collections detailed above are chosen before a realization is generated. Said differently, the realization of the RNIC model is a function of these independent random variables.

All that remains is to give the timing of the different chemical reactions. One method is the following. Let $\{Y_r\}_{r\in\mathcal{R}}$ be a collection of independent (of each other, and all other random objects so far), unit-rate Poisson processes, one for each possible reaction in $\mathcal{I}_\mathcal{K}$. Moreover, for each $r \in \mathcal{R}$, let $\{u_i^r\}$, $i = 1, 2, \ldots$ be a collection of independent uniform random variables. Then, for $r \in \mathcal{R}$, we let

$$R_r(t) = Y_r\left(\sum_{\substack{i\geq 0\\ T_i\leq t}} \int_{T_i}^{T_{i+1}\wedge t} \sum_{j=1}^{M_C(T_i)} \lambda_r(X_j^i(s))ds\right),$$

where the T_i are the jump times of the process M_C, $X_j^i(s)$ is the state of the process in compartment j at time s, and λ_r is given as in (2.2). Then R_r is the counting process that jumps by $+1$ when the rth reaction takes place in *some* compartment. When $R_r(T) - R_r(T-) = 1$, meaning a reaction has taken place somewhere, we use $u^r_{R_r(T)}$ to determine

the compartment within which the reaction took place. In particular, the probability that it took place in compartment k is simply

$$\frac{\lambda_r(X_k^i(T-))}{\sum_{j=1}^{M_C(T_i)} \lambda_r(X_j^i(T-))}.$$

2.3.3 A coarse-grained representation

While the description (and construction) above is often convenient for the sake of analysis and simulation, it is sometimes not the most natural way to think about these models. For example, suppose we have a model with a single species, denoted S, and for which there are two compartments at time t, so that $M_C(t) = 2$. It is reasonable to think that we would not care to distinguish the situation in which there are 6 molecules of species S in the first compartment and 2 in the second, which is the state $(6, 2)$, versus the situation of 2 molecules of S in the first and 6 in the second, which is the state $(2, 6)$. In this situation, we would simply care that we have one compartment with two S molecules, another with six, and there are no other compartments.

To handle this, we consider a function $n : \mathbb{Z}_{\geq 0} \to \mathbb{Z}_{\geq 0}$ in which $n_x := n(x)$ gives the number of compartments present with precisely x molecules of S (hence the notation that "n" gives the $\underline{n}$umber of compartments with different counts). In this case, the state of the example system described above would simply be the function with

$$n_x = \begin{cases} 1, & \text{if } x = 2 \\ 1, & \text{if } x = 6 \\ 0, & \text{else.} \end{cases}$$

Note that in this one-dimensional case we can also think of n as an "infinite vector." For example, in our example above we would have

$$n = (0, 0, 1, 0, 0, 0, 1, 0, 0, \dots),$$

with only zeros continuing on.

For another example, we could consider the case discussed in Example 2.3.3, where there are two species A and B and the state for the simulation representation was

$$\left(\begin{bmatrix} 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 12 \end{bmatrix} \right).$$

In this case, the state could naturally be described by the function

$$n_x = \begin{cases} 2, & \text{if } x = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \\ 1, & \text{if } x = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ 1, & \text{if } x = \begin{bmatrix} 1 \\ 12 \end{bmatrix} \\ 0, & \text{else.} \end{cases}$$

Note that in this example, it is not natural to view n as an "infinite vector." Instead, it would be natural to view it as an "infinite array" with a two in the $(2,2)$ component, ones in the $(0,1)$ and $(1,12)$ components, and zeros elsewhere.

Thus, we may take the following approach, as done in [1]. The state space of the coarse-grained model will be

$$\begin{aligned} \mathcal{N} &:= \{\text{functions } n : \mathbb{Z}^d_{\geq 0} \to \mathbb{Z}_{\geq 0} \text{ with compact support}\} \\ &= \{\text{functions } n : \mathbb{Z}^d_{\geq 0} \to \mathbb{Z}_{\geq 0} \text{ with finite support}\} \\ &= \{\text{functions } n : \mathbb{Z}^d_{\geq 0} \to \mathbb{Z}_{\geq 0} \text{ with finite } \ell^1 \text{ norm}\}, \end{aligned} \tag{2.7}$$

where we observe that all three sets are the same. Given $n : \mathbb{Z}^d_{\geq 0} \to \mathbb{Z}_{\geq 0}$, we write

$n = (n_x)_{x \in \mathbb{Z}^d_{\geq 0}}$. For each possible state $x \in \mathbb{Z}^d_{\geq 0}$ of the chemistry, $n_x \in \mathbb{Z}_{\geq 0}$ represents the number of compartments whose chemistry has that particular state. Given Markov chains M_C and X^i_j as defined in Section 2.3.1, let N be the process where $N_x(t)$ is the number of compartments in state $x \in \mathbb{Z}^d_{\geq 0}$ at time $t \geq 0$:

$$N_x(t) = \sum_{i=0}^{\infty} \mathbb{I}\{t \in [T_i, T_{i+1})\} \sum_{j=1}^{M_C(T_i)} \mathbb{I}\{X^i_j(t) = x\}.$$

Note that the total number of compartments at time $t \geq 0$ can be recovered from $N(t)$ via

$$M_C(t) = \|N(t)\|_{\ell^1} := \sum_{x \in \mathbb{Z}^d_{\geq 0}} N_x(t).$$

Note also that the process N transitions iff F^{sim} does. This fact is important enough that we state it as a lemma:

Lemma 2.3.6: Let F^{sim} and N be as above. Then N undergoes a transition at time t iff F^{sim} does.

Proof. On the one hand, N is defined as a function of F^{sim} and so N cannot transition if F^{sim} does not. On the other hand, all possible transitions of F^{sim} cause a change in N: If F^{sim} transitions because M_C does, then $\|N\|_{\ell^1} = M_C$ changes, whereas if F^{sim} changes otherwise then the contents of some single compartment updated, which changes N. □

For the lemma below, we recall that for $x \in \mathbb{Z}^d_{\geq 0}$ we define e_x to be the function taking the value of one at x and zero otherwise.

Lemma 2.3.7: Let $N(t)$ be as defined above. Then $\{N(t)\}_{t \geq 0}$ is a Markov chain taking values in $\mathcal{N}$, defined in (2.7). Moreover, for $n \in \mathcal{N}$, the transitions rates are as follows:

Transition type		Rate
Compartment inflow	$n \mapsto n + e_x$	$\kappa_I \mu(x)$
Compartment exit	$n \mapsto n - e_x$	$\kappa_E n_x$
Compartment coagulation, $x \neq y$	$n \mapsto n + e_{x+y} - e_x - e_y$	$\kappa_C n_x n_y$
Compartment coagulation	$n \mapsto n + e_{2x} - 2e_x$	$\kappa_C \binom{n_x}{2}$
Compartment fragmentation ($x = y$ allowed here)	$n \mapsto n - e_{x+y} + e_x + e_y$	$\kappa_F n_{x+y} \varphi(x+y, x)$
Internal reaction $r \in \mathcal{R}$	$n \mapsto n - e_x + e_{x+\nu_r' - \nu_r}$	$n_x \kappa_r \binom{x}{\nu_r}$

where

$$\varphi(z, x) := \prod_{k=1}^{d} \binom{z_k}{x_k} 2^{-z_k}$$

so that the distribution of the resulting compartments after a fragmentation is independently binomial in each species. Note that each row mentioning x or y corresponds to an infinite family of transitions and in the last row $r \in \mathcal{R}$ also ranges over all reactions of the reaction network $\mathcal{I}$.

Proof. The fact that N has finite support follows from the fact that F^{sim} is always a finite tuple, proved in Lemma 2.3.4.

The fact that N is Markovian with the rates given follows from consideration of the infinitesimal behavior of F^{sim}. For example, for $x \neq y \in \mathbb{Z}^d_{\geq 0}$,

$$\mathbb{P}(N(t+h) = n + e_{x+y} - e_x - e_y | N(t) = n) = \kappa_C n_x n_y h + o(h), \text{ as } h \to 0,$$

since, to leading order, the probability that some compartment in state x merges with a compartment in state y in the time interval $[t, t+h)$ is $\kappa_C n_x n_y h$. The other rows of the table follow similarly. □

Example 2.3.8: Consider the following possible compartment model:

$$0 \underset{\kappa_d}{\overset{\kappa_b}{\rightleftarrows}} S \qquad\qquad 0 \underset{\kappa_E}{\overset{\kappa_I}{\rightleftarrows}} C \xleftarrow{\kappa_C} 2C \qquad\qquad \tfrac{1}{2}\delta_5 + \tfrac{1}{2}\delta_{17}$$

Here we are keeping track of some chemical S which forms with rate κ_b and degrades with rate κ_d. Compartments are allowed to enter with rate κ_I, and new compartments that enter this way have either 5 or 17 molecules of S, each with probability $1/2$. Compartments can also exit with rate constant κ_E, and merge (or coagulate) with rate constant κ_C. Since there is only one species, the state space for the chemistry is $\mathbb{Z}^1_{\geq 0} = \mathbb{Z}_{\geq 0}$. As we detail below, we will be assuming mass-action kinetics; in this case that means when the model is in state $n \in \mathcal{N}$ the transition rates are given by

Transition type		Rate
Compartment inflow	$n \mapsto n + e_5$	$\kappa_I/2$
Compartment inflow	$n \mapsto n + e_{17}$	$\kappa_I/2$
Compartment exit	$n \mapsto n - e_x$	$\kappa_E n_x$
Compartment coagulation ($x \neq y$)	$n \mapsto n + e_{x+y} - e_x - e_y$	$\kappa_C n_x n_y$
Compartment coagulation	$n \mapsto n + e_{2x} - 2e_x$	$\kappa_C \binom{n_x}{2}$
S birth	$n \mapsto n - e_x + e_{x+1}$	$\kappa_b n_x$
S death	$n \mapsto n - e_x + e_{x-1}$	$\kappa_d n_x x$

As before, each row mentioning x or y corresponds to an infinite family of transitions, one for each $x \neq y \in \mathbb{Z}^d_{\geq 0}$, and as always e_x is the unit vector in direction x. △

Chapter 3

Non-Explosivity, Recurrence, and Positive Recurrence

3.1 Non-Explosivity

A Markov Chain is explosive if it can undergo infinitely many transitions in finite time. The formal definition is below; see e.g. [17] for more.

Definition 3.1.1 (Explosivity)**:** Let $\{X(t)\}_{t\geq 0}$ be a continuous-time Markov chain with countable state space $\mathbb{S}$. For each $m \in \mathbb{Z}_{\geq 0}$, let τ_m be the time of the m-th transition of X (formally, $\tau_0 = 0$ and $\tau_m = \inf\{t > \tau_{m-1} : X(t) \neq X(\tau_{m-1})\}$), and let $\tau_\infty = \lim_{m\to\infty} \tau_m$. We say that X *explodes* if $\tau_\infty < \infty$. If there is some state $x \in \mathbb{S}$ such that with positive probability X explodes when started in state x, we say that X is *explosive*.

We will show that explosivity for the RNIC model $\mathcal{F} = (\mathcal{I}_\mathcal{K}, \mathcal{H}_\mathcal{K}, \mu)$ is determined by explosivity for the internal reaction network $\mathcal{I}_\mathcal{K}$. But to even talk about explosivity for $\mathcal{F}$ instead of just the Markov chains F^{sim} or N, we need the following simple proposition.

Proposition 3.1.2: Suppose we have a RNIC $\mathcal{F} = (\mathcal{I}_\mathcal{K}, \mathcal{H}_\mathcal{K}, \mu)$. Let F^{sim} and N be the corresponding simulation and coarse-grained representations. Then F^{sim} is explosive iff N is explosive.

Proof. This is immediate from lemma 2.3.6, which says that F^{sim} and N transition at the same times. □

In light of the proposition, we will speak merely of $\mathcal{F} = (\mathcal{I}_{\mathcal{K}}, \mathcal{H}_{\mathcal{K}}, \mu)$ being explosive, and check the explosivity of either F^{sim} or N depending on convenience. As it turns out, it will be most convenient to check explosivity for F^{sim}. (Indeed, the fact that explositivity is more easily checked for F^{sim} is one of the major reasons for introducing F^{sim} in the first place.)

Theorem 3.1.3: Suppose we have a RNIC $\mathcal{F} = (\mathcal{I}_{\mathcal{K}}, \mathcal{H}_{\mathcal{K}}, \mu)$. Then $\mathcal{F}$ is explosive iff $\mathcal{I}_{\mathcal{K}}$ is explosive.

Proof. First, suppose that $\mathcal{I}_{\mathcal{K}}$ is explosive. As discussed above, we intend to show that F^{sim} is explosive. By assumption, there is some $x \in \mathbb{Z}^d$ such that when the Markov chain corresponding to $\mathcal{I}_{\mathcal{K}}$ is started in state x it explodes with positive probability. In particular, there is some finite (nonrandom) time t so that the chemistry undergoes infinitely many transitions before time t with positive probability. Start F^{sim} in the state with one compartment whose state is x. With positive probability, no compartment transitions happen before time t. But the compartment transition times are independent of what is happening inside them by construction, and the compartment evolves according to $\mathcal{I}_{\mathcal{K}}$, so on the event that no compartment transition happens before time t the compartment undergoes infinitely many transitions before time t with positive probability. It follows that F^{sim} is explosive.

Conversely, suppose that $\mathcal{I}_{\mathcal{K}}$ is not explosive. Note that $\mathcal{H}$, the compartment network, is not explosive for any choice of rate constants (see e.g. [24]). So with probability one F^{sim} undergoes only finitely many compartment transitions in finite time. But between each pair of consecutive compartment transitions there are finitely many compartments each evolving according to $\mathcal{I}_{\mathcal{K}}$, and by assumption each of these undergoes only finitely many reactions in finite time a.s.. It follows that F^{sim} undergoes only finitely many transitions total in finite time, and hence is not explosive. □

3.2 Transience, recurrence, and positive recurrence

The following definitions are standard. For example, see [17].

Definition 3.2.1: Let M be a Markov chain with countable state space $\mathbb{S}$, and for $x \in \mathbb{S}$ let $T_x = \inf\{t > 0 : M_t = x \text{ but } \exists s \in [0,t], M_s \neq x\}$ be the first time the process returns to x (or just arrives at x, if the process does not start from x). If $\mathbb{P}_x(T_x < \infty) = 1$, we say that the state x is *recurrent*, and if $\mathbb{E}_x(T_x) < \infty$ we say that the state x is *positive recurrent*. A state which is not recurrent is called *transient*, and a recurrent state which is not positive recurrent is *null recurrent*. If $\mathbb{P}_x(T_y < \infty) > 0$ we say that y is *reachable* from x. If every state $x \in \mathbb{S}$ is positive recurrent, null recurrent, or transient, we say M is *positive recurrent*, *null recurrent*, or *transient*, respectively.

A standard fact about (positive) recurrence is that it is a class property:

Proposition 3.2.2 (Theorems 3.4.1(iv) and 3.5.3(i) $\Longleftrightarrow$ (ii) in [17])**:** Suppose that y is reachable from x and x is recurrent (resp. positive recurrent). Then y is recurrent (resp. positive recurrent).

In other words, if you can get between x and y with positive probability (in both directions), then x and y are either both transient, both null recurrent, or both positive recurrent. So for irreducible chains (ones where you can pass between any two points of the state space with positive probability), the chain M is always positive recurrent, null recurrent, or transient.

Before proceeding with the theory, we summarize the results of this section with a table. The way to read Table 3.1 is as follows:

- Suppose we have a RNIC $(\mathcal{I}_\mathcal{K}, \mathcal{H}_\mathcal{K}, \mu)$, and N is the associated coarse-grained model.
- The top row indicates possible dynamics (transient, null recurrent, or positive recurrent) for $\mathcal{I}_\mathcal{K}$, the chemical model, and the left column indicates possible dynamics for $\mathcal{H}_\mathcal{K}$, the compartment model. Since the possible dynamics for N will turn out to depend crucially on whether the compartments can exit ($\kappa_E > 0$) or not ($\kappa_E = 0$), the left column is further subdivided along these lines.

- Several cells are marked "Impossible", because $\mathcal{H}_\mathcal{K}$ cannot be null recurrent if $\kappa_E = 0$.
- The numbers inside each cell refer to the relevant theorems, lemmas, or examples that demonstrates the result.

		Chemistry ($\mathcal{I}_\mathcal{K}$)		
Compartments ($\mathcal{H}_\mathcal{K}$)		Transient (Trans.)	Null Recurrent (NR)	Positive Recurrent (PR)
Transient		N must be Transient Remark 3.2.4		
NR	$\kappa_E = 0$	Impossible Lemma 3.2.3		
	$\kappa_E > 0$	N must be Null Recurrent Theorem 3.2.5		
PR	$\kappa_E = 0$	N can be Trans. Ex 3.2.9 N can be PR Ex 3.2.11	N can be Trans. Ex 3.2.19 N can be PR Ex 3.2.13	N can be Trans. Ex 3.2.15, 3.2.17, 3.2.21 N can be PR Ex 3.2.17
	$\kappa_E > 0$	N must be Positive Recurrent Theorem 3.2.5		

Table 3.1: The possibly dynamics for N, classified in terms of the dynamics for $\mathcal{H}_\mathcal{K}$ and $\mathcal{I}_\mathcal{K}$. In the above "NR" and "PR" stand for "null recurrent" and "positive recurrent", respectively, whereas "Trans." stands for "transient."

Note that in all cases where we give an example of a recurrent N, the example is actually positive recurrent. We suspect that null recurrent examples will also exist, but we felt it more interesting to cover the behavioral extremes.

Moving to our theory, we begin by considering the dynamics of the compartment model of section 2.2, which takes the form of a relatively simple reaction network, namely,

$$0 \underset{\kappa_E}{\overset{\kappa_I}{\rightleftarrows}} C \underset{\kappa_C}{\overset{\kappa_F}{\rightleftarrows}} 2C \tag{3.1}$$

The (positive) recurrence of this model is already completely classified; see e.g. [24]. We state this classification now as a lemma.

Lemma 3.2.3: Consider the CRN in (3.1).

- Suppose $\kappa_I = 0$. Then 0 is an absorbing state. If some other rate constant is non-zero then all other states are transient, whereas if all four rate constants are zero

then all states are absorbing.

- Suppose $\kappa_I > 0$ and $\kappa_E > 0$. The irreducible state space is $\{0, 1, 2, \dots\}$ and:
  - If $\kappa_C > 0$, then the chain is positive recurrent.
  - If $\kappa_C = 0$ but $\kappa_F < \kappa_E$, then the chain is positive recurrent.
  - If $\kappa_C = 0$ and $\kappa_F > \kappa_E$, then the chain is transient.
  - If $\kappa_C = 0$ and $\kappa_F = \kappa_E$, then either $\kappa_I \leq \kappa_E$ and the chain is null recurrent, or $\kappa_I > \kappa_E$ and the chain is transient.
- Suppose $\kappa_I > 0$ and $\kappa_E = 0$. Then all statements remain the same as in the case $\kappa_I > 0$ and $\kappa_E > 0$ except the irreducible state space is now $\{1, 2, \dots\}$ (and the state 0 is transient).

Now we begin with our positive results. The first fact is simple enough to be stated as a remark:

Remark 3.2.4: Notice that if N is the course-grained representation for $\mathcal{F} = (\mathcal{I}_\mathcal{K}, \mathcal{H}_\mathcal{K}, \mu)$ and n is a (positive) recurrent state for N, then the number of compartments in n, $\|n\|_{\ell^1}$, is a (positive) recurrent state for $\mathcal{H}_\mathcal{K}$, since the return time to $\|n\|_{\ell^1}$ is bounded by the return time to n. △

Said succinctly, if n is a positive recurrent state of the full model, then so is $\|n\|_{\ell^1}$ for the compartment model. One might hope that the converse would be true, and it turns out under relatively mild assumptions it is:

Theorem 3.2.5: Consider a non-explosive model $\mathcal{F} = (\mathcal{I}_\mathcal{K}, \mathcal{H}_\mathcal{K}, \mu)$ where $\kappa_E > 0$, and let N be its course-grained representation. Then a state n is (positive) recurrent for N iff n is reachable from the empty state $\vec{0}$ for N and the state $\|n\|_{\ell^1}$ is (positive) recurrent for $\mathcal{H}_\mathcal{K}$.

Proof. If $\kappa_I = 0$ the conclusions of the theorem are clear, since by Lemma 3.2.3 the state with no compartments is absorbing for both N and $\mathcal{H}_\mathcal{K}$ and all other states are transient. From here on we assume $\kappa_I > 0$.

Let $M_C = \|N\|_{\ell^1}$ be the number of compartments; recall that M_C is a Markov chain which evolves according to $\mathcal{H}_\mathcal{K}$. Suppose first that n is recurrent for N. By Remark 3.2.4, $\|n\|_{\ell^1}$ is recurrent for $\mathcal{H}_\mathcal{K}$. Since $\kappa_E > 0$ and $\kappa_I > 0$, by Lemma 3.2.3 $\mathcal{H}_\mathcal{K}$ is irreducible, so $\mathcal{H}_\mathcal{K}$ eventually hits zero with probability one when started from $\|n\|_{\ell^1}$. But when M_C hits zero, $N = \vec{0}$. Since n is recurrent for N, it must be that N eventually returns to state n after hitting state $\vec{0}$. This proves that n is reachable from $\vec{0}$ for N.

Now suppose that n is reachable from $\vec{0}$ and the state $\|n\|_{\ell^1}$ is positive recurrent (resp. recurrent) for $\mathcal{H}_\mathcal{K}$. Since $\mathcal{H}_\mathcal{K}$ is irreducible as in the previous paragraph, it follows that zero is positive recurrent (resp. recurrent) for $\mathcal{H}_\mathcal{K}$. But $N = \vec{0}$ exactly when M_C is 0, so $\vec{0}$ is positive recurrent (resp. recurrent) for N. But positive recurrence (resp. recurrence) is a class property and by assumption n is reachable from $\vec{0}$, so we conclude that n is positive recurrent (resp. recurrent) for N, as desired. □

The same theorem holds, *mutatis mutandis*, for F^{sim}. The proof is the same so we omit it.

Theorem 3.2.6: Consider a non-explosive model $\mathcal{F} = (\mathcal{I}_\mathcal{K}, \mathcal{H}_\mathcal{K}, \mu)$ where $\kappa_E > 0$, and let F^{sim} be its simulation representation. Then a state $(x_1, \ldots, x_k)$ is (positive) recurrent for F^{sim} iff $(x_1, \cdots, x_k)$ is reachable from the empty vector () for F^{sim} and the state k is (positive) recurrent for $\mathcal{H}_\mathcal{K}$.

Remark 3.2.7: Theorems 3.2.5 and 3.2.6 hold under more general assumptions. Note that the key idea of both is that 0 is (positive) recurrent for $\mathcal{H}_\mathcal{K}$. Hence, one can generalize to the situation in which $\mathcal{F} = (\mathcal{I}_\mathcal{K}, \mathcal{H}_\mathcal{K}, \mu)$ has non-mass action kinetics for either $\mathcal{I}_\mathcal{K}$ or $\mathcal{H}_\mathcal{K}$, so long as the system is non-explosive and 0 is (positive) recurrent for $\mathcal{H}_\mathcal{K}$. △

3.2.1 Instructive examples

We now consider some examples. The first is an application of Theorem 3.2.5, and the rest show the various ways the conclusion of the theorem can fail if the hypothesis $\kappa_E > 0$ is not satisfied. These examples also serve to illustrate various techniques that are useful

for analysing recurrence and transience of RNIC models. In Example 3.2.11, positive recurrence for the RNIC is shown via a Lyapunov function, applying Theorem 6.1.2. In Example 3.2.19, transience for the RNIC is shown via a Lyapunov function, applying Theorem 6.1.3. And in Example 3.2.21, transience for the RNIC is shown with the help of the construction of F^{sim} given in section 2.3.2.

In the following, any rate constants not specified are assumed to be positive.

Example 3.2.8: Consider the following RNIC.

$$0 \xrightarrow{\kappa_b} 2S \qquad\qquad 0 \underset{\kappa_E}{\overset{\kappa_I}{\rightleftarrows}} C \xleftarrow{\kappa_C} 2C \qquad\qquad \delta_0$$

where δ_0 is the point mass at zero (so each compartment enters empty). Even though $\mathcal{I}_{\mathcal{K}}$ is transient, by Theorem 3.2.5 the empty state is positive recurrent for N. Any state where every compartment has an even number of S molecules is reachable from the empty state, hence positive recurrent. Any state where any compartment has an odd number of S molecules is not reachable from the empty state, hence transient. △

In all of the remaining examples in this section, we have $\kappa_E = 0$ and hence the state 0 will be transient for $\mathcal{H}_{\mathcal{K}}$. Hence, when discussing the properties of the model we restrict ourselves to the state space $\mathcal{N}\backslash\{0\}$ that does not include the state with zero compartments.

The case where $\kappa_E = 0$ is more complicated than the $\kappa_E \neq 0$ case. For one thing, it is no longer enough just to look at $\mathcal{H}_{\mathcal{K}}$ to decide if all states are transient. Indeed, if Example 3.2.8 is modified so that $\kappa_E = 0$ then every state becomes transient, despite the fact that all states are positive recurrent for the compartment network $\mathcal{H}_{\mathcal{K}}$:

Example 3.2.9: Consider the model $\mathcal{F} = (\mathcal{I}_{\mathcal{K}}, \mathcal{H}_{\mathcal{K}}, \mu)$ described by

$$0 \xrightarrow{\kappa_b} 2S \qquad\qquad 0 \xrightarrow{\kappa_I} C \xleftarrow{\kappa_C} 2C \qquad\qquad \delta_0 \tag{3.2}$$

where δ_0 is again the point mass at zero.

We reiterate that this is exactly the same as the previous example but with κ_E set to zero. However, that is enough to make every state transient for $\mathcal{F}$:

Proposition 3.2.10: In the RNIC model (3.2), $\mathcal{I}_\mathcal{K}$ is transient, $\mathcal{H}_\mathcal{K}$ is positive recurrent on the irreducible state space $\{1, 2, \dots\}$, and N (the coarse-grained model corresponding to $\mathcal{F}$) is transient.

Proof. Except for the zero-compartment state (which cannot be returned to), all states are positive recurrent for $\mathcal{H}_\mathcal{K}$ by Lemma 3.2.3. However, the total number of S molecules across all compartments can never shrink, and grows with some positive rate (at least κ_b, and larger if there are more compartments), so all states are transient for N. □

Thus we see that, in this example, the long-term behavior of $\mathcal{H}_\mathcal{K}$ and the course-grained model N are different. △

The above example shows that when $\kappa_E = 0$ and $\mathcal{I}_\mathcal{K}$ is transient, $\mathcal{F}$ may be transient even if $\mathcal{H}_\mathcal{K}$ is not. However, this need not always be the case. Below we have an example that demonstrates that, when $\kappa_E = 0$ and $\mathcal{I}_\mathcal{K}$ is transient, it is still possible for $\mathcal{F}$ to be positive recurrent.

Example 3.2.11: Consider the model $\mathcal{F} = (\mathcal{I}_\mathcal{K}, \mathcal{H}_\mathcal{K}, \mu)$ described by

$$2A + B \xrightarrow{1} B \xrightarrow{1} 0 \xrightarrow{1} A \qquad 0 \xrightarrow{1} C \xleftarrow{6} 2C \qquad \delta_{(0,1)}(a, b) \quad (3.3)$$

where $\delta_{(0,1)}$ is a point mass with zero A molecules and one B molecule. We will show that the chemical model $\mathcal{I}_\mathcal{K}$ is transient but that the course-grained model, N, is positive recurrent. Intuitively, this can be understood in the following manner: B should be thought of as an enzyme that degrades the substrate A. Without the compartment model, the enzyme would simply disappear over time, and then the substrate would grow without bound (from the reaction $0 \to A$). However, each compartment brings in a new enzyme allowing for the further degradation of A.

Proposition 3.2.12: In the RNIC model (3.3), $\mathcal{I}_\mathcal{K}$ is transient, $\mathcal{H}_\mathcal{K}$ is positive recurrent on the irreducible state space $\{1, 2, \dots\}$, and N (the coarse-grained model corresponding to $\mathcal{F}$) is positive recurrent.

Proof. $\mathcal{H}_{\mathcal{K}}$ is positive recurrent by Lemma 3.2.3. $\mathcal{I}_{\mathcal{K}}$ is transient by the discussion above.

It just remains to check positive recurrence of N. For $n \in \mathcal{N}$, let $C(n) = \|n\|_{\ell^1} = \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} n_{(a,b)}$ denote the number of compartments, and let $A(n) = \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} an_{(a,b)}$ and $B(n) = \sum_{a=0}^{\infty}\sum_{b=0}^{\infty} bn_{(a,b)}$ be the total number of A and B molecules, respectively, across all compartments. Define $V : \mathcal{N} \to [0, \infty)$ via

$$V(n) = \begin{cases} A(n) + B(n) + 5C(n) - 1 & B(n) \neq 0 \\ A(n) + B(n) + 5C(n) + 7 & B(n) = 0. \end{cases}$$

We claim that this is a Lyapunov function for N. An upper bound for $\mathcal{L}V(n)$, the generator applied to V at n, is given by

$$\mathcal{L}V(n) \leq \begin{cases} -B(n) + 7 - 15C(n)(C(n) - 1) & B(n) \geq 2 \text{ and } C(n) \geq 2 \\ 14 - 15C(n)(C(n) - 1) & B(n) = 1 \text{ and } C(n) \geq 2 \\ -1 - 15C(n)(C(n) - 1) & B(n) = 0 \\ -A(n)(A(n) - 1)B(n) - B(n) + 7 & B(n) \geq 2 \text{ and } C(n) = 1 \\ -A(n)(A(n) - 1) + 14 & B(n) = 1 \text{ and } C(n) = 1 \end{cases}$$

Note that the first two rows are upper bounds and the last three rows are exact. Specifically, in the first two rows we neglected the contribution of the $2A + B \to B$ reaction — unlike everything else it crucially depends on how the A and B molecules are distributed across the compartments.

We see that $\mathcal{L}V(n) \leq -1$ for all n outside a finite set of states—for instance, you could take the states where there is exactly one compartment and it has at most 7 B and at most 4 A. So V is indeed a Lyapunov function for N, and hence N is positive recurrent by Theorem 6.1.2. □

△

In the previous example we saw that even when $\kappa_E = 0$, positive recurrent compart-

ments $\mathcal{H}_\mathcal{K}$ can still tame transient chemistry $\mathcal{I}_\mathcal{K}$. It should not be surprising, then, that positive recurrent compartments can tame null recurrent chemistry in the same manner. For the sake of filling in Table 3.1 completely, we present a modification of Example 3.2.11 where $\mathcal{I}_\mathcal{K}$ is null recurrent instead of transient.

Example 3.2.13: Consider the model $\mathcal{F} = (\mathcal{I}_\mathcal{K}, \mathcal{H}_\mathcal{K}, \mu)$ described by

$$2A + B \xrightarrow{1} B \xrightarrow{1} 0 \underset{1}{\overset{1}{\rightleftarrows}} A \xrightarrow{1} 2A \qquad 0 \xrightarrow{1} C \xleftarrow{6} 2C \qquad \delta_{(0,1)}(a,b) \tag{3.4}$$

where $\delta_{(0,1)}$ is a point mass with zero A molecules and one B molecule.

The verification of this example is similar enough to that of Example 3.2.11 that we provide only a sketch.

Proposition 3.2.14: In the RNIC model (3.4), $\mathcal{I}_\mathcal{K}$ is null recurrent on the irreducible state space $\{0, 1, 2, \dots\} \times \{0\}$, $\mathcal{H}_\mathcal{K}$ is positive recurrent on the irreducible state space $\{1, 2, \dots\}$, and N (the coarse-grained model corresponding to $\mathcal{F}$) is positive recurrent.

Proof Sketch. Similarly to Example 3.2.11, $\mathcal{H}_\mathcal{K}$ is positive recurrent and $\mathcal{I}_\mathcal{K}$ is eventually reduces (after all the B molecules degrade) to the network

$$0 \underset{1}{\overset{1}{\rightleftarrows}} A \xrightarrow{1} 2A.$$

This model is null recurrent by Lemma 3.2.3.

As for N, let V be the very same Lyapunov function used to prove positive recurrence in Example 3.2.11. The only difference between this example and that one is the addition of the reactions $A \to 0$ and $A \to 2A$. But notice that the contribution of $A \to 0$ in $\mathcal{L}V(n)$ is $-A(n)$, and the contribution of $A \to 2A$ is $A(n)$. These are equal and opposite, so $\mathcal{L}V(n)$ is exactly the same in this example and Example 3.2.11. Thus the remainder of the proof is identical. □

△

Examples 3.2.9 and 3.2.11 showed that $\mathcal{F} = (\mathcal{I}_\mathcal{K}, \mathcal{H}_\mathcal{K}, \mu)$ can be either positive recurrent or transient when $\kappa_E = 0$ and $\mathcal{I}_\mathcal{K}$ is transient. The next few examples are dedicated to

showing the same when $\mathcal{I}_{\mathcal{K}}$ is recurrent. First, if new compartments enter with a huge number of molecules, it can overwhelm otherwise positive recurrent chemistry:

Example 3.2.15: Consider the RNIC model $\mathcal{F} = (\mathcal{I}_{\mathcal{K}}, \mathcal{H}_{\mathcal{K}}, \mu)$ described by

$$0 \underset{\kappa_d}{\overset{\kappa_b}{\rightleftarrows}} S \qquad\qquad 0 \xrightarrow{\kappa_I} C \underset{\kappa_C}{\overset{\kappa_F}{\rightleftarrows}} 2C \qquad\qquad \mu, \tag{3.5}$$

where μ is not yet specified.

Proposition 3.2.16: Let N be the coarse-grained model associated with the RNIC model (3.5). For any choice of non-negative rate constants such that $\kappa_I > 0$, there is a distribution μ on the non-negative integers such that N is transient.

Proof. We will show that in the case $\kappa_b = 0$, μ can be chosen so that the total number of S molecules is itself a transient Markov chain. The case of $\kappa_b > 0$ then immediately follows by a coupling argument. That portion of the proof is straightforward and is omitted.

Let $M(t)$ denote the number of S molecules across all compartments at time t. Under the assumption that $\kappa_b = 0$, M is a Markov chain which transitions from state $m \in \mathbb{N}$ to state $m-1$ with rate $\kappa_d m$ and to state $m+j$ with rate $\kappa_I \mu(j)$.

Our plan is the following: we will recursively define an increasing sequence of integers m_k for $k = 1, 2, 3, \ldots$, and define $\mu(m_k) = 2^{-k}$ and $\mu(j) = 0$ otherwise. For $k = 2, 3, 4, \ldots$, we will let A_k denote the event that the process M reaches m_{k-1} before it reaches (or exceeds) m_{k+1}. It then suffices to show that $\sup_k \mathbb{P}_{m_k}(A_k) < 1/2$ to prove transience of M.

Continuing, we begin by letting $m_1 = 0$. Now suppose $m_1, \ldots, m_{k-1}$ have been defined. We will show that for any $\varepsilon > 0$ it is possible to pick m_k so that $\mathbb{P}_{m_k}(A_k) < \varepsilon$ regardless of the values chosen for $m_{k+1}, m_{k+2}, \ldots$. To show this, we make the following observations.

1. Since M can only go down by one at a time, to get from m_k to m_{k-1} before hitting a state equal to or larger than m_{k+1}, the process must visit every state $m_k, m_k - 1, \cdots, m_{k-1} + 1$ at least once.

2. On the event A_k, during each visit to each of the states $m_{k-1}+1, \ldots, m_k$ there was no transition of size $+m_{k+1}$ (for in that case the state of M would would necessarily reach or exceed m_{k+1}).

The probability of the process M transitioning up by m_{k+1} while in state m is $\frac{2^{-(k+1)}\kappa_I}{\kappa_I+\kappa_d m}$ because the total rate out of state m is $\kappa_I + \kappa_d m$, and the rate of inflows of size m_{k+1} in state m is $\mu(m_{k+1})\kappa_I = 2^{-k-1}\kappa_I$. Hence, combining the above observations we see

$$\begin{aligned}\mathbb{P}_{m_k}(A_k) &\leq \prod_{m=m_{k-1}+1}^{m_k} \left(1 - \frac{2^{-(k+1)}\kappa_I}{\kappa_I + \kappa_d m}\right) \\ &\leq \prod_{m=m_{k-1}+1}^{m_k} \exp\left(-\frac{2^{-(k+1)}\kappa_I}{\kappa_I + \kappa_d m}\right) = \exp\left(-2^{-(k+1)}\kappa_I \sum_{m=m_{k-1}+1}^{m_k} \frac{1}{\kappa_I + \kappa_d m}\right),\end{aligned}$$

where above we use the bound $1 - x \leq e^{-x}$.

If m_{k-1} is fixed and we send $m_k \to \infty$ in the sum above, we get ∞ (it's a tail of a harmonic series). Therefore, $\mathbb{P}_{m_k}(A_k)$ can be made as small as we like by choosing m_k big enough. We conclude that for appropriate choice of m_k, the process M is transient, and hence so is N. □

Hence, so long as $\kappa_E = 0$, a distribution μ that is "bad enough" can cause the whole model to be transient even if the chemical model $\mathcal{I}_\mathcal{K}$ is positive recurrent. △

In the previous example, the distribution μ of incoming compartments was unbounded. As it turns out, $\mathcal{F} = (\mathcal{I}_\mathcal{K}, \mathcal{H}_\mathcal{K}, \mu)$ can be transient even when $\mathcal{I}_\mathcal{K}$ and $\mathcal{H}_\mathcal{K}$ are positive recurrent and μ is bounded. The simplest, though not only, reason this can occur is the existence of some conservation law, as the next example demonstrates. Put simply, the total amount of species A and B is preserved by the chemistry, so any inflow of those species, no matter how small, will overwhelm it.

Example 3.2.17: Consider the RNIC model $\mathcal{F} = (\mathcal{I}_\mathcal{K}, \mathcal{H}_\mathcal{K}, \mu)$ described by

$$A \underset{\kappa_b}{\overset{\kappa_a}{\rightleftarrows}} B \qquad\qquad 0 \xrightarrow{\kappa_I} C \underset{\kappa_C}{\overset{\kappa_F}{\rightleftarrows}} 2C \qquad\qquad \mu \tag{3.6}$$

where μ is not yet specified.

Proposition 3.2.18: Let N be the coarse-grained model associated with the system $\mathcal{F}$ from (3.6). If μ is any measure on $\mathbb{Z}^2_{\geq 0}$ other than the trivial measure $\delta_{(0,0)}$, then N is transient even though all states are positive recurrent for $\mathcal{I}_\mathcal{K}$. On the other hand, if $\mu = \delta_{(0,0)}$ then N is positive recurrent.

Proof. $\mathcal{I}_\mathcal{K}$ is not irreducible, but when it is partitioned into closed irreducible communicating classing, all are finite, and hence all states are positive recurrent. As always when $\kappa_E = 0$ but $\kappa_C > 0$, the empty state is transient for $\mathcal{H}_\mathcal{K}$ but all other states are positive recurrent.

For $n \in \mathcal{N}$, let $S(n) = \sum_{a=0}^{\infty} \sum_{b=0}^{\infty} (a+b) n_{(a,b)}$ denote the sum of the number of A and B molecules, combined across all compartments in n.

First suppose that $\mu \neq \delta_{(0,0)}$. Then $S(N(t))$ cannot shrink, and grows with positive probability every time a compartment enters. So N is transient in this case.

Now suppose $\mu = \delta_{(0,0)}$. For $n \in \mathcal{N}$, let $C(n) = \|n\|_{\ell^1}$ be the number of compartments in state n, and let $V(n) = 2C(n)$. Then

$$\mathcal{L}V(n) = 2\kappa_I + 2\kappa_F C(n) - \kappa_C C(n)(C(n) - 1),$$

where $\mathcal{L}$ is the generator of N. This is less than -1 outside a finite set because it is quadratic in $C(n)$ with negative leading term, provided we restrict the state space to $\{n \in \mathcal{N} : S(n) = S(N(0))\}$. So Theorem 6.1.2 applies and N is positive recurrent, as claimed. □

△

A natural question at this point is whether, if the behaviors in the last two examples are ruled out, N can still be transient when $\mathcal{I}_\mathcal{K}$ and $\mathcal{H}_\mathcal{K}$ are both separately recurrent. Specifically, if $\mathcal{I}_\mathcal{K}$ and $\mathcal{H}_\mathcal{K}$ are both recurrent, there are no conservation laws, and the number of molecules that an incoming compartment can have is bounded, can N be transient? The answer is yes, as the next example demonstrates.

Example 3.2.19: Consider the RNIC model $\mathcal{F} = (\mathcal{I}_\mathcal{K}, \mathcal{H}_\mathcal{K}, \mu)$ described by

$$0 \underset{1}{\overset{1}{\rightleftarrows}} S \xrightarrow{1} 2S \qquad\qquad 0 \xrightarrow{1} C \xleftarrow{1} 2C \qquad\qquad \delta_1, \tag{3.7}$$

where δ_1 is the point mass at one S.

Proposition 3.2.20: Let N be the coarse-grained model associated to the network $\mathcal{F} = (\mathcal{I}_\mathcal{K}, \mathcal{H}_\mathcal{K}, \mu)$ from (3.7). Then $\mathcal{I}_\mathcal{K}$ is recurrent with no conservation laws and the number of molecules in new compartments is bounded, however every state is transient for N.

Proof. $\mathcal{I}_\mathcal{K}$ is (null) recurrent, and $\mathcal{H}_\mathcal{K}$ is positive recurrent on the irreducible state space $\{1, 2, \dots\}$, by Lemma 3.2.3.

It remains to show that every state is transient for N. As in all examples with $\kappa_E = 0$, the state with zero compartments can never be returned to and we restrict the state space of the chain to $\mathcal{N} \setminus \{0\}$. With this assumption the state space is a closed irreducible set, so it suffices to pick one state and show that it is transient. We will show e_0 (the state with one empty compartment) is transient. Denoting a state of N by n, let $C(n) = \sum_{x=0}^{\infty} n_x$ and $S(n) = \sum_{x=0}^{\infty} x \cdot n_x$ denote the total number of compartments and S molecules, respectively. Define $V : \mathcal{N} \to [0, 1]$ by

$$V(n) = \frac{S(n)}{1 + S(n)}.$$

If $\mathcal{L}$ denotes the generator of N, notice that

$$\begin{aligned}
\mathcal{L}V(n) &= (C(n) + S(n) + 1)\left(\frac{S(n)+1}{S(n)+2} - \frac{S(n)}{S(n)+1}\right) + S(n)\left(\frac{S(n)-1}{S(n)} - \frac{S(n)}{S(n)+1}\right) \\
&= \frac{C(n) + S(n) + 1}{(S(n)+2)(S(n)+1)} - \frac{1}{S(n)+1} \\
&= \frac{C(n) - 1}{(S(n)+2)(S(n)+1)} \\
&\geq 0
\end{aligned}$$

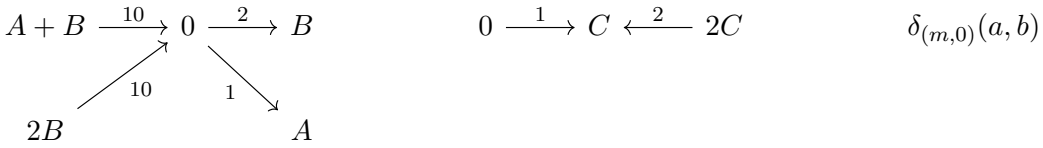
for all $n \in \mathcal{N} \setminus \{0\}$. In particular, if $B = \{e_0\}$, we can apply Theorem 6.1.3 to conclude

that when N is started from $e_0 + e_1$ (the state with two compartments, one empty and the other with one S), then the probability of reaching B is less than 1. But when N is started from e_0, it reaches $e_0 + e_1$ with positive probability (the transition from e_0 to $e_0 + e_1$ corresponds to an inflow event). Putting these together, when N is started from e_0 it fails to return with positive probability, and hence e_0 is transient. As discussed, this is enough to conclude that all states are transient for N. □

△

In the previous example $\mathcal{I}_\mathcal{K}$ was *null* recurrent. One may still be tempted to think that perhaps if it were *positive* recurrent then the whole process must be. The next example demonstrates that even this is not guaranteed.

Example 3.2.21: Consider the compartment model described by



(3.8)

where m is some non-negative integer and $\delta_{(m,0)}$ is the point mass at m molecules of A and zero of B. Let $\gamma > 0$ denote the expected number of compartments in stationarity.

Proposition 3.2.22: Let $\mathcal{F} = (\mathcal{I}_\mathcal{K}, \mathcal{H}_\mathcal{K}, \mu)$ be the compartment model from (3.8), and let N be the associated coarse-grained model. Then $\mathcal{I}_\mathcal{K}$ is positive recurrent, but N is transient when $m > \gamma$.

Proof. That $\mathcal{I}_\mathcal{K}$ is positive recurrent is witnessed by the Lyapunov function

$$V(a,b) = \begin{cases} 3a + 3 & b = 0 \\ 3a + 3b - 2 & b \geq 1 \end{cases}$$

Indeed, if $\mathcal{A}$ denotes the generator of $\mathcal{I}_{\mathcal{K}}$, then

$$\mathcal{A}V(a,b) = \begin{cases} 3(1) - 2(2) & b = 0 \\ 3(1) + 3(2) - 1(10a) & b = 1 \\ 3(1) + 3(2) - 6(20a) - 1(10) & b = 2 \\ 3(1) + 3(2) - 6(10ab) - 6(5b(b-1)) & b \geq 3 \end{cases}$$

$$= \begin{cases} -1 & b = 0 \\ 9 - 10a & b = 1 \\ -1 - 120a & b = 2 \\ 9 - 60ab - 30b(b-1) & b \geq 3 \end{cases}$$

This is at most -1 away from $(0,1)$, so by Theorem 6.1.2 $\mathcal{I}_{\mathcal{K}}$ is positive recurrent.

Now regarding transience of N, let F^{sim} be the simulation representation of $\mathcal{F}$, so that N is a fuction of F^{sim}. Let X_A and X_B denote the total number of A and B molecules, respectively, across all compartments in N (equivalently, across all compartments in F^{sim}). To show that N is transient, we will show that $X_A(t) \to \infty$ a.s., as $t \to \infty$. To do this, we will make use of the construction of F^{sim} from section 2.3.2. Let Y_I and Y_C be as in that section, so that the process M_C for the number of compartments is given by

$$M_C(t) = M_C(0) + Y_I\left(t\right) - Y_C\left(\int_0^t \frac{M_C(s)(M_C(s)-1)}{2} ds\right).$$

Similarly, for $r \in \{A + B \to 0, 0 \to B, 2B \to 0, 0 \to A\}$ let Y_r be as in section 2.3.2, and let R_r be the associated counting process for the number of times reaction r has occurred across all compartments, so that

$$R_r(t) = Y_r\left(\sum_{\substack{i \geq 0 \\ T_i \leq t}} \int_{T_i}^{T_{i+1} \wedge t} \sum_{j=1}^{M_C(T_i)} \lambda_r(X_j^i(s)) ds\right),$$

where the T_i are the jump times of the process M_C, $X_j^i(s)$ is the state of the process in compartment j at time s, and λ_r is given as in (2.2). Then

$$\begin{aligned} X_A(t) &= X_A(0) + R_{0\to A}(t) + mY_I\left(t\right) - R_{A+B\to 0}(t) \\ &= X_A(0) + Y_{0\to A}\left(\int_0^t M_C(s)ds\right) + mY_I\left(t\right) - R_{A+B\to 0}(t). \end{aligned}$$

Notice that in the last line above we were able to simplify the expression for $R_{0\to A}$ in terms of $Y_{0\to A}$ from the expression given above for R_r in general. This was done by making use of the fact that the total rate of this reaction across all compartments, $\sum_j \lambda_{0\to A}(X_j^i(s))$, is exactly the total number of compartments $M_C(s)$. We cannot hope to do the same for $R_{A+B\to 0}$ because the rate of that reaction depends on how the molecules are distributed across the compartments. However, notice that the total number of times the reaction $A + B \to 0$ fires is at most the total number of B molecules ever present in the system:

$$\begin{aligned} R_{A+B\to 0}(t) &\leq X_B(0) + R_{0\to B}(t) \\ &= X_B(0) + Y_{0\to B}\left(2\int_0^t M_C(s)ds\right). \end{aligned}$$

Therefore,

$$X_A(t) \geq X_A(0) - X_B(0) + Y_{0\to A}\left(\int_0^t M_C(s)ds\right) + mY_I\left(t\right) - Y_{0\to B}\left(2\int_0^t M_C(s)ds\right).$$

Recall that γ denotes the expected number of C in the CRN $\mathcal{H}_\mathcal{K}$ at stationarity. By the CTMC ergodic theorem (see Theorem 45 in Chapter 4 of [25]), $\frac{1}{t}\int_0^t M_C(s)ds \to \gamma$ almost surely as $t \to \infty$. This will matter in its own right; it also follows that $\int_0^t M_C(s)ds \to \infty$ a.s. as $t \to \infty$. It is a standard fact about unit Poisson processes Y that $Y(t)/t \to 1$ a.s. as $t \to \infty$. Composing this Poisson limit with the limit from the previous sentence, we get

that

$$\frac{Y_{0\to B}\left(2\int_0^t M_C(s)ds\right)}{2\int_0^t M_C(s)ds} \to 1$$

a.s. as $t \to \infty$, and similarly for $Y_{0\to A}$. Putting this all together we have

$$\begin{aligned}\lim_{t\to\infty}\frac{X_A(t)}{t} \geq \lim_{t\to\infty}\Bigg[&\frac{Y_{0\to A}\left(\int_0^t M_C(s)ds\right)}{\int_0^t M_C(s)ds}\cdot\frac{1}{t}\int_0^t M_C(s)ds + m\frac{Y_I(t)}{t} \\ &- \frac{Y_{0\to B}\left(2\int_0^t M_C(s)ds\right)}{2\int_0^t M_C(s)ds}\cdot\frac{2}{t}\int_0^t M_C(s)ds\Bigg] \\ = \gamma + m - 2\gamma.&\end{aligned}$$

almost surely. Therefore, as long as the integer m is (strictly) larger than γ, $X_A(t)/t$ is converging almost surely to a positive number. In this case $X_A(t) \to \infty$ a.s. as $t \to \infty$, and hence N is transient. □

Note that the above example shows the potential usefulness of the RNIC representation provided in section 2.3.2. △

Chapter 4

Stationary Distribution in a Special Case

In light of Theorem 3.2.5, whenever $\mathcal{H}_\mathcal{K}$ is positive recurrent and $\kappa_E > 0$, then N, the coarse-grained model associated to $\mathcal{F} = (\mathcal{I}_\mathcal{K}, \mathcal{H}_\mathcal{K}, \mu)$, is positive recurrent for at least some states. In this case, the standard theory of Markov chains tells us that there is a stationary distribution supported on those states. Ideally, it would be possible to write down a formula for this stationary distribution in terms of information about the CRNs $\mathcal{I}_\mathcal{K}$ and $\mathcal{H}_\mathcal{K}$. Under the further assumption that $\kappa_C = 0 = \kappa_F$ (so that compartments are not interacting), we are able to do so.

4.1 Theorem

First we recall some general theory.

Definition 4.1.1: Let M be a continuous-time Markov chain with discrete state space $\mathbb{S}$ and transition rate matrix $Q = (q_{ij})_{i,j\in\mathbb{S}}$. Let $\pi = (\pi(i))_{i\in\mathbb{S}} = (\pi_i)_{i\in\mathbb{S}}$ be a probability measure on $\mathbb{S}$. If $\pi Q := \left(\sum_{i\in\mathbb{S}} \pi_i q_{ij}\right)_{j\in\mathbb{S}} = 0$, we say π is an *invariant distribution*. If for every $t > 0$ we have $\mathbb{P}_\pi(X(t) = j) = \pi(j)$, then we say π is a *stationary distribution*.

The previous definition is not totally standard (in particular, Norris [17] uses *stationary*

distribution as a synonym of *invariant distribution*). However, the following theorem ensures there should be no confusion.

Theorem 4.1.2 (3.5.5 in [17])**:** Let M be a continuous-time Markov chain with state space $\mathbb{S}$. Let π be a probability measure on $\mathbb{S}$. If π is an invariant distribution for M, and M is irreducible and recurrent on the support of π, then π is a stationary distribution for M.

The idea is that having a stationary distribution π is useful (among other things it can be shown that π is the limiting distribution for the chain), whereas the condition $\pi Q = 0$ is more easily checkable. As discussed above, when $\kappa_E > 0$ and $\kappa_C = 0 = \kappa_F$ the compartment system has a stationary distribution π; the next theorem gives a formula for this π, and is proven by checking the condition $\pi Q = 0$.

Theorem 4.1.3: Consider a non-explosive model $\mathcal{F} = (\mathcal{I}_\mathcal{K}, \mathcal{H}_\mathcal{K}, \mu)$ with $\kappa_F = \kappa_C = 0$, and $\kappa_E > 0$:

$$\mathcal{I}_\mathcal{K} \qquad\qquad 0 \underset{\kappa_E}{\overset{\kappa_I}{\rightleftarrows}} C \qquad\qquad \mu$$

Let N be the coarse-grained model associated to $\mathcal{F}$. For $x \in \mathbb{Z}^d_{\geq 0}$ and $t \in [0, \infty)$, let $P_\mu(x,t)$ denote the probability that $\mathcal{I}_\mathcal{K}$ is in state x at time t when started from time zero with initial distribution μ. For $x \in \mathbb{Z}^d_{\geq 0}$ define $\alpha(x)$ via

$$\alpha(x) = \int_0^\infty P_\mu(x,t)\kappa_E e^{-\kappa_E t}\mathrm{d}t,$$

and define a distribution π on $\mathcal{N}$ via

$$\pi(n) = \left(\prod_{x \in \mathbb{Z}^d_{\geq 0}} \frac{\alpha(x)^{n_x}}{n_x!} \right) \cdot \left[e^{-\kappa_I/\kappa_E} \cdot \left(\frac{\kappa_I}{\kappa_E} \right)^{\|n\|_{\ell^1}} \right]$$

Then π is the unique stationary distribution for N.

Remark 4.1.4: To apply Theorem 4.1.3, one needs to know not just the stationary distribution for the chemistry, but the distribution for all time. This requirement may seem

daunting, and indeed for many models this distribution is not known. One class of models where it *is* know are the DR models of [15]. A second class of models are monomolecular reaction networks with arbitrary initial conditions — see [26]. Note that [15] allows for more general networks (all monomolecular networks satisfy the DR condition), but [26] allows for more general initial conditions (the DR paper requires Poisson initial conditions). $\triangle$

Proof of Theorem 4.1.3. Note that by Theorem 3.2.5, any state which is reachable from the zero state is positive recurrent, and all other states are transient. Furthermore, notice that N is irreducible if restricted to the set of states which are reachable from the zero state, since zero is reachable from any state. Thus there is a unique stationary distribution. To prove that the π given above is indeed this unique stationary distribution, it suffices to show that π is a distribution and $\pi Q = 0$, where Q is the transition rate matrix for N. That π is a distribution follows from the fact that α is a distribution, which we will check later in the proof. So fix $n \in \mathcal{N}$; we wish to show that $\sum_{n' \in \mathcal{N}} \pi(n') q(n', n) = 0$.

Note that there are only three possible types of transitions: inflow of compartment, outflow of compartment, and transition of reaction network. Expanding the sum above into three terms, one for each of these types of transitions, the desired equality can be written

$$\begin{aligned}\sum_{x \in \mathbb{Z}^d_{\geq 0}} \Bigg[& \pi(n - e_x) q(n - e_x, n) + \pi(n + e_x) q(n + e_x, n) \\ & + \sum_j \pi(n - e_x + e_{x - \nu'_j + \nu_j}) q(n - e_x + e_{x - \nu'_j + \nu_j}, n) \Bigg] \\ & = \pi(n) \sum_{x \in \mathbb{Z}^d_{\geq 0}} \left(q(n, n + e_x) + q(n, n - e_x) + \sum_j q(n, n - e_x + e_{x + \nu'_j - \nu_j}) \right)\end{aligned}$$

or

$$\sum_{x\in\mathbb{Z}^d_{\geq 0}}\Bigg[\pi(n-e_x)\kappa_I\mu(x)+\pi(n+e_x)\kappa_E(n_x+1) + \sum_j \pi(n-e_x+e_{x-\nu'_j+\nu_j})(n_{x-\nu'_j+\nu_j}+1)\kappa_j\binom{x-\nu'_j+\nu_j}{\nu_j}\Bigg] \tag{4.1}$$
$$= \pi(n)\sum_{x\in\mathbb{Z}^d_{\geq 0}}\left(\kappa_I\mu(x)+\kappa_E n_x+\sum_j n_x\kappa_j\binom{x}{\nu_j}\right)$$

To prove this equality, we will consider two cases. Suppose first that n is such that $n_y > 0$ for some $y \in \mathbb{Z}^d_{\geq 0}$ with $\alpha(y) = 0$, and fix such a y. Then $\alpha(y)$ participates in the product defining $\pi(n)$, and hence $\pi(n) = 0$. Thus the right-hand side of (4.1) is zero; we claim that the left-hand side is also zero. Specifically, we will argue for each x and each j, each of the three terms in the sum is zero. So fix x and j:

- $\pi(n-e_x)\kappa_I\mu(x)$: Notice that if $x \neq y$ then $\pi(n-e_x) = 0$ for the same reason that $\pi(n) = 0$. If $x = y$ then $\mu(x) = 0$, since if $\mu(y) > 0$ it would be the case that $P_\mu(y,t) > 0$ for all small enough t, and hence the integral defining $\alpha(y)$ would be positive.

- $\pi(n+e_x)\kappa_E(n_x+1)$: Regardless of x, $\pi(n+e_x) = 0$ for the same reason that $\pi(n) = 0$.

- $\pi(n-e_x+e_{x-\nu'_j+\nu_j})(n_{x-\nu'_j+\nu_j}+1)\kappa_j\binom{x-\nu'_j+\nu_j}{\nu_j}$: As before, if $x \neq y$ then $\pi(n-e_x+e_{x-\nu'_j+\nu_j}) = 0$. Suppose towards a contradiction that $\pi(n-e_y+e_{y-\nu'_j+\nu_j}) \neq 0$ and that $\kappa_j\binom{y-\nu'_j+\nu_j}{\nu_j} \neq 0$. Then $\pi(n-e_y+e_{y-\nu'_j+\nu_j}) \neq 0$ implies that $\alpha(y-\nu'_j+\nu_j) \neq 0$, and hence $P_\mu(y-\nu'_j+\nu_j,t) \neq 0$ for some t. But this means that the state $y-\nu'_j+\nu_j$ is reachable for $\mathcal{I}$ when started with initial distribution μ. But $\kappa_j\binom{y-\nu'_j+\nu_j}{\nu_j} \neq 0$ implies that y is reachable from $y-\nu'_j+\nu_j$ for $\mathcal{I}$ via the j-th reaction, so we conclude that y is reachable from μ. But this implies that $P_\mu(y,t) \neq 0$ for $t > 0$, which in turn means that $\alpha(y) > 0$. This contradicts our choice of y, so it must be that our assumption was wrong: either $\pi(n-e_y+e_{y-\nu'_j+\nu_j}) = 0$ or $\kappa_j\binom{y-\nu'_j+\nu_j}{\nu_j} = 0$. But either of those imply the desired equality $\pi(n-e_y+e_{y-\nu'_j+\nu_j})(n_{y-\nu'_j+\nu_j}+1)\kappa_j\binom{y-\nu'_j+\nu_j}{\nu_j} = 0$.

This proves that (4.1) reduces to $0 = 0$ in this case. The reminder of the proof will be devoted to the second case; namely, the case where n is such that $n_y = 0$ for all $y \in \mathbb{Z}^d_{\geq 0}$ with $\alpha(y) = 0$.

Let $\mathbb{X} = \{x \in \mathbb{Z}^d_{\geq 0} : \alpha(x) \neq 0\}$. We claim that for every $x \notin \mathbb{X}$ and every j, every summand in (4.1) is zero. So fix $x \notin \mathbb{X}$ and j:

- $\pi(n - e_x)\kappa_I \mu(x)$: Since $\alpha(x) = 0$, by choice of n we have $n_x = 0$. But this means that $n - e_x$ is negative at x and hence $n - e_x \notin \mathcal{N}$, so $\pi(n - e_x) = 0$.

- $\pi(n + e_x)\kappa_E(n_x + 1)$: Notice that $\alpha(x) = 0$ participates in the product defining $\pi(n + e_x)$, and hence $\pi(n + e_x) = 0$.

- $\pi(n - e_x + e_{x-\nu'_j+\nu_j})(n_{x-\nu'_j+\nu_j} + 1)\kappa_j \binom{x-\nu'_j+\nu_j}{\nu_j}$: As before, $n - e_x + e_{x-\nu'_j+\nu_j} \notin \mathcal{N}$ and hence $\pi(n - e_x + e_{x-\nu'_j+\nu_j}) = 0$.

- $\kappa_I \mu(x)$: Since $\alpha(x) = 0$, it must be the case that $\mu(x) = 0$, as otherwise $P_\mu(x, t)$ would be positive for sufficiently small t.

- $\kappa_E n_x$: Since $\alpha(x) = 0$, by choice of n we have $n_x = 0$.

- $n_x \kappa_j \binom{x}{\nu_j}$: Once again, $n_x = 0$.

Thus we have shown that terms with $x \notin \mathbb{X}$ do not contribute to (4.1). So to complete the proof, we have only to show that

$$\begin{aligned} \sum_{x \in \mathbb{X}} \Bigg[& \pi(n - e_x)\kappa_I \mu(x) + \pi(n + e_x)\kappa_E(n_x + 1) \\ & + \sum_j \pi(n - e_x + e_{x-\nu'_j+\nu_j})(n_{x-\nu'_j+\nu_j} + 1)\kappa_j \binom{x - \nu'_j + \nu_j}{\nu_j} \Bigg] \\ & = \pi(n) \sum_{x \in \mathbb{X}} \left(\kappa_I \mu(x) + \kappa_E n_x + \sum_j n_x \kappa_j \binom{x}{\nu_j} \right). \end{aligned} \tag{4.2}$$

Let $x \in \mathbb{X}$ be arbitrary. Integration by parts gives

$$\begin{aligned}\int_0^\infty \left(\frac{\mathrm{d}}{\mathrm{d}t} P_\mu(x,t)\right) \kappa_E e^{-\kappa_E t} \mathrm{d}t &= \kappa_E e^{-\kappa_E t} P_\mu(x,t)\Big|_{t=0}^{t=\infty} - \int_0^\infty P_\mu(x,t)(-\kappa_E^2 e^{-\kappa_E t})\mathrm{d}t \\ &= -\kappa_E \mu(x) + \kappa_E \alpha(x).\end{aligned}$$

Because P_μ is the distribution for $\mathcal{I}_\mathcal{K}$, the Kolmogorov forward equations for $\mathcal{I}_\mathcal{K}$ tell us that

$$\frac{\mathrm{d}}{\mathrm{d}t} P_\mu(x,t) = \sum_{\nu_j \to \nu_j'} \kappa_j \binom{x - \nu_j' + \nu_j}{\nu_j} P_\mu(x - \nu_j' + \nu_j, t) - \sum_{\nu_j \to \nu_j'} \kappa_j \binom{x}{\nu_j} P_\mu(x,t)$$

for each t. Plugging this in above and rearranging yields

$$\begin{aligned}\sum_{\nu_j \to \nu_j'} \kappa_j \binom{x - \nu_j' + \nu_j}{\nu_j} \alpha(x - \nu_j' + \nu_j) - \sum_{\nu_j \to \nu_j'} \kappa_j \binom{x}{\nu_j} \alpha(x) &= -\kappa_E \mu(x) + \kappa_E \alpha(x) \\ \kappa_E \frac{\mu(x)}{\alpha(x)} + \sum_{\nu_j \to \nu_j'} \kappa_j \binom{x - \nu_j' + \nu_j}{\nu_j} \frac{\alpha(x - \nu_j' + \nu_j)}{\alpha(x)} &= \kappa_E + \sum_{\nu_j \to \nu_j'} \kappa_j \binom{x}{\nu_j}.\end{aligned}$$

Note that we did not divide by zero in the second line because $\alpha(x) \neq 0$ by definition of $\mathbb{X}$. Since $x \in \mathbb{X}$ was arbitrary, we can multiply through by n_x and sum over x, which yields

$$\begin{aligned}\sum_{x \in \mathbb{X}} \left(n_x \kappa_E \frac{\mu(x)}{\alpha(x)} + n_x \sum_{\nu_j \to \nu_j'} \kappa_j \binom{x - \nu_j' + \nu_j}{\nu_j} \frac{\alpha(x - \nu_j' + \nu_j)}{\alpha(x)} \right) \\ = \sum_{x \in \mathbb{X}} \left(n_x \kappa_E + n_x \sum_{\nu_j \to \nu_j'} \kappa_j \binom{x}{\nu_j} \right). \quad (4.3)\end{aligned}$$

Now we claim that μ and α are both probability measures supported on $\mathbb{X}$. We know that μ is a probability measure by assumption; it is supported on $\mathbb{X}$ because if $\mu(x) > 0$ then $P_\mu(x,t) > 0$ for small enough t and hence $\alpha(x) > 0$. We know that α is supported on $\mathbb{X}$ by definition of $\mathbb{X}$; to see that it is a probability measure, use the fact that the integrand in the definition of α is non-negative to interchange a sum over x with the integral and then apply the fact that $P_\mu(x,t)$ is a probability measure for each t. Therefore μ and α

are both probability measures supported on $\mathbb{X}$, as claimed; it follows that $\sum_{x\in\mathbb{X}} \kappa_I \mu(x) = \kappa_I = \sum_{x\in\mathbb{X}} \kappa_I \alpha(x)$. So adding κ_I to both sides of (4.3) gives

$$\sum_{x\in\mathbb{X}} \left(n_x \kappa_E \frac{\mu(x)}{\alpha(x)} + \kappa_I \alpha(x) + n_x \sum_{\nu_j \to \nu_j'} \kappa_j \binom{x - \nu_j' + \nu_j}{\nu_j} \frac{\alpha(x - \nu_j' + \nu_j)}{\alpha(x)} \right)$$
$$= \sum_{x\in\mathbb{X}} \left(\kappa_I \mu(x) + n_x \kappa_E + n_x \sum_{\nu_j \to \nu_j'} \kappa_j \binom{x}{\nu_j} \right). \tag{4.4}$$

Now notice that, directly from the definition of π, we have

$$\begin{aligned}
\frac{\pi(n - e_x)}{\pi(n)} &= \frac{n_x}{\alpha(x)} \frac{\kappa_E}{\kappa_I} \\
\frac{\pi(n + e_x)}{\pi(n)} &= \frac{\alpha(x)}{n_x + 1} \frac{\kappa_I}{\kappa_E} \\
\frac{\pi(n - e_x + e_{x-\nu_j'+\nu_j})}{\pi(n)} &= \frac{\alpha(x - \nu_j' + \nu_j)}{\alpha(x)} \frac{n_x}{n_{x-\nu_j'+\nu_j} + 1},
\end{aligned}$$

where the last equality holds for each reaction $\nu_j \to \nu_j'$. Applying these three in order on the left-hand side of (4.4), we get

$$\begin{aligned}
&\sum_{x\in\mathbb{X}} \Bigg(\kappa_I \mu(x) \frac{\pi(n - e_x)}{\pi(n)} + \kappa_E (n_x + 1) \frac{\pi(n + e_x)}{\pi(n)} \\
&\quad + \sum_{\nu_j \to \nu_j'} \kappa_j \binom{x - \nu_j' + \nu_j}{\nu_j} (n_{x-\nu_j'+\nu_j} + 1) \frac{\pi(n - e_x + e_{x-\nu_j'+\nu_j})}{\pi(n)} \Bigg) \\
&\quad = \sum_{x\in\mathbb{X}} \left(\kappa_I \mu(x) + n_x \kappa_E + n_x \sum_{\nu_j \to \nu_j'} \kappa_j \binom{x}{\nu_j} \right) \\
&\sum_{x\in\mathbb{X}} \Bigg(\kappa_I \mu(x) \pi(n - e_x) + \kappa_E (n_x + 1) \pi(n + e_x) \\
&\quad + \sum_{\nu_j \to \nu_j'} \kappa_j \binom{x - \nu_j' + \nu_j}{\nu_j} (n_{x-\nu_j'+\nu_j} + 1) \pi(n - e_x + e_{x-\nu_j'+\nu_j}) \Bigg) \\
&\quad = \pi(n) \sum_{x\in\mathbb{X}} \left(\kappa_I \mu(x) + n_x \kappa_E + n_x \sum_{\nu_j \to \nu_j'} \kappa_j \binom{x}{\nu_j} \right),
\end{aligned}$$

which is exactly the desired equality, (4.2). □

4.2 Specific Examples

Let us now consider some examples of applying this result.

Example 4.2.1: Let $\lambda \geq 0$, and consider the compartment system

$$0 \underset{\kappa_d}{\overset{\kappa_b}{\rightleftarrows}} S \qquad\qquad 0 \underset{\kappa_E}{\overset{\kappa_I}{\rightleftarrows}} C \qquad\qquad \text{Poisson}(\lambda)$$

Then the stationary distribution of the system is given by

$$\pi(n) = \left(\prod_{x=0}^{\infty} \frac{\alpha(x)^{n_x}}{n_x!} \right) \cdot \left[e^{-\kappa_I/\kappa_E} \cdot \left(\frac{\kappa_I}{\kappa_E} \right)^{\|n\|_{\ell^1}} \right],$$

where

$$\alpha(x) = \int_0^{\infty} \exp\left\{ -(\lambda - \kappa_b/\kappa_d)e^{-\kappa_d t} - \kappa_b/\kappa_d \right\} \frac{((\lambda - \kappa_b/\kappa_d)e^{-\kappa_d t} + \kappa_b/\kappa_d)^x}{x!} \kappa_E e^{-\kappa_E t} \mathrm{d}t.$$

Proof. Check that the distribution

$$P_\lambda(x,t) := \exp\left\{ -(\lambda - \kappa_b/\kappa_d)e^{-\kappa_d t} - \kappa_b/\kappa_d \right\} \frac{((\lambda - \kappa_b/\kappa_d)e^{-\kappa_d t} + \kappa_b/\kappa_d)^x}{x!}$$

is Poisson(λ) at time $t = 0$ and satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t} P_\lambda(x,t) = \kappa_b P_\lambda(x-1,t) + \kappa_d(x+1)P_\lambda(x+1,t) - \kappa_b P_\lambda(x,t) - \kappa_d x P_\lambda(x,t)$$

for each x and t, and apply Theorem 4.1.3. □

△

In the previous example, notice that the expected value of α is

$$\begin{aligned}
&\sum_{x=0}^{\infty} x\alpha(x) \\
&= \int_0^{\infty} \sum_{x=0}^{\infty} \exp\left\{-(\lambda - \kappa_b/\kappa_d)e^{-\kappa_d t} - \kappa_b/\kappa_d\right\} \frac{((\lambda - \kappa_b/\kappa_d)e^{-\kappa_d t} + \kappa_b/\kappa_d)^{x+1}}{x!} \kappa_E e^{-\kappa_E t} \mathrm{d}t \\
&= \int_0^{\infty} (\lambda - \kappa_b/\kappa_d)\kappa_E e^{-(\kappa_d+\kappa_E)t} + \frac{\kappa_b\kappa_E}{\kappa_d} e^{-\kappa_E t} \mathrm{d}t \\
&= \frac{(\lambda - \kappa_b/\kappa_d)\kappa_E}{\kappa_d + \kappa_E} + \frac{\kappa_b}{\kappa_d} \\
&= \frac{\lambda\kappa_E + \kappa_b}{\kappa_d + \kappa_E}.
\end{aligned}$$

This matches [1], where the same example is consider in section 2.A (see specifically their equation [20] and the following discussion). Note that in [1], though the expected value of α is calculated in general, an explicit formula for $\alpha(x)$ (which, in their notation, is written $P_\infty(x)$) is given in only two cases. The first is the case where $\lambda = \kappa_b/\kappa_d$, where (in section S7.4 of their SI Appendix) they remark that P_∞ is Poission with mean λ. This matches the formula we give above in Example 4.2.1. The second case they cover is the one where $\kappa_d = 0$. In that case they obtain

$$P_\infty(x) = (1-\xi)\xi^x e^{\lambda(1/\xi - 1)} \frac{\Gamma(1+x, \lambda/\xi)}{x!}, \tag{4.5}$$

where $\xi = \kappa_b/(\kappa_b + \kappa_E)$ and Γ is the upper incomplete Gamma function. In Proposition 4.2.3, we will check that (4.5) agrees with our next example, Example 4.2.2, in the case where μ is taken to be Poission with parameter λ.

The following example is interesting for a few reasons. First, the chemistry is not converging to any sort of stationary distribution, and yet the whole compartment model is. Second, notice that when μ is not a Poisson distribution, $P_\mu(x,t)$ is not a Poisson distribution in x for all t unlike the previous example or more generally the DR models from Remark 4.1.4. Third, as discussed above, it generalizes an example from [1].

Example 4.2.2: Let μ be a probability distribution on $\mathbb{Z}_{\geq 0}$, and consider the compart-

ment system

$$0 \xrightarrow{\kappa_b} S \qquad\qquad 0 \underset{\kappa_E}{\overset{\kappa_I}{\rightleftarrows}} C \qquad\qquad \mu,$$

Then the stationary distribution of the system is given by

$$\pi(n) = \left(\prod_{x=0}^{\infty} \frac{\alpha(x)^{n_x}}{n_x!} \right) \cdot \left[e^{-\kappa_I/\kappa_E} \cdot \left(\frac{\kappa_I}{\kappa_E} \right)^{\|n\|_{\ell^1}} \right],$$

where

$$\alpha(x) = \int_0^{\infty} e^{-\kappa_b t} \left(\sum_{m=0}^{x} \frac{\kappa_b^m t^m}{m!} \mu(x-m) \right) \kappa_E e^{-\kappa_E t} \mathrm{d}t. \tag{4.6}$$

Proof. Check that the distribution

$$P_\mu(x,t) := e^{-\kappa_b t} \left(\sum_{m=0}^{x} \frac{\kappa_b^m t^m}{m!} \mu(x-m) \right)$$

satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t} P_\mu(x,t) = \kappa_b P_\mu(x-1,t) - \kappa_b P_\mu(x,t),$$

with initial condition $P_\mu(x,0) = \mu(x)$, and apply Theorem 4.1.3. □

△

As discussed, in the special case of the above where μ is Poisson with parameter λ, [1] obtained the expression in (4.5) for α. So our expression in (4.6) should, by all rights, agree with theirs in that case, and indeed it does:

Proposition 4.2.3: Let $\mu(x) = \lambda^x e^{-\lambda}/x!$, let α be as in (4.6), and let P_∞ be as in (4.5). Then $\alpha(x) = P_\infty(x)$.

Proof. The binomial theorem gives

$$\begin{aligned}
\alpha(x) &= \int_0^\infty e^{-\kappa_b t}\left(\sum_{m=0}^{x}\frac{\kappa_b^m t^m}{m!}\frac{\lambda^{x-m}e^{-\lambda}}{(x-m)!}\right)\kappa_E e^{-\kappa_E t}\mathrm{d}t \\
&= \int_0^\infty e^{-(\kappa_b t+\lambda)}\frac{1}{x!}\left(\sum_{m=0}^{x}\binom{x}{m}(\kappa_b t)^m\lambda^{x-m}\right)e^{-\kappa_E t}\kappa_E\mathrm{d}t \\
&= \int_0^\infty e^{-(\kappa_b t+\lambda)}\frac{1}{x!}(\kappa_b t+\lambda)^x e^{-\kappa_E t}\kappa_E\mathrm{d}t
\end{aligned}$$

Let $\xi u = \kappa_b t + \lambda$. Then multiplying by κ_E/κ_b and rearrainging yields

$$\begin{aligned}
\kappa_E t &= \frac{\kappa_E}{\kappa_b+\kappa_E}u - \lambda\frac{\kappa_E}{\kappa_b} \\
&= (1-\xi)u - \lambda(1/\xi - 1).
\end{aligned}$$

So $\kappa_E \mathrm{d}t = (1-\xi)\mathrm{d}u$ and

$$\begin{aligned}
\alpha(x) &= \int_{\lambda/\xi}^\infty e^{-\xi u}\frac{1}{x!}(\xi u)^x e^{-(1-\xi)u+\lambda(1/\xi-1)}(1-\xi)\mathrm{d}u \\
&= (1-\xi)\xi^x e^{\lambda(1/\xi-1)}\int_{\lambda/\xi}^\infty e^{-\xi u}\frac{1}{x!}u^x e^{-u+\xi u}\mathrm{d}u \\
&= P_\infty(x)
\end{aligned}$$

as claimed. □

Chapter 5

A Generalization: Content-dependent Fragmentation

This chapter is the first whose material is not present in [2]. The models studied in this chapter do not fall under the framework from that paper because the rate of certain compartment transitions will depend on the content of the compartments. Note that these models will fall under the more general framework from [1].

Let $\mathcal{I}_{\mathcal{K}}$ be some reaction network where S appears as a species. The models of this chapter will be represented by diagrams of the form

$$\mathcal{I}_{\mathcal{K}} \qquad 0 \underset{\kappa_E}{\overset{\kappa_I}{\rightleftarrows}} C \underset{\kappa_C}{\overset{\kappa_F S_C}{\rightleftarrows}} 2C \qquad \mu, \tag{5.1}$$

where the intent is that everything is the same as in previous chapters except that a given compartment with s molecules of S fragments at rate $\kappa_F \cdot s$ instead of rate κ_F.

Remark 5.0.1: The usual convention in the study of reaction networks is to write rate constants over reaction arrows, but $\kappa_F S_C$ is not a rate constant, which may seem strange. One way to think about the rate constants we have been writing is that any given compartment exits the system with rate κ_E, any given pair of compartments coagulates with rate κ_C, etc.. Viewed in this light, the new notation makes more sense, since if S_C is the

number of S in compartment C, then $\kappa_F S_C$ is exactly the rate at which compartment C is fragmenting, consistent with the old notation. $\triangle$

To be precise, let d (as always) be the number of species which appear in $\mathcal{I}_\mathcal{K}$, and for $x \in \mathbb{Z}^d_{\geq 0}$ let $S(x)$ denote the projection onto the S coordinate: when in state x there are $S(x)$ molecules of S. For reactions $\nu \to \nu'$ of $\mathcal{I}_\mathcal{K}$, let $\lambda_{\nu\to\nu'}(x)$ denote the rate of reaction $\nu \to \nu'$ when in state x, so that the generator $\mathcal{A}$ of $\mathcal{I}_\mathcal{K}$ is given by

$$\mathcal{A}f(x) = \sum_{\nu\to\nu'} \lambda_{\nu\to\nu'}(x)(f(x+\nu'-\nu) - f(x)).$$

Let $\mathcal{N}$ be as in previous chapters, and for $n \in \mathcal{N}$ let $C(n) = \sum_{x\in\mathbb{Z}^d_{\geq 0}} n_x$ be the total number of compartments and let $S(n) = \sum_{x\in\mathbb{Z}^d_{\geq 0}} S(x)n_x$ be the total number of S molecules across all compartments. For each $x, y \in \mathbb{Z}^d_{\geq 0}$, let $\psi(x,y)$ denote the probability that when a compartment in state x fragments, the daughter compartment produced is in state y (and the original compartment is in state $x-y$). For example, earlier in this thesis we worked only with $\psi(x,y) = 2^{-(x_1+\cdots+x_d)}\binom{x}{y}$, and in Remark 5.2.6 we will take $y \mapsto \psi(x,y)$ to be uniform over possible pairs of resulting compartments. In general, we require only that $\psi(x,y) = 0$ if $y_i > x_i$ for some i, and that $y \mapsto \psi(x,y)$ is a probability measure for each x. Then in this chapter, we are studying the Markov Chain N with generator $\mathcal{L}$, where for functions $V : \mathcal{N} \to \mathbb{R}$,

$$\begin{aligned}
\mathcal{L}V(n) = \sum_{x\in\mathbb{Z}^d_{\geq 0}} \Bigg[& \left(\sum_{\nu\to\nu'} n_x \lambda_{\nu\to\nu'}(x) \big(V(n - e_x + e_{x+\nu'-\nu}) - V(n)\big) \right) \\
& + \kappa_I \mu(x) \big(V(n+e_x) - V(n)\big) + \kappa_E n_x \big(V(n-e_x) - V(n)\big) \\
& + \kappa_F S(x) n_x \left(\sum_{y\in\mathbb{Z}^d_{\geq 0}} \psi(x,y) \big(V(n - e_x + e_y + e_{x-y}\big) - V(n)\big) \right) \\
& + \kappa_C \binom{n_x}{2} \big(V(n - 2e_x + e_{2x}) - V(n)\big) \\
& + \sum_{\substack{y\in\mathbb{Z}^d_{\geq 0} \\ y\neq x}} \kappa_C \frac{n_x n_y}{2} \big(V(n - e_x - e_y + e_{x+y}) - V(n)\big) \Bigg].
\end{aligned}$$

As in previous chapters, we will refer to N as the *coarse-grained model associated to* (5.1). We will frequently be illustrating our results or techniques by considering the simpler model

$$0 \underset{\kappa_d}{\overset{\kappa_b}{\rightleftarrows}} S \qquad\qquad 0 \underset{\kappa_E}{\overset{\kappa_I}{\rightleftarrows}} C \underset{\kappa_C}{\overset{\kappa_F S_C}{\rightleftarrows}} 2C \qquad\qquad \mu, \tag{5.2}$$

which has generator

$$\begin{aligned}\mathcal{L}V(n) = \sum_{x=0}^{\infty} \Bigg[& \kappa_b n_x \big(V(n - e_x + e_{x+1}) - V(n)\big) + \kappa_d n_x x \big(V(n - e_x + e_{x-1}) - V(n)\big) \\ & + \kappa_I \mu(x) \big(V(n + e_x) - V(n)\big) + \kappa_E n_x \big(V(n - e_x) - V(n)\big) \\ & + \kappa_F x n_x \left(\sum_{y=0}^{\infty} \psi(x,y) \big(V(n - e_x + e_y + e_{x-y}) - V(n)\big) \right) \\ & + \kappa_C \binom{n_x}{2} \big(V(n - 2e_x + e_{2x}) - V(n)\big) \\ & + \sum_{\substack{y=0 \\ y \neq x}}^{\infty} \kappa_C \frac{n_x n_y}{2} \big(V(n - e_x - e_y + e_{x+y}) - V(n)\big) \Bigg].\end{aligned}$$

Note that we may take any rate constants to be zero, including possibly κ_b or κ_d. The case where $\kappa_b = \kappa_d = 0$ and the number of S in each compartment is only changed by the compartment events was studied in [1].

One assumption is frequent enough in what follows that it deserves attention here:

Condition 5.0.2: Let λ denote the expectation under μ of the total molecular count of new compartments: $\lambda = \sum_{x \in \mathbb{Z}^d_{\geq 0}} \mu(x) \sum_{j=1}^{d} x_j$. Assume μ is such that $\lambda < \infty$.

Unfortunately, translating results about $\mathcal{I}_{\mathcal{K}}$ into results about N will not be as straightforward as in the previous chapters. One case where we *can* get results about N, however, is the one where $\mathcal{I}_{\mathcal{K}}$ has a Lyapunov function, and that Lyapunov function happens to be linear. We are far from the first to consider specifically linear Lyapunov functions. For example, in [27], a linear Lyapunov function condition (which they refer to as condition DD1), stronger than any of the Lyapunov conditions appearing in this chapter, is studied. They give conditions for their condition to be satisfied for unimolecular and bimolecular

reaction networks (see sections S3 and S4, respectively, of their supplementary material), and they verify the condition for a number of biologically motivated examples (see sections S5 through S12 of the same).

5.1 Non-Explosivity

In Theorem 3.1.3, we showed that when the fragmentation rate does not depend on compartment contents, compartmentalizing a CRN does not affect whether or not it explodes: N is explosive iff $\mathcal{I}_{\mathcal{K}}$ is. However, things are more delicate in the present setting, as the next example shows.

Example 5.1.1: Consider the following model

$$E + 2S \xrightarrow{2} E + 3S \qquad\qquad C \xrightarrow{S_C} 2C \tag{5.3}$$

Let $\psi_0, \psi_1 : \mathbb{Z}^2_{\geq 0} \times \mathbb{Z}^2_{\geq 0} \to \{0, 1\}$ be the functions

$$\psi_0((e,s),(e',s')) = \begin{cases} 1 & e' = e \text{ and } s' = s \\ 0 & \text{else} \end{cases}$$

$$\psi_1((e,s),(e',s')) = \begin{cases} 1 & e' = e \text{ and } s' = 0 \\ 0 & \text{else} \end{cases}.$$

The idea is that when the compartment fragmentation distribution is given by ψ_0, a compartment in state (s, e) always splits into one compartment in state (s, e) and one empty compartment. And when it is given by ψ_1, a compartment in state (s, e) always splits into one compartment in state $(s, 0)$ and one in state $(0, e)$. Let N_i be the coarse-grained mode associated to (5.3) with fragmentation distribution ψ_i, for $i = 0, 1$.

Proposition 5.1.2: The network $E+2S \xrightarrow{2} E+3S$ is explosive. When the distribution of fragmented compartments is given by ψ_0, then the corresponding coarse-grained model N_0 associated to (5.3) is also explosive. When the distribution of fragmented compartments

is given by ψ_1, however, then the corresponding coarse-grained model N_1 associated to (5.3) is not explosive.

Proof. Let X denote the Markov chain associated to $E + 2S \xrightarrow{2} E + 3S$. Notice that since the number of E cannot change, the S-component of X is just an instance of the CRN $2S \xrightarrow{2E_0} 3S$, where E_0 is the initial amount of E present in the system. But this is well-known to be explosive whenever the rate $2E_0$ is positive — see for instance [24].

Now we argue that N_0 is explosive. Specifically, we will argue that it explodes with positive probability when started in the state with exactly one compartment, where this one compartment contains one E and two S. If $\mathcal{N}$, as always, denotes the state space for N_0, define $f : \mathcal{N} \to \mathbb{Z}^2_{\geq 0}$ so that $f(n)$ is the state $x = (s, e)$ with the largest value of s, breaking ties in favor of larger values of e, such that $n_x \neq 0$. Consider the stochastic process $f(N_0)$. We claim that $f(N_0)$ is not only a Markov chain, but actually an instance of the Markov chain X from the beginning of this proof (with starting state $(2, 1)$, from which we know X explodes).

Indeed, notice that fragmentations only produce new empty compartments, that totally empty compartments have no way of producing either more E or more S, and that existing S and E are never removed from the system, so the state of N_0 is always some number of empty compartments plus one compartment with at least two S and one E. It follows that $f(N_0)$ is just the state of this one non-empty compartment, and that fragmentation cannot affect $f(N_0)$. So $f(N_0)$ starts out in state $(2, 1)$, can only have transitions of the form $(s, 1) \to (s + 1, 1)$, and these transitions have independent exponential holding times with rate $s(s - 1)$. So $f(N_0)$ is indeed an instance of X, and thus $f(N_0)$, and hence N_0, is explosive.

Now we argue that N_1 is not explosive. Notice that if N_1 is in a state where no compartment has both S and E, then it cannot explode. Indeed, (i) in such a state the only possible transition is $C \to 2C$, (ii) once in such a state it will always be in such a state and the number of S will not further change, so (iii) the network gains C at constant rate equal to the number of S. Since the rate of the only reaction is constant, it cannot

be explosive. In light of this, we aim to show that, no matter what state N_1 is started in, it will a.s. eventually reach such a state where no compartment has both S and E before having the chance to explode.

Different compartment are independent, so it suffices to consider the case of a single compartment in state (s, e). The idea is that such a compartment fragments with rate s, which because of our choice of ψ_1 splits the S and E into different compartments and kills off any explosion. Since the compartment gains one S with rate $es(s-1)$, the probability of fragmenting before leaving the state is $\frac{s}{s+es(s-1)} = \frac{1}{1+e(s-1)}$. The number e is fixed and this quantity is not summable in s, so the second Borel–Cantelli lemma tells us that the compartment fragments before exploding with probability one. (Strictly speaking the Borel–Cantelli lemma applies to independent events, so to make this idea precise one would need to, for instance, construct the process N_1 "by hand" from independent exponential random variables. But this is straightforward, if messy, so we omit it.) □

△

The example above shows that N, the coarse-grained model associated to (5.1), can fail to be explosive even when the associated chemistry $\mathcal{I}_\mathcal{K}$ is explosive, in contrast to the case where the fragementation rate does not depend on the compartment contents. The converse, however, still holds provided we have an additional technical assumption on $\mathcal{I}_\mathcal{K}$:

Theorem 5.1.3: Let N be the coarse-grained model associated to (5.1), and let $\mathcal{A}$ denote the generator of the associated chemistry $\mathcal{I}_\mathcal{K}$. Suppose there exists some linear function $f : \mathbb{Z}^d_{\geq 0} \to \mathbb{R}$ which satisfies the hypotheses of Theorem 6.1.1; that is, suppose f is of the form $f(x) = w \cdot x$ for some $w \in \mathbb{R}^d_{>0}$, and for some constants c and d, we have $\mathcal{A}f(x) \leq cf(x) + d$ for every x. Then N is not explosive.

Proof. First, we will assume that Condition 5.0.2 is satisfied, i.e., that $\lambda < \infty$. We will use the result for this case in the $\lambda = \infty$ case. Let $V(n) = C(n) + \sum_{x \in \mathbb{Z}^d_{\geq 0}} n_x(w \cdot x)$. The assumption that every coordinate of w is strictly positive means that $V \to \infty$ in the sense of Theorem 6.1.1. Notice that neglecting the κ_E and κ_C terms and computing the other

terms exactly gets us the upper bound

$$\begin{aligned}
\mathcal{L}V(n) &\leq \sum_{x\in\mathbb{Z}^d_{\geq 0}} \kappa_F S(x)n_x + \kappa_I\mu(x)(1+w\cdot x) + \sum_{\nu\to\nu'} n_x\lambda_{\nu\to\nu'}(x)\big(w\cdot(x+\nu'-\nu)-w\cdot x\big)\\
&= \sum_{x\in\mathbb{Z}^d_{\geq 0}} \kappa_F S(x)n_x + \kappa_I\mu(x)(1+w\cdot x) + n_x\mathcal{A}f(x)\\
&\leq \sum_{x\in\mathbb{Z}^d_{\geq 0}} \kappa_F S(x)n_x + \kappa_I\mu(x) + \kappa_I\mu(x)(\max_j w_j)\left(\sum_j x_j\right) + n_x c f(x) + dn_x\\
&\leq (\kappa_F w_S^{-1} + c + d)V(n) + \big(\kappa_I + \kappa_I(\max_j w_j)\lambda\big).
\end{aligned}$$

It follows from Theorem 6.1.1 that N is not explosive when $\lambda < \infty$.

Now we deal with the case where $\lambda = \infty$. Let $\tau_0 = 0$ and for $j = 1, 2, 3, \cdots$, let τ_j denote the time of the j-th inflow for N. For $j = 0, 1, 2, \cdots$, let $N^{(j)}(t) = N(t)\mathbb{I}_{t\leq\tau_j}$. (Strictly speaking, $N^{(j)}$ is not a Markov chain, but we can expand the state space from $\mathcal{N}$ to $\mathbb{Z}_{\geq 0}\times\mathcal{N}$, where the first coordinate indicates the number of inflows that have happened. We will ignore this detail going forward.)

We claim that $N^{(j)}$ is not explosive, for any j.

We proceed by induction on j. $N^{(0)}$ isn't explosive, since it's just a constant. Suppose that $N^{(j-1)}$ is not explosive. $N^{(j)} = N^{(j-1)}$ up to time τ_{j-1}, so $N^{(j)}$ cannot explode before time τ_{j-1}, and $N^{(j)}$ is constant after time τ_j, so it remains only to consider what happens for $\tau_{j-1} \leq t \leq \tau_j$. But for $\tau_{j-1} \leq t \leq \tau_j$, by the strong Markov property $N^{(j)}(t)$ has the same distribution as $N(t)$ with $0 \leq t \leq \tau_1$ started from $N^{(j)}(\tau_{j-1})$. But we can construct another Markov chain $\tilde{N}$ coupled to N by deleting all inflow transitions, we have $N(t) = \tilde{N}(t)$ for $0 \leq t \leq \tau_1$, and $\tilde{N}(t)$ is not explosive by the argument given above in the $\lambda < \infty$ case.

Now with the claim proven, let $t, \varepsilon > 0$ be arbitrary. We will show that the probability that N explodes by time t is at most ε. Pick j large enough that $\tau_j > t$ with probability at least $1-\varepsilon$. Then since $N_s^{(j)} = N_s$ when $s < \tau_j$ and since $N^{(j)}$ is not explosive, it follows that $\{N \text{ explodes by time } t\} \subseteq \{\tau_j < t\}$, and hence the probability N explodes by time t

is at most ε. Since $\varepsilon > 0$ was arbitrary, N explodes by time t with probability zero. Since t was arbitrary, N is not explosive. □

We now illustrate Theorem 5.1.3 by applying it to the simpler model with chemistry $0 \leftrightarrows S$.

Corollary 5.1.4: Let N be the coarse-grained model associated to (5.2):

$$0 \underset{\kappa_d}{\overset{\kappa_b}{\rightleftarrows}} S \qquad\qquad 0 \underset{\kappa_E}{\overset{\kappa_I}{\rightleftarrows}} C \underset{\kappa_C}{\overset{\kappa_F S_C}{\rightleftarrows}} 2C \qquad\qquad \mu.$$

For any choice of parameters, N is not explosive. (We do not assume μ has finite expectation here.)

Proof. Let $f(x) = x$. Notice that

$$\mathcal{A}f(x) = \kappa_b - \kappa_d x \leq \kappa_b.$$

Therefore, N is not explosive by Theorem 5.1.3. □

5.2 Positive Recurrence

Once again the key property of $\mathcal{I}_\mathcal{K}$ will be the existence of a linear Lyapunov function.

Theorem 5.2.1: Let N be the coarse-grained model associated to (5.1):

$$\mathcal{I}_\mathcal{K} \qquad\qquad 0 \underset{\kappa_E}{\overset{\kappa_I}{\rightleftarrows}} C \underset{\kappa_C}{\overset{\kappa_F S_C}{\rightleftarrows}} 2C \qquad\qquad \mu.$$

Let $\mathcal{A}$ denote the generator of the associated chemistry $\mathcal{I}_\mathcal{K}$. Suppose there exists some linear function $f : \mathbb{Z}^d_{\geq 0} \to \mathbb{R}$ of the form $f(x) = w \cdot x$ for some $w \in \mathbb{R}^d_{>0}$, such that $\sup_x \mathcal{A}f(x) < \infty$. Suppose condition 5.0.2 is satisfied: $\lambda < \infty$. If $\kappa_C > 0$ and $\kappa_E > 0$, then the state with no compartments is positive recurrent for N, all states reachable from that one are also positive recurrent for N, and all other states are transient with finite expected time to reach the set of positive recurrent states.

Remark 5.2.2: If f were a Lyapunov function witnessing (via Theorem 6.1.2) that $\mathcal{I}_{\mathcal{K}}$ was positive recurrent, we would have $\mathcal{A}f(x) \leq -1$ outside some finite set. Similarly, if f were witnessing that $\mathcal{I}_{\mathcal{K}}$ was recurrent, we would have $\mathcal{A}f(x) \leq 0$ outside some finite set. Either of these would imply the assumption $\sup_x \mathcal{A}f(x) < \infty$ from the theorem. This assumption $\sup_x \mathcal{A}f(x) < \infty$ is in fact *strictly* weaker than having a Lyapunov function for recurrence, as we will see in 5.2.3 when we apply it with the transient chemistry $0 \to S$. However, the assumption $\sup_x \mathcal{A}f(x) < \infty$ is stronger than having a Lyapunov function for non-explosivity. Indeed, if $\sup_x \mathcal{A}f(x) < \infty$ and $f \to \infty$, then a suitable shift g of f will satisfy $\mathcal{A}g(x) \leq g(x)$ (this is easiest to see by taking g to be a shift of f satisfying $\mathcal{A}g(x) = \mathcal{A}f(x) \leq \sup_y \mathcal{A}f(y) \leq \inf_y g(y) \leq g(x)$). △

Proof of Theorem 5.2.1. Let $V(n) = C(n) + \alpha \sum_{x \in \mathbb{Z}^d_{\geq 0}} n_x(w \cdot x)$, for some constant $\alpha > 0$ to be chosen later. The assumption that every coordinate of w is strictly positive means that $V \to \infty$ in the sense of Theorem 6.1.2.

$$\begin{aligned}
\mathcal{L}V(n) = &\sum_{x \in \mathbb{Z}^d_{\geq 0}} \kappa_F S(x) n_x + \kappa_I \mu(x)(1 + \alpha w \cdot x) + \alpha \sum_{\nu \to \nu'} n_x \lambda_{\nu \to \nu'}(x)\big(w \cdot (x + \nu' - \nu) - w \cdot x\big) \\
&\qquad - \kappa_E n_x(1 + \alpha w \cdot x) - \kappa_C \binom{n_x}{2} - \sum_{\substack{y \in \mathbb{Z}^d_{\geq 0} \\ y \neq x}} \kappa_C \frac{n_x n_y}{2} \\
= &\left(\sum_{x \in \mathbb{Z}^d_{\geq 0}} \kappa_F S(x) n_x - \alpha \kappa_E n_x (w \cdot x) + \kappa_I \mu(x)(1 + \alpha w \cdot x) + \alpha n_x \mathcal{A}f(x) \right) \\
&- \kappa_C \binom{C(n)}{2} - \kappa_E C(n) \\
\leq &-\kappa_C \binom{C(n)}{2} + \left(\alpha \sup_x \mathcal{A}f(x) - \kappa_E \right) C(n) + (\kappa_F w_S^{-1} - \alpha \kappa_E) \left(\sum_{x \in \mathbb{Z}^d} n_x (w \cdot x) \right) \\
&+ \kappa_I + \alpha \kappa_I (\max_j w_j) \lambda.
\end{aligned}$$

Pick α large enough that $\alpha \kappa_E > \kappa_F w_S^{-1}$ (here we used $\kappa_E > 0$). Then we claim this upper bound is at most -1 outside a finite set of n. Indeed, the term $(\kappa_F w_S^{-1} - \alpha \kappa_E)\left(\sum_{x \in \mathbb{Z}^d} n_x (w \cdot x)\right)$ is non-positive by choice of α, so $\mathcal{L}V(n)$ is at most a polynomial

in $C(n)$ (with no dependence on n other than through $C(n)$ and) with negative leading term (here we used $\kappa_C > 0$). So for large enough values of $C(n)$ we have $\mathcal{L}V(n) \leq -1$. For each individual value of $C(n)$ less than this threshold, the only term that varies as n varies is $(\kappa_F w_S^{-1} - \alpha\kappa_E)\left(\sum_{x\in\mathbb{Z}^d} n_x(w\cdot x)\right)$, and this approaches $-\infty$ as $\sum_{x\in\mathbb{Z}^d} n_x(w\cdot x) \to \infty$. It follows that $\mathcal{L}V(n) \leq -1$ outside some finite set, as claimed.

Therefore, by Theorem 6.1.2 any state in a closed, irreducible component of $\mathcal{N}$ is positive recurrent, and from any given state the expected time for the process to enter the union of the closed irreducible components is finite. So to complete the proof it remains only to point out that the state with no compartment is a member of a (indeed, *the*) closed irreducible component of $\mathcal{N}$; this fact follows from the fact that $\kappa_E > 0$ and hence the state with no compartments is reachable from every state in $\mathcal{N}$. □

We now illustrate this theorem by applying it to the simpler model with chemistry $0 \leftrightarrows S$. Note that the result holds even when $\mathcal{I}_{\mathcal{K}}$ is transient.

Corollary 5.2.3: Let N be the coarse-grained model associated to (5.2):

$$0 \underset{\kappa_d}{\overset{\kappa_b}{\rightleftarrows}} S \qquad\qquad 0 \underset{\kappa_E}{\overset{\kappa_I}{\rightleftarrows}} C \underset{\kappa_C}{\overset{\kappa_F S_C}{\rightleftarrows}} 2C \qquad\qquad \mu.$$

Suppose that condition 5.0.2 is satisfied: $\lambda < \infty$. If both $\kappa_C > 0$ and $\kappa_E > 0$, then the state $(0, 0, 0, \cdots)$ with no compartments is positive recurrent. Moreover,

- If $\kappa_I > 0$ and either $\kappa_b > 0$ or $\mu \neq \delta_0$, then all states are reachable from $(0, 0, 0, \cdots)$ and hence positive recurrent.
- If $\kappa_I > 0$ and both $\kappa_b = 0$ and $\mu = \delta_0$, then all states with zero S are positive recurrent and all other states are transient. These other states all have finite expected time to be absorbed by the collection of zero-S states.
- If $\kappa_I = 0$, then all states with a positive number of compartments are transient, but are absorbed by the state with zero compartments in finite expected time.

Proof. Let $f(x) = x$, and notice that $\mathcal{A}f(x) \leq \kappa_b$. Therefore, the state with no compartments is positive recurrent for N by Theorem 5.2.1. The "moreover" part of this corollary follows straightforward considerations about which states are reachable from the state with no compartments. □

In any specific model, one may be able to do better than Theorem 5.2.1 by using a more tailored Lyapunov function. The next proposition provides an example of this by extending Corollary 5.2.3.

Proposition 5.2.4: Let N be the coarse-grained model associated to (5.2):

$$0 \underset{\kappa_d}{\overset{\kappa_b}{\rightleftarrows}} S \qquad\qquad 0 \underset{\kappa_E}{\overset{\kappa_I}{\rightleftarrows}} C \underset{\kappa_C}{\overset{\kappa_F S_C}{\rightleftarrows}} 2C \qquad\qquad \mu.$$

Suppose that condition 5.0.2 is satisfied: $\lambda < \infty$. Consider the following conditions on the parameters of the model:

(a) $\kappa_C > 0$ and $\kappa_E > 0$.

(b) $\kappa_E^2 + \kappa_E \kappa_d > \kappa_b \kappa_F$

(c) $\kappa_C > 0$ and $\kappa_d > 0$ and $\kappa_E = 0$

If condition (a) or (b) is satisfied, the state $(0, 0, 0, \cdots)$ with no compartments is positive recurrent. Moreover,

- If $\kappa_I > 0$ and either $\kappa_b > 0$ or $\mu \neq \delta_0$, then all states are reachable from $(0, 0, 0, \cdots)$ and hence positive recurrent.
- If $\kappa_I > 0$ and both $\kappa_b = 0$ and $\mu = \delta_0$, then all states with zero S are positive recurrent and all other states are transient. These other states all have finite expected time to be absorbed by the collection of zero-S states.
- If $\kappa_I = 0$, then all states with a positive number of compartments are transient, but are absorbed by the state with zero compartments in finite expected time.

If condition (c) is satisfied, the state $(1, 0, 0, \cdots)$ with one empty compartment is positive recurrent. Moreover,

- If $\kappa_I \kappa_b > 0$, or $\kappa_F \kappa_b > 0$, or $\kappa_I > 0$ and $\mu \neq \delta_0$, then all states other than $(0, 0, 0, \cdots)$ are reachable from $(1, 0, 0, \cdots)$ and hence positive recurrent.

- If $\kappa_I > 0$ and both $\kappa_b = 0$ and $\mu = \delta_0$, then all states with zero S and a positive number of compartments are positive recurrent and all other states are transient. These other states all have finite expected time to be absorbed by the collection of zero-S states.

- If $\kappa_b > 0$ but $\kappa_I = 0 = \kappa_F$, then all states with one compartment are positive recurrent, and all other states are transient and have finite expected time to be absorbed by the one-compartment states.

- If $\kappa_I = 0$ and $\kappa_b = 0$, then the state with no compartments is absorbing, as is the state with one empty compartment, and all other states are transient. These other states have finite expected time to be absorbed by the state with one empty compartment.

Remark 5.2.5: Note that the above covers conditions (a), (b), and (c) completely: All possible combinations of other parameters are listed in the bullet points following each "Moreover". Furthermore, note that (a) and (c) together cover every possible case where $\kappa_C > 0$ except the one where $\kappa_E = \kappa_d = 0$, which is not a very interesting case since the number of S cannot decrease. So the proposition is essentially complete except for the parameter regime where $\kappa_C = 0$ and $\kappa_E^2 + \kappa_E \kappa_d \leq \kappa_b \kappa_F$. Later, in Proposition 5.2.7, we will prove transience in part (though not all) of this remaining regime. △

Proof of Proposition 5.2.4. The claims about condition (a) are just repeating Corollary 5.2.3. For the remaining conditions, let $\mathcal{L}$ denote the generator of N. For $n \in \mathcal{N}$, let $C(n) = \sum_{x=0}^{\infty} n_x$ and $S(n) = \sum_{x=0}^{\infty} x n_x$ be the total number of compartments and the total number of S across compartments, respectively. Let $V(n) = \alpha S(n) + C(n)$ for some

constant $\alpha > 0$ to be chosen later. Then

$$\begin{aligned}\mathcal{L}V(n) &= -\kappa_C \binom{C(n)}{2} + \kappa_F S(n) - \kappa_E C(n) - \alpha\kappa_E S(n) + \alpha\kappa_b C(n) - \alpha\kappa_d S(n) \\ &\qquad + \sum_{x=0}^{\infty} \kappa_I \mu(x)\big(1 + xa\big) \\ &= -\kappa_C \binom{C(n)}{2} + (\kappa_F - \alpha(\kappa_E + \kappa_d))S(n) + (\alpha\kappa_b - \kappa_E)C(n) + \kappa_I + \kappa_I \lambda\alpha.\end{aligned}$$

If $\kappa_C > 0$ and $\kappa_d + \kappa_E > 0$, then picking $\alpha > \kappa_F/(\kappa_d + \kappa_E)$ we have that $\mathcal{L}V(n) \leq -1$ outside a finite set. The claims about condition (c) now follow from Theorem 6.1.2 and straightforward reachability considerations. (This also gives an alternate proof of the claims about condition (a).)

If $\kappa_E^2 + \kappa_E\kappa_d > \kappa_b\kappa_F$, then $\kappa_F/(\kappa_E + \kappa_d) < \kappa_E/\kappa_b$; pick α to be some number satisfying $\kappa_F/(\kappa_E + \kappa_d) < \alpha < \kappa_E/\kappa_b$ (with the convention that $\kappa_E/\kappa_b = \infty$ if $\kappa_b = 0$). Then the coefficients of both $S(n)$ and $C(n)$ are negative, so $\mathcal{L}V(n) \leq -1$ outside a finite set. The claims about condition (b) follow just as above. □

Remark 5.2.6: In Section 2.B of [1], the model (5.2) is studied via simulation and moment closure methods in the case where $\kappa_b = 0 = \kappa_d$, where μ is Poisson with parameter λ, and where $\psi(x, y)$ is uniform over possible unordered pairs $\{y, x - y\}$. They study a moment-closure approximation of the model for finite time scales, and do not deal with the long term behavior of their model. As it turns out, their model is positive recurrent. This follows from Proposition 5.2.4; one could use either condition (a) or (b). In fact, since Proposition 5.2.4(a) is just Corollary 5.2.3, the existence of the stationary distribution is essentially free from Theorem 5.2.1. △

In the simple model with chemistry $0 \leftrightarrows S$, we are also able to prove results about transience. We will assume that $\kappa_C = 0$, since by Proposition 5.2.4 conditions (a) and (c) we know that if $\kappa_C > 0$ then N is positive recurrent as long as $\kappa_d > 0$ or $\kappa_E > 0$, and the case where $\kappa_d = 0 = \kappa_E$ is not interesting because the total number of molecules of S across all compartments cannot shrink.

Proposition 5.2.7: Let N be the coarse-grained model associated to (5.2)

$$0 \underset{\kappa_d}{\overset{\kappa_b}{\rightleftarrows}} S \qquad\qquad 0 \underset{\kappa_E}{\overset{\kappa_I}{\rightleftarrows}} C \underset{\kappa_C}{\overset{\kappa_F S_C}{\rightleftarrows}} 2C \qquad\qquad \mu,$$

where we again assume condition 5.0.2: $\lambda < \infty$. If $(\kappa_F - \kappa_E)\kappa_b > (\kappa_E + \kappa_d)\kappa_E$ and $\kappa_I > 0$ and $\kappa_C = 0$, then all states are transient for N.

Proof. Let $C(n) = \sum_{x=0}^{\infty} n_x$ and $C_{>0}(n) = \sum_{x=1}^{\infty} n_x$, and let $W(n) = C(n) + \alpha C_{>0}(n)$ for some constant $\alpha > 0$ to be chosen later. Let $V(n) = 1 - \frac{1}{W(n)+1}$.

Notice that (with the convention that $0/0 = 0$),

$$\begin{aligned}
\mathcal{L}V(n) = \sum_{x=0}^{\infty} &\kappa_b n_x \big(V(n - e_x + e_{x+1}) - V(n)\big) + \kappa_d n_x x \big(V(n - e_x + e_{x-1}) - V(n)\big) \\
&+ \kappa_I \mu(x)\big(V(n + e_x) - V(n)\big) + \kappa_E n_x \big(V(n - e_x) - V(n)\big) \\
&+ \kappa_F x n_x \left(\sum_{y=0}^{\infty} \psi(x,y)\big(V(n - e_x + e_y + e_{x-y}\big) - V(n)\big)\right) \\
\geq\ & \kappa_b n_0 \left(\frac{1}{W(n)+1} - \frac{1}{W(n)+1+\alpha}\right) + \kappa_d C_{>0}(n) \left(\frac{1}{W(n)+1} - \frac{1}{W(n)+1-\alpha}\right) \\
&+ \kappa_E n_0 \left(\frac{1}{W(n)+1} - \frac{1}{W(n)}\right) + \kappa_E C_{>0}(n) \left(\frac{1}{W(n)+1} - \frac{1}{W(n)-\alpha}\right) \\
&+ (\kappa_F C_{>0}(n) + \kappa_I) \left(\frac{1}{W(n)+1} - \frac{1}{W(n)+2}\right) \\
=\ & \frac{\alpha\kappa_b n_0}{(W(n)+1)(W(n)+1+\alpha)} + \frac{\kappa_F C_{>0}(n) + \kappa_I}{(W(n)+1)(W(n)+2)} \\
&- \frac{\kappa_E n_0}{(W(n)+1)(W(n))} - \frac{(1+\alpha)\kappa_E C_{>0}(n)}{(W(n)+1)(W(n)-\alpha)} - \frac{\alpha\kappa_d C_{>0}(n)}{(W(n)+1)(W(n)+1-\alpha)}
\end{aligned}$$

Multiplying through by $(W(n)+1)(W(n)+2)$, the above becomes

$$\begin{aligned}
(W(n)+1)(W(n)+2)\mathcal{L}V(n) = \alpha\kappa_b n_0 \frac{W(n)+2}{W(n)+1+\alpha} + \kappa_F C_{>0}(n) + \kappa_I - \kappa_E n_0 \frac{W(n)+2}{W(n)} \\
- (1+\alpha)\kappa_E C_{>0}(n)\frac{W(n)+2}{W(n)-\alpha} - \alpha\kappa_d C_{>0}(n)\frac{W(n)+2}{W(n)+1-\alpha}.
\end{aligned}$$

Let $\varepsilon > 0$ be another constant that we will pick later. Notice that all four fractions immediately above are converging to 1 as $W(n) \to \infty$. Let B_ε be a set of the form

$\{n : W(n) \leq k_\varepsilon\}$ for some number k_ε, such that the first fraction is bounded below by $1-\varepsilon$ on B_ε^c and the latter three are bounded above by $1+\varepsilon$ on B_ε^c. But then off of B_ε,

$$\begin{aligned}(W(n)+1)(W(n)+2)\mathcal{L}V(n) &\geq \alpha\kappa_b n_0(1-\varepsilon) + \kappa_F C_{>0}(n) + \kappa_I - \kappa_E n_0(1+\varepsilon) \\ &\quad - (1+\alpha)\kappa_E C_{>0}(n)(1+\varepsilon) - \alpha\kappa_d C_{>0}(n)(1+\varepsilon) \\ &= \big(\kappa_F - (1+\alpha)\kappa_E(1+\varepsilon) - \alpha\kappa_d(1+\varepsilon)\big)C_{>0}(n) \\ &\quad + \big(\alpha\kappa_b(1-\varepsilon) - \kappa_E(1+\varepsilon)\big)n_0 + \kappa_I\end{aligned}$$

Since $(\kappa_F - \kappa_E)\kappa_b > (\kappa_E + \kappa_d)\kappa_E$, in particular $(\kappa_F - \kappa_E)/(\kappa_E + \kappa_d) > \kappa_E/\kappa_b$; pick $\alpha > 0$ so that $(\kappa_F - \kappa_E)/(\kappa_E + \kappa_d) > \alpha > \kappa_E/\kappa_b$. Rearranging these two inequalities gets us $\alpha\kappa_b - \kappa_E > 0$ and $\kappa_F - (1+\alpha)\kappa_E - \alpha\kappa_d > 0$. So for some sufficiently small $\varepsilon > 0$, we have $\alpha\kappa_b(1-\varepsilon) - \kappa_E(1+\varepsilon) > 0$ and $\kappa_F - (1+\alpha)\kappa_E(1+\varepsilon) - \alpha\kappa_d(1+\varepsilon) > 0$. Let ε be such; outside B_ε, we have

$$\begin{aligned}(W(n)+1)(W(n)+2)\mathcal{L}V(n) &\geq 0 \\ \mathcal{L}V(n) &\geq 0.\end{aligned}$$

Then, since $\sup_{n\in B_\varepsilon} V(n) < \inf_{n\in B_\varepsilon^c} V(n)$ and since N is not explosive by Corollary 5.1.4, Theorem 6.1.3 tells us that when N is started outside of B_ε, the probability that it eventually reaches B_ε is less than one. Notice that (since $\kappa_I > 0$), when started from $(0,0,0,\cdots)$ the Markov chain N reaches $(k_\varepsilon,0,0,0,\cdots) \in B_\varepsilon^c$ with positive probability, and then by above with positive probability it never returns to $(0,0,0,\cdots)$. So $(0,0,0,\cdots)$ is transient. Now either $\kappa_E = 0$ and every state is obviously transient, or $\kappa_E > 0$ and the transient state $(0,0,0,\cdots)$ is reachable from every state. So in either case, every state is transient, as claimed. □

Remark 5.2.8: Notice the gap between the previous two results: If $\kappa_C = 0$ and $\kappa_I > 0$, the former says that N is positive recurrent if $\kappa_E^2 + \kappa_E\kappa_d > \kappa_b\kappa_F$ whereas the latter says that N is transient if $\kappa_E^2 + \kappa_E\kappa_d < \kappa_b\kappa_F - \kappa_b\kappa_E$. We conjecture that N is transient in the

case where $\kappa_b\kappa_F > \kappa_E^2 + \kappa_E\kappa_d \geq \kappa_b\kappa_F - \kappa_b\kappa_E$.

Limited numerical simulation supports this conjecture. With a team of four undergraduate students (Carina Guo, Olivia Guo, Leo Shen, and Yikai Zhang), we simulated the model of this chapter using a combination of the Gillespie and next reaction algorithms (see [6], specifically Algorithms 1 and 2, for background on these two methods). Specifically, simulations were done with parameters $\kappa_C = 0$, $\kappa_b = \kappa_d = \kappa_E = \kappa_I = 1$, and $\kappa_F \in \{1.9, 2.0, 2.1\}$. By the theorem above, the system should be positive recurrent with $\kappa_F = 1.9$, whereas $\kappa_F = 2.0$ and $\kappa_F = 2.1$ fall into the gap (with $\kappa_F = 2$ right on the boundary and not covered by the conjecture above). Indeed, the trajectories kept returning to zero in the $\kappa_F = 1.9$ case, they were steadily increasing when $\kappa_F = 2.1$, and the $\kappa_F = 2.0$ case was less clear. △

Chapter 6

Technical Results

This chapter is devoted to stating and proving any technical results needed for the thesis.

6.1 Lyapunovfunctionology

The theory of Lyapunov functions for Markov chains provides powerful tools for checking (in)stability properties. This section is devoted to introducing the extent of the theory we use.

The statement and proof of Theorem 6.1.1 are heavily inspired by Theorem 2.1 of [28], and one can check that our theorem is a corollary of theirs. Our proof is slightly different because we prove the result by truncating the Lyapunov function instead of truncating the process. The key idea for this alternate truncation was drawn from [29].

Theorem 6.1.1: Let X be a continuous-time Markov chain on a countable state space $\mathbb{S}$ with generator $\mathcal{L}$. Suppose V is a function on $\mathbb{S}$ which satisfies $V \to \infty$ in the sense that $\{x \in \mathcal{S} : V(x) < B\}$ is finite for every $B > 0$. If there are constants $c, d \geq 0$ so that $\mathcal{L}V(x) \leq cV(x) + d$ for all x, then X is not explosive.

Proof. Observe that it is enough to consider the case where $V \geq 0$ and $\mathcal{L}V(x) \leq cV(x)$. Indeed, the condition $V \to \infty$ grants that V attains a global minimum value. So a suitable shift of V, say W, will have minimum at least $d \geq 0$ and so the condition $\mathcal{L}V(x) \leq cV(x)+d$

implies $\mathcal{L}W(x) \leq (c+1)W(x)$.

Now fix an enumeration $x_1, x_2, \cdots$ of $\mathbb{S}$, and for $m = 1, 2, 3, \cdots$ let $O_m = \{x_1, \cdots, x_m\}$ and let T_m denote the first time the process (X_t) is not in O_m. Notice that to show that X is not explosive, it suffices to show that $\zeta := \lim_{m\to\infty} T_m$ is infinite $\mathbb{P}_x$-a.s. for each $x \in \mathbb{S}$.

For $m = 1, 2, 3, \cdots$, set $g(x,t) = V(x)e^{-ct}$ and set $f(x,t) = g(x,t)\mathbb{I}_{O_m}(x)$. The indicator function ensures that the support of f in the variable x lies in O_m and that f is uniformly bounded in both variables. So we can apply Dynkin's formula (Lemma 6.2.2) to f and conclude that if τ is any a.s. bounded stopping time,

$$\mathbb{E}_x[f(X_\tau, \tau)] = f(x,0) + \mathbb{E}_x\left[\int_0^\tau \mathcal{L}f(X_s, s)\mathrm{d}s\right].$$

One can check that since V was assumed non-negative, $\mathcal{L}f(x,s) \leq \mathcal{L}g(x,s)$ for all $x \in O_m$ and all s. This follows since $x \in O_m$ one has

$$\mathcal{L}g(x,s) - \mathcal{L}f(x,s) = e^{-ct}\sum_{y\neq x} q_{xy}\mathbb{I}_{y\notin O_m}V(y) \geq 0.$$

So if $x \in O_m$, then for any a.s. bounded stopping time τ, we get

$$\begin{aligned}
\mathbb{E}_x[f(X_{\tau\wedge T_m}, \tau\wedge T_m)] &\leq f(x,0) + \mathbb{E}_x\left[\int_0^{\tau\wedge T_m} \mathcal{L}g(X_s, s)\mathrm{d}s\right] \\
&= V(x) + \mathbb{E}_x\left[\int_0^{\tau\wedge T_m} e^{-cs}\Big(\mathcal{L}V(X_s) - cV(X_s)\Big)\mathrm{d}s\right] \\
&\leq V(x),
\end{aligned}$$

where in the first line we use the fact that $X_s \in O_m$ when $s < \tau \wedge T_m$.

Now consider the process $M_t := f(X_t, t)\mathbb{I}_{t<T_m} = V(X_t)e^{-ct}\mathbb{I}_{t<T_m}$. Notice that M is the product of right-continuous functions and hence itself right-continuous; we claim that M is a supermartingale. Toward this end, fix $s < t$; we wish to show that $\mathbb{E}[M_t|\mathcal{F}_s] \leq M_s$ where $\mathcal{F}_s$ denotes the σ-algebra generated by $(X_r)_{r\leq s}$. Since the event $\{s \geq T_m\} \in \mathcal{F}_s$,

we can decompose

$$\mathbb{E}[M_t|\mathcal{F}_s] = \mathbb{E}[M_t|\mathcal{F}_s]\mathbb{I}_{s\geq T_m} + \mathbb{E}[M_t|\mathcal{F}_s]\mathbb{I}_{s<T_m},$$

so it is enough to show the desired inequality both on and off the event $\{s \geq T_m\}$. On this event, we have $M_s = 0 = M_t$, and so $\mathbb{E}[M_t|\mathcal{F}_s]\mathbb{I}_{s\geq T_m} = 0 = M_s\mathbb{I}_{s\geq T_m}$. On the event $\{s < T_m\}$, notice that $X_s \in O_m$, and hence the bound $\mathbb{E}_{X_s}[f(X_{\tau\wedge T_m}, \tau \wedge T_m)] \leq V(X_s)$ from above applies for any a.s. bounded stopping time τ (and in particular, for $\tau = t - s$). So by the Markov property,

$$\begin{aligned}
\mathbb{E}[M_t|\mathcal{F}_s]\mathbb{I}_{s<T_m} &= \mathbb{E}[V(X_t)e^{-ct}\mathbb{I}_{t<T_m}|\mathcal{F}_s]\mathbb{I}_{s<T_m} \\
&= e^{-cs}\mathbb{E}_{X_s}[V(X_{t-s})e^{-c(t-s)}\mathbb{I}_{t-s<T_m}]\mathbb{I}_{s<T_m} \\
&= e^{-cs}\mathbb{E}_{X_s}[f(X_{t-s}, t-s)\mathbb{I}_{t-s<T_m}]\mathbb{I}_{s<T_m} \\
&\leq e^{-cs}\mathbb{E}_{X_s}[f(X_{(t-s)\wedge T_m}, (t-s)\wedge T_m)]\mathbb{I}_{s<T_m} \\
&\leq e^{-cs}V(X_s)\mathbb{I}_{s<T_m} \\
&= M_s\mathbb{I}_{s<T_m}.
\end{aligned}$$

We conclude that M is a supermartingale, as claimed. Now fix $x \in \mathbb{S}$ and $\lambda > 0$. By Doob's Supermartingale inequality (Lemma 6.2.4),

$$\begin{aligned}
\lambda^{-1}V(x)\mathbb{I}_{O_m}(x) &\geq \mathbb{P}_x\left[\sup_{t\in[0,\infty)} M_t \geq \lambda\right] \\
&= \mathbb{P}_x\left[\sup_{0\leq t<T^m} V(X_t)e^{-ct} \geq \lambda\right].
\end{aligned}$$

Sending $m \to \infty$ and applying continuity of probability yields

$$\mathbb{P}_x\left[\sup_{0\leq t<\zeta} V(X_t)e^{-ct} \geq \lambda\right] \leq \lambda^{-1}V(x). \tag{6.1}$$

Notice that

$$\sup_{0\le t<\zeta} V(X_t)e^{-ct} \ge e^{-c\zeta} \sup_{0\le t<\zeta} V(X_t).$$

This lower bound is ∞ when $\zeta < \infty$ thanks to the condition $V \to \infty$, and hence we have $\sup_{0\le t<\zeta} V(X_t)e^{-ct} \ge \lambda$ on the set where $\zeta < \infty$, regardless of λ. From (6.1), then,

$$\mathbb{P}_x(\zeta < \infty) \le \lambda^{-1} V(x).$$

But $\lambda > 0$ was arbitrary; sending $\lambda \to \infty$ yields that $\mathbb{P}_x(\zeta < \infty) = 0$. Since x was arbitrary, it follows that X is not explosive. □

The following theorem is well-known. In full generality, it is due to Meyn and Tweedie — see [28]. The version below is a specialization to the countable state space case. For a proof of the version given below, see the more recent paper [30].

Theorem 6.1.2: Let X be a continuous-time Markov chain on a countable state space $\mathbb{S}$ with generator $\mathcal{L}$. Suppose there exists a finite set $K \subset \mathbb{S}$ and a positive function V on $\mathbb{S}$ such that

$$\mathcal{L}V(x) \le -1$$

for all $x \in \mathbb{S} \setminus K$. Suppose further that $V \to \infty$ in the sense that $\{x \in \mathbb{S} : V(x) < B\}$ is finite for every $B > 0$. Then each state in a closed, irreducible component of $\mathbb{S}$ is positive recurrent. Moreover, if τ_{x_0} is the time for the process to enter the union of the closed irreducible components given an initial condition x_0, then $\mathbb{E}_{x_0}[\tau_{x_0}] < \infty$.

We will also need the following, which provides a method to check for transience.

Theorem 6.1.3: Let X be a non-explosive continuous-time Markov chain on a countable discrete state space $\mathbb{S}$ with generator $\mathcal{L}$. Let $B \subset \mathbb{S}$, and let τ_B be the time for the process to enter B. Suppose there is some bounded function V such that for all $x \in B^c$,

$$\mathcal{L}V(x) \ge 0.$$

Then $\mathbb{P}_{x_0}(\tau_B < \infty) < 1$ for any x_0 such that

$$\sup_{x \in B} V(x) < V(x_0).$$

For a version of the theorem above that applies in much greater generality, see Theorem 3.3(i) in [31]. Our theorem is not an immediate corollary of theirs (they define restricted versions of the chain X and state their theorem in terms of the generators of the restricted processes), so we will provide a proof (which, just like the theorem statement, draws heavy inspiration from [31]):

Proof of Theorem 6.1.3. Define W on $\mathbb{S}$ via $W = V - \sup_{x \in B} V(x)$. Notice that $W(x_0)$ is strictly positive, W is nonpositive on B, and $\mathcal{L}W = \mathcal{L}V$. Fix some enumeration of $\mathbb{S}$ in which x_0 is the first element, and for $m \in \mathbb{N}$ let $\mathbb{S}_m$ denote the first m elements of $\mathbb{S}$. Let τ_m be the first time X is not in $\mathbb{S}_m$. Let Δ be a new state not in $\mathbb{S}$, and for $m \in \mathbb{N}$ define a new Markov chain X^m via

$$X_t^m = \begin{cases} X_t & t < \tau_m \\ \Delta & t \geq \tau_m \end{cases}$$

Notice that X^m has finite state space $\mathbb{S}_m \cup \{\Delta\}$. Notice that W is bounded since V is, let $C = \sup_{x \in \mathbb{S}} W(x)$, and extend W to a function on $\mathbb{S} \cup \{\Delta\}$ by setting $W(\Delta) = C$. Let $\mathcal{L}_m$ denote the generator of the process X^m; we claim that $\mathcal{L}W(x) \leq \mathcal{L}_m W(x)$ whenever

$x \in \mathbb{S}_m$. Indeed, notice that

$$\begin{aligned}
\mathbb{E}_x[W(X_t)] &= \sum_{y\in\mathbb{S}} W(y)\mathbb{P}_x(X_t = y) \\
&= \sum_{y\in\mathbb{S}} W(y)\mathbb{P}_x(X_t = y, t < \tau_m) + \sum_{y\in\mathbb{S}} W(y)\mathbb{P}_x(X_t = y, t \geq \tau_m) \\
&\leq \sum_{y\in\mathbb{S}} W(y)\mathbb{P}_x(X_t = y, t < \tau_m) + \sum_{y\in\mathbb{S}} C\mathbb{P}_x(X_t = y, t \geq \tau_m) \\
&= \sum_{y\in\mathbb{S}_m} W(y)\mathbb{P}_x(X_t^m = y) + W(\Delta)\mathbb{P}_x(X_t^m = \Delta) \\
&= \mathbb{E}_x[W(X_t^m)],
\end{aligned}$$

and hence

$$\mathcal{L}W(x) = \lim_{t\searrow 0} \frac{\mathbb{E}_x[W(X_t)] - W(x)}{t} \leq \lim_{t\searrow 0} \frac{\mathbb{E}_x[W(X_t^m)] - W(x)}{t} = \mathcal{L}_m W(x),$$

as claimed. Now for any m, applying Dynkin's Formula (Lemma 6.2.1) to the chain X^m with finite stopping time $\tau_B \wedge \tau_m \wedge m$ yields

$$\mathbb{E}_{x_0}[W(X^m_{\tau_B\wedge\tau_m\wedge m})] = W(x_0) + \mathbb{E}_{x_0}\left[\int_0^{\tau_B\wedge\tau_m\wedge m} \mathcal{L}_m W(X_s^m)\mathrm{d}s\right].$$

But for $s < \tau_B \wedge \tau_m$ we have $X_s^m = X_s \in B^c \cap \mathbb{S}_m$ and hence

$$\mathcal{L}_m W(X_s^m) = \mathcal{L}_m W(X_s) \geq \mathcal{L}W(X_s) = \mathcal{L}V(X_s) \geq 0.$$

So the integrand in Dynkin's Formula is non-negative, and

$$\begin{aligned}
W(x_0) &\leq \mathbb{E}_{x_0}[W(X^m_{\tau_B\wedge\tau_m\wedge m})] \\
&= \mathbb{E}_{x_0}[W(X^m_{\tau_B})\mathbb{I}_{\tau_B<\tau_m\wedge m}] + \mathbb{E}_{x_0}[W(X^m_{\tau_m\wedge m})\mathbb{I}_{\tau_B\geq\tau_m\wedge m}] \\
&\leq \mathbb{E}_{x_0}[W(X^m_{\tau_B})\mathbb{I}_{\tau_B<\tau_m\wedge m}] + C\mathbb{P}_{x_0}(\tau_B \geq \tau_m \wedge m).
\end{aligned}$$

Note that $X^m_{\tau_B} \in B$ on the event $\tau_B < \tau_m \wedge m$. Hence $W(X^m_{\tau_B})\mathbb{I}_{\tau_B<\tau_m\wedge m} \leq 0$, and

$$W(x_0) \leq C\mathbb{P}_{x_0}(\tau_B \geq \tau_m \wedge m)$$

Since X is assumed to be non-explosive, $\tau_m \to \infty$ as $m \to \infty$, so taking $m \to \infty$ above gives

$$W(x_0) \leq C\mathbb{P}_{x_0}(\tau_B = \infty).$$

But $W(x_0)$ is strictly positive and $0 < W(x_0) \leq C < \infty$, so $\mathbb{P}_{x_0}(\tau_B = \infty) \neq 0$. That is, $\mathbb{P}_{x_0}(\tau_B < \infty) < 1$, as desired. □

Remark 6.1.4: Note that the proof above gives us a lower bound for the probability that the process never returns to the set B:

$$\frac{W(x_0)}{C} \leq \mathbb{P}_{x_0}(\tau_B = \infty),$$

where $C = \sup_{x\in\mathbb{S}} W(x)$ and $W = V - \sup_{x\in B} V(x)$. We do not make use of this fact. △

The following is a simplified version of Theorem 8.4.3 from [32] (expect that their theorem is in discrete time, but that doesn't matter), and this proof is essentially theirs.

Theorem 6.1.5: Suppose X is an irreducible CTMC on $\mathbb{S}$. Suppose $B \subseteq \mathbb{S}$ is finite, and suppose $V : X \to \mathbb{R}$ with $V \to \infty$ (in the sense that $\{V < M\}$ is finite for each M) satisfies $\mathcal{L}V(x) \leq 0$ for $x \notin B$. Then X is recurrent.

Proof. By shifting V, we may suppose without loss of generality that $V : X \to [0, \infty)$. Let Y be the jump chain for X. Then $\mathcal{L}V(x) \leq 0$ iff $\Delta V(x) := \mathbb{E}_x[V(Y_1)] - V(x) \leq 0$ (you can check that $\mathcal{L}V(x)$ and $\Delta V(x)$ differ by a positive ratio, namely, the total rate of X out of state x). So our approach will be to prove that Y is recurrent using $V(x) \geq \mathbb{E}_x[V(Y_1)]$ for $x \notin B$.

Suppose toward a contradiction that $\{Y_k\}$ is transient. Then there is some $x_0 \notin B$ such that $\mathbb{P}_{x_0}(T_B = \infty) > 0$ where $T_B = \inf\{k \geq 0 : Y_k \in B\}$; let x_0 be such. Let M

denote some number large enough that

$$V(x_0) < M\mathbb{P}_{x_0}(T_B = \infty).$$

Define a new process $\widehat{Y}$ via $\widehat{Y}_k = Y_{k\wedge T_B}$ (that is, take Y and stop it upon entering B). Notice that $V(x) \geq \mathbb{E}_x[V(\widehat{Y}_1)]$ for all $x \in \mathbb{S}$ (if $x \in B$ we have equality, and if $x \notin B$ we have $V(x) \geq \mathbb{E}_x[V(Y_1)] = \mathbb{E}_x[V(\widehat{Y}_1)]$). By induction, we get $V(x) \geq \mathbb{E}_x[V(\widehat{Y}_k)]$ for all k (expand as a telescoping sum and apply the Markov property). But then for all k, and all x,

$$\begin{aligned} V(x) &\geq \mathbb{E}_x[V(\widehat{Y}_k)] \\ &= \sum_{y\in\mathbb{S}} V(y)\widehat{P}^k(x,y) \\ &\geq \sum_{\substack{y\notin B \\ V(y)\geq M}} V(y)\widehat{P}^k(x,y) \\ &\geq M\mathbb{P}_x(\widehat{Y}_k \notin B, V(\widehat{Y}_k) \geq M), \end{aligned} \tag{6.2}$$

where $\widehat{P}$ is the transition matrix for $\widehat{Y}$. We claim $\mathbb{P}_{x_0}(\widehat{Y}_k \notin B, V(\widehat{Y}_k) \geq M) \to \mathbb{P}_{x_0}(T_B = \infty)$ as $k \to \infty$. Indeed, we can decompose

$$\mathbb{P}_{x_0}(\widehat{Y}_k \notin B) = \mathbb{P}_{x_0}(\widehat{Y}_k \notin B, V(\widehat{Y}_k) \geq M) + \mathbb{P}_{x_0}(\widehat{Y}_k \notin B, V(\widehat{Y}_k) < M),$$

and it is clear that $\mathbb{P}_{x_0}(\widehat{Y}_k \notin B) = \mathbb{P}_{x_0}(Y_0, \cdots, Y_k \notin B) \to \mathbb{P}_{x_0}(T_B = \infty)$ by continuity of probability. And

$$\mathbb{P}_{x_0}(\widehat{Y}_k \notin B, V(\widehat{Y}_k) < M) = \mathbb{P}_{x_0}(\widehat{Y}_k \notin B, V(Y_k) < M) \leq \mathbb{P}_{x_0}(V(Y_k) < M),$$

which goes to 0 as $y \to \infty$ since x_0 is transient by assumption and $\{y : V(y) < M\}$ is

finite by assumption on V. It follows that

$$\mathbb{P}_{x_0}(\widehat{Y}_k \notin B, V(\widehat{Y}_k) \geq M) \to \mathbb{P}_{x_0}(T_B = \infty)$$

as $k \to \infty$, as claimed. But then taking $k \to \infty$ in equation (6.2) yields

$$V(x) \geq M\mathbb{P}_{x_0}(T_B = \infty),$$

which contradicts the definition of M. Thus our assumption was wrong, and Y (and hence X) is recurrent. □

6.2 Other Technical Lemmata

The following was used in the proof of 6.1.3, and is well-known.

Lemma 6.2.1 (Dynkin's Formula)**:** Suppose X is a Markov chain with finite state space $\mathbb{S}$, and let $\mathcal{L}$ be the generator of X. Then for any a.s. bounded stopping time τ and any $x \in \mathbb{S}$, we have

$$\mathbb{E}_x[f(X_\tau)] = f(x) + \mathbb{E}_x\left[\int_0^\tau \mathcal{L}f(X_s)\mathrm{d}s\right]$$

The version of Dynkin's formula stated above was the one used in [2]. We do not provide a proof because this thesis requires a more general version of Dynkin's formula, proven below.

Lemma 6.2.2 (Dynkin's Formula)**:** Suppose $f : \mathbb{S} \times [0, \infty) \to \mathbb{R}$ has finite support in the first argument uniformly in the second argument; that is, suppose there exists $F \subset \mathbb{S}$ finite with the property that for all $t \in [0, \infty)$ and all $y \notin F$ we have $f(y, t) = 0$. Suppose moreover that f is bounded uniformly in both variables. Then f is in the domain of the generator, and for any a.s. bounded stopping time τ,

$$\mathbb{E}_x[f(X_\tau, \tau)] = f(x, 0) + \mathbb{E}_x\left[\int_0^\tau \mathcal{L}f(X_s, s)\mathrm{d}s\right].$$

Proof. Since f has finite support, the sum in the definition of $\mathcal{L}f$ (Definition 1.0.1) is a finite sum, and in particular is absolutely convergent. So f is in the domain of the generator.

Let $M_t = f(X_t, t) - \int_0^t \mathcal{L}f(X_s, s)\mathrm{d}s$. We claim that (M_t) is a martingale. Toward this end, fix $0 \leq r < t$ and notice that (if $\mathcal{F}_r$ denotes the σ-algebra of information available up to time r),

$$
\begin{aligned}
&\mathbb{E}\left[\int_0^t \mathcal{L}f(X_s, s)\mathrm{d}s - \int_0^r \mathcal{L}f(X_s, s)\mathrm{d}s \middle| \mathcal{F}_r\right] \\
&= \mathbb{E}\left[\int_r^t \mathcal{L}f(X_s, s)\mathrm{d}s \middle| \mathcal{F}_r\right] \\
&= \mathbb{E}\left[\int_r^t \left(\sum_{y\in\mathbb{S}} f(y, s)q_{X_s y}\right) + \partial_2 f(X_s, s)\mathrm{d}s \middle| \mathcal{F}_r\right] \\
&= \mathbb{E}_{X_r}\left[\int_0^{t-r} \left(\sum_{y\in\mathbb{S}} f(y, r+s)q_{X_s y}\right) + \partial_2 f(X_s, r+s)\mathrm{d}s\right] && \text{(Markov property)} \\
&= \sum_{z\in\mathbb{S}} \int_0^{t-r} \sum_{y\in\mathbb{S}} f(y, r+s)q_{zy}p_{X_r z}(s)\mathrm{d}s + \sum_{y\in\mathbb{S}} \int_0^{t-r} \partial_2 f(y, r+s)p_{X_r y}(s)\mathrm{d}s \\
&= \sum_{y\in\mathbb{S}}\sum_{z\in\mathbb{S}} \int_0^{t-r} f(y, r+s)q_{zy}p_{X_r z}(s)\mathrm{d}s + \sum_{y\in\mathbb{S}} \int_0^{t-r} \partial_2 f(y, r+s)p_{X_r y}(s)\mathrm{d}s && \text{(See below)} \\
&= \sum_{y\in\mathbb{S}} \int_0^{t-r} \sum_{z\in\mathbb{S}} f(y, r+s)q_{zy}p_{X_r z}(s)\mathrm{d}s + \sum_{y\in\mathbb{S}} \int_0^{t-r} \partial_2 f(y, r+s)p_{X_r y}(s)\mathrm{d}s && \text{(See below)} \\
&= \sum_{y\in\mathbb{S}} \int_0^{t-r} f(y, r+s)p'_{X_r y}(s) + \partial_2 f(y, r+s)p_{X_r y}(s)\mathrm{d}s && \text{(Kolmogorov forward eq.)} \\
&= \sum_{y\in\mathbb{S}} \int_0^{t-r} \frac{\mathrm{d}}{\mathrm{d}s}\Big(f(y, r+s)p_{X_r y}(s)\Big)\mathrm{d}s \\
&= \sum_{y\in\mathbb{S}} \Big(f(y, t)p_{X_r y}(t-r) - f(y, r)p_{X_r y}(0)\Big) \\
&= \mathbb{E}_{X_r}\left[f(X_{t-r}, t) - f(X_0, r)\right] \\
&= \mathbb{E}\left[f(X_t, t) - f(X_r, r)|\mathcal{F}_r\right]. && \text{(Markov property)}
\end{aligned}
$$

The only steps left to justify are the ones marked "See below" where we interchange various sums and integrals. Note that since $y \mapsto f(y, \cdot)$ has finite support uniformly in

the second argument, the sums over y are finite. This justifies the first "See below" step. By the Fubini–Tonelli theorems, to justify the second such step it suffices to check that for each y we have

$$\int_0^{t-r} \sum_{z\in\mathbb{S}} |f(y, r+s) q_{zy} p_{X_r z}(s)| \, \mathrm{d}s < \infty.$$

But f is uniformly bounded and $0 \leq p_{X_r z}(s) \leq 1$, so it suffices to bound

$$\begin{aligned}
\int_0^{t-r} \sum_{z\in\mathbb{S}} |q_{zy}| \, p_{X_r z}(s) \mathrm{d}s &= \int_0^{t-r} -q_{yy} p_{X_r z}(s) + \sum_{z\neq y} q_{zy} p_{X_r z}(s) \mathrm{d}s \\
&= \int_0^{t-r} -2q_{yy} p_{X_r z}(s) + \sum_{z\in\mathbb{S}} q_{zy} p_{X_r z}(s) \mathrm{d}s \\
&\leq \int_0^{t-r} -2q_{yy} + \sum_{z\in\mathbb{S}} q_{zy} p_{X_r z}(s) \mathrm{d}s \\
&= \int_0^{t-r} -2q_{yy} + p'_{X_r y}(s) \mathrm{d}s \\
&= -2q_{yy}(t-r) + p_{X_r y}(t-r) - p_{X_r y}(0).
\end{aligned}$$

This last quantity is finite, so interchanging the sums and integral is justified, and we conclude that

$$\mathbb{E}\left[\int_0^t \mathcal{L} f(X_s, s) \mathrm{d}s - \int_0^r \mathcal{L} f(X_s, s) \mathrm{d}s \middle| \mathcal{F}_r \right] = \mathbb{E}\left[f(X_t, t) - f(X_r, r) | \mathcal{F}_r \right].$$

Rearranging, we see that $\mathbb{E}[M_t - M_r | \mathcal{F}_r] = 0$, so (M_t) is indeed a martingale. To get the desired result, now apply some form of the optional stopping theorem (e.g., Theorem 2.13 in chapter 2 of [18]) to (M_t) with stopping time τ. □

In the proof of Theorem 6.1.1, in addition to the stronger version of Dynkin's formula given above we also required Doob's Supermartingale inequality. We now prove the latter by way of an intermediate result:

Lemma 6.2.3: Let $(X_t)_{t\in[0,T]}$ be a right-continuous submartingale, and let $\lambda > 0$ a real constant. Then

$$\lambda \mathbb{P}\left[\inf_{t\in[0,T]} X_t \leq -\lambda \right] \leq \mathbb{E}(X_T^+) - \mathbb{E}[X_0].$$

where $X_T^+ = \max\{X_T, 0\}$.

Proof. See (for example) Theorem 3.8 (ii) in Chapter 1 of [33]. □

The following result is sometimes also called "Kolmogorov's Inequality", for example in the proof of Theorem 2.1 of [28].

Lemma 6.2.4 (Doob's Supermartingale inequality)**:** Let $(X_t)_{t\in[0,\infty)}$ be a right-continuous $[0,\infty)$-valued supermartingale, and let $\lambda > 0$ a real constant. Then

$$\mathbb{P}\left[\sup_{t\in[0,\infty)} X_t \geq \lambda\right] \leq \lambda^{-1}\mathbb{E}[X_0].$$

Proof. Letting $Y_t = -X_t$, and applying Lemma 6.2.3 to Y_t yields, for each fixed $T \in (0,\infty)$,

$$\mathbb{P}\left[\sup_{t\in[0,T]} X_t \geq \lambda\right] \leq \lambda^{-1}\mathbb{E}[X_0].$$

Sending $T \to \infty$ and applying continuity of probability gives the desired result. □

Bibliography

[1] Lorenzo Duso and Christoph Zechner. "Stochastic reaction networks in dynamic compartment populations". In: *Proceedings of the National Academy of Sciences* 117.37 (2020), pp. 22674–22683. ISSN: 0027-8424. DOI: 10.1073/pnas.2003734117. eprint: https://www.pnas.org/content/117/37/22674.full.pdf. URL: https://www.pnas.org/content/117/37/22674.

[2] David F. Anderson and Aidan S. Howells. "Stochastic Reaction Networks Within Interacting Compartments". In: *Bulletin of Mathematical Biology* 85 (10 2023). ISSN: 1522-9602. DOI: 10.1007/s11538-023-01185-w. URL: https://doi.org/10.1007/s11538-023-01185-w.

[3] Daniel T. Gillespie. "A general method for numerically simulating the stochastic time evolution of coupled chemical reactions". In: *J. Comput. Phys.* 22 (1976), pp. 403–434.

[4] Daniel T. Gillespie. "Exact Stochastic Simulation of Coupled Chemical Reactions". In: *J. Phys. Chem.* 81.25 (1977), pp. 2340–2361.

[5] Michael A. Gibson and Jehoshua Bruck. "Efficient exact stochastic simulation of chemical systems with many species and many channels". In: *J. Phys. Chem. A* 105 (2000), pp. 1876–1889.

[6] David F. Anderson. "A modified Next Reaction Method for simulating chemical systems with time dependent propensities and delays". In: *J. Chem. Phys.* 127.21 (2007), p. 214107.

[7] David F. Anderson and Thomas G. Kurtz. *Stochastic Analysis of Biochemical Systems*. Mathematical Biosciences Institute Lecture Series. Springer International Publishing, 2015. ISBN: 9783319168951. URL: http://www.springer.com/us/book/9783319168944.

[8] Samuel A. Isaacson. "A convergent reaction-diffusion master equation". In: *The Journal of Chemical Physics* 139.5 (2013), p. 054101. DOI: 10.1063/1.4816377. eprint: https://doi.org/10.1063/1.4816377. URL: https://doi.org/10.1063/1.4816377.

[9] Lea Popovic, Scott A. McKinley, and Michael C. Reed. "A Stochastic Compartmental Model for Fast Axonal Transport". In: *SIAM Journal on Applied Mathematics* 71.4 (2011), pp. 1531–1556. DOI: 10.1137/090775385. eprint: https://doi.org/10.1137/090775385. URL: https://doi.org/10.1137/090775385.

[10] Masao Doi. "Stochastic theory of diffusion-controlled reaction". In: *Journal of Physics A: Mathematical and General* 9.9 (Sept. 1976), p. 1479. DOI: 10.1088/0305-4470/9/9/009. URL: https://dx.doi.org/10.1088/0305-4470/9/9/009.

[11] Radek Erban and Hans G. Othmer. "Editorial: Special Issue on Stochastic Modelling of Reaction–Diffusion Processes in Biology". In: *Bulletin of Mathematical Biology* 76 (4 2014), pp. 761–765. ISSN: 1522-9602. DOI: 10.1007/s11538-013-9929-z. URL: https://doi.org/10.1007/s11538-013-9929-z.

[12] Ikemefuna C. Agbanusi and Samuel A. Isaacson. "A Comparison of Bimolecular Reaction Models for Stochastic Reaction–Diffusion Systems". In: *Bulletin of Mathematical Biology* 76 (4 2014). ISSN: 1522-9602. DOI: 10.1007/s11538-013-9833-6. URL: https://doi.org/10.1007/s11538-013-9833-6.

[13] Mauricio J. del Razo et al. "Chemical diffusion master equation: Formulations of reaction–diffusion processes on the molecular level". In: *Journal of Mathematical Physics* 64.1 (Jan. 2023). 013304. ISSN: 0022-2488. DOI: 10.1063/5.0129620. eprint: https://pubs.aip.org/aip/jmp/article-pdf/doi/10.1063/5.0129620/16694154/013304_1_online.pdf. URL: https://doi.org/10.1063/5.0129620.

[14] A. J. McKane and T. J. Newman. "Stochastic models in population biology and their deterministic analogs". In: *Phys. Rev. E* 70 (4 Oct. 2004), p. 041902. DOI: 10.1103/PhysRevE.70.041902. URL: https://link.aps.org/doi/10.1103/PhysRevE.70.041902.

[15] David F. Anderson, David Schnoerr, and Chaojie Yuan. "Time-dependent product-form Poisson distributions for reaction networks with higher order complexes". In: *Journal of Mathematical Biology* 80 (6 May 2020). URL: https://doi.org/10.1007/s00285-020-01485-y.

[16] Fred Brauer. "Compartmental Models in Epidemiology". In: *Mathematical Epidemiology*. Ed. by Fred Brauer, Pauline van den Driessche, and Jianhong Wu. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 19–79. ISBN: 978-3-540-78911-6. DOI: 10.1007/978-3-540-78911-6_2. URL: https://doi.org/10.1007/978-3-540-78911-6_2.

[17] James R. Norris. *Markov Chains*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1997. ISBN: 978-0-521-63396-3. DOI: 10.1017/CBO9780511810633.

[18] Stewart N. Ethier and Thomas G. Kurtz. *Markov Processes: Characterization and Convergence*. 2nd ed. John Wiley & Sons, 2005. ISBN: 978-0-471-76986-6.

[19] Guy Shinar and Martin Feinberg. "Structural sources of robustness in biochemical reaction networks". In: *Science* 327.5971 (2010), pp. 1389–1391.

[20] Saul C. Leite and Ruth J. Williams. "A constrained Langevin approximation for chemical reaction networks". In: *The Annals of Applied Probability* 29.3 (2019), pp. 1541–1608.

[21] David F. Anderson et al. "On constrained Langevin equations and (bio) chemical reaction networks". In: *Multiscale Modeling & Simulation* 17.1 (2019), pp. 1–30.

[22] Darren J. Wilkinson. *Stochastic modelling for systems biology*. CRC press, 2018.

[23] Thomas G. Kurtz. "Strong approximation theorems for density dependent Markov chains". In: *Stochastic Processes and their Applications* 6.3 (1978), pp. 223–240.

[24] Chuang Xu, Mads Christian Hansen, and Carsten Wiuf. "Full classification of dynamics for one-dimensional continuous-time Markov chains with polynomial transition rates". In: *Advances in Applied Probability* (2022), pp. 1–35. DOI: 10.1017/apr.2022.20.

[25] Richard Serfozo. *Basics of Applied Stochastic Processes*. Probability and Its Applications. Springer-Verlag Berlin Heidelberg, 2009. ISBN: 978-3-540-89331-8. URL: https://link.springer.com/book/10.1007/978-3-540-89332-5.

[26] Tobias Jahnke and Wilhelm Huisinga. "Solving the chemical master equation for monomolecular reaction systems analytically". In: *Journal of Mathematical Biology* 54 (1 2007), pp. 1–26. ISSN: 1432-1416. DOI: 10.1007/s00285-006-0034-x. URL: https://doi.org/10.1007/s00285-006-0034-x.

[27] Ankit Gupta, Corentin Briat, and Mustafa Khammash. "A Scalable Computational Framework for Establishing Long-Term Behavior of Stochastic Reaction Networks". In: *PLOS Computational Biology* 10.6 (June 2014), pp. 1–16. DOI: 10.1371/journal.pcbi.1003669. URL: https://doi.org/10.1371/journal.pcbi.1003669.

[28] Sean P. Meyn and R. L. Tweedie. "Stability of Markovian processes III: Foster–Lyapunov criteria for continuous-time processes". In: *Advances in Applied Probability* 25.3 (1993), pp. 518–548. DOI: 10.2307/1427522.

[29] Martina Bonelli. "Foster-Lyapunov criteria with stopping times and applications to Stochastic Reaction Networks". MA thesis. Politecnico di Torino, 2024. Unpublished.

[30] David F. Anderson et al. "Tier structure of strongly endotactic reaction networks". In: *Stochastic Processes and their Applications* 130.12 (2020), pp. 7218–7259. ISSN: 0304-4149. DOI: https://doi.org/10.1016/j.spa.2020.07.012. URL: https://www.sciencedirect.com/science/article/pii/S0304414920303239.

[31] O. Stramer and R.L. Tweedie. "Stability and Instability of Continuous-time Markov Processes". In: *Probability, Statistics and Optimization: A Tribute to Peter Whittle*. Ed. by F.P. Kelly. Wiley Chichester, UK, 1994. Chap. 12, pp. 173–184.

[32] Sean Meyn and Richard L. Tweedie. "Transience and recurrence". In: *Markov Chains and Stochastic Stability*. 2nd ed. Cambridge Mathematical Library. Cambridge University Press, 2009. Chap. 8, pp. 171–198. DOI: 10.1017/CBO9780511626630.011.

[33] Ioannis Karatzas and Steven E. Shreve. *Brownian Motion and Stochastic Calculus*. Corrected Second. Springer Science+Business Media New York, 2019. ISBN: 978-0-387-97655-6.