Stationary solutions of classical Markov chains and Lindblad equations

Zur Erlangung des Grades eines Doktors der Naturwissenschaften (Dr. rer. nat.)
Genehmigte Dissertation von Bernd Michael Fernengel aus Aschaffenburg
Tag der Einreichung: 11.04.2023, Tag der Prüfung: 26.04.2023

1. Gutachten: Prof. Dr. Barbara Drossel
2. Gutachten: Prof. Dr. Enno Giese
Darmstadt, Technische Universität Darmstadt

TECHNISCHE UNIVERSITÄT DARMSTADT

Fachbereich Physik
Institute for Condensed Matter Physics
Theory of Complex Systems

Stationary solutions of classical Markov chains and Lindblad equations

Genehmigte Dissertation von Bernd Michael Fernengel

Tag der Einreichung: 11.04.2023
Tag der Prüfung: 26.04.2023

Darmstadt, Technische Universität Darmstadt

Die Veröffentlichung steht unter folgender Creative Commons Lizenz:
Namensnennung 4.0 International
https://creativecommons.org/licenses/by/4.0/

Inhaltsverzeichnis

Zusammenfassung

Mastergleichungen spielen eine große Rolle in den Naturwissenschaften, da sie die Zeitentwicklung der Wahrscheinlichkeitsverteilung eines Systems beschreiben. Während sie oft als fundamental zitiert werden, werden für Lösungen entweder numerische Verfahren oder Näherungsmethoden bedient. In dieser Arbeit präsentieren wir einen analytischen Ausdruck einer stationären Lösung der Mastergleichung für endliche Systeme, welche auf der Struktur des Netzwerks besteht, welches die Übergänge des Systems beschreibt. Dabei wurde der Begriff der kleinsten absorbierenden Menge eingeführt. Diese Gleichung ist übertragbar auf Markov Ketten mit diskreter Zeit.
Im zweiten Teil dieser Arbeit berechnen wir sie stationäre Lösung der Lindblad-Gleichung, in dem wir deren Sprungdynamik als einen stückweise deterministischen Prozesses betrachten. Durch Vertauschen von Zeit- und Ensemblemittel ist es möglich, Zeitmittel einer Einzeltrajektorie zu berechnen, indem man die stationären Wahrscheinlichkeiten von klassischen Markovketten verwendet und einen klassischen Zustand durch einen zeitgemittelten quantenmechanischen Zustand ersetzt. Das Ensemblemittel ergibt sich aus den möglichen Langzeitverhalten der Trajektorien, die den kleinsten absorbierenden Mengen des quantenmechanischen Übergangsnetzwerks entsprechen.
Unsere Methode ist limitiert durch die Forderung, dass die Anzahl an quantenmechanischen Zuständen direkt nach einem Sprung für jede Trajektorie endlich ist. Am Ende dieser Arbeit diskutieren wir mögliche Verallgemeinerung zu einem abzählbar unendlichen Zustandsraum oder dem Fall, dass der quantenmechanische Zustand von einem kontinuierlichen Parameter abhängt. In beiden Fällen benötigt man eine entsprechende Verallgemeinerung für stationäre Lösungen von klassische Mastergleichungen auf abzählbar unendlich großen Systemen.

Abstract

Master equations play a crucial role in natural science, as they describe the time evolution of a probability distribution in a system. While they are often referred to as being essential, computing a solution is often avoided and people refer to numerical methods or approximation techniques. In this thesis we present an analytical expression of the stationary solution of a master equation for a finite-size system, which is based on the structure of the associated state transition network and the notion of minimal absorbing sets. This formula is also applicable to discrete-time Markov chains.
In the second part of this thesis we compute the stationary solution of the Lindblad equation by using the quantum jump unravelling. After interchanging the time average with the ensemble average, evaluating the time average of a single quantum trajectory is possible using the stationary solutions of classical discrete-time Markov chains and by replacing the classical states with time-averages quantum states. The ensemble average corresponds to the possible long-term behaviors, given by the minimal absorbing sets of a quantum state transition network.
So far our method is restricted to the case that for every quantum trajectory the number of states directly after a quantum jump is finite. At the end of this thesis, we discuss possible generalizations, either to a countable infinite state space or to states that depend on a continuous parameter. Both cases require an analogue expression for stationary solutions of classical master equations on a countable infinite state space.

1. Introduction

This thesis has two major topics: While Chapter 2 focuses on Markov chains and classical master equations, Chapter 3 is about quantum master equations and their differences to the classical case.
Markov chains (both discrete- and continuous-time versions) appear in multiple contexts throughout science, whether to describe chemical reactions [Van92; Haa78; Bre14], stochastic systems in biology [GR74; QG21; Bre14], quantum optics [Aga73], population dynamics or migration models [HD83]. More recent applications also include weather forecasting and election models [Xun21].
A Markov chain is a stochastic process, where the probability for the next state depends only on the current state [Pri13; Dou+18; Bré20]. Despite (or perhaps because of) its simple structure, it contains a rich theory, which is still subject of today's research [Hoh98; EN00; Van92; Hon12].
An equivalent way to study continuous-time Markov chains is the so-called master equation. It describes the time evolution of all probabilities of the states of the system and is determined by their transition rates. The master equation is usually formulated as an initial value problem of a linear differential equation with constant coefficients. The idea behind it is that the probability flows between the states of the system like a fluid, where its total amount is being conserved. In this picture, the links between the states can be interpreted as 'pipes' and the link strength as 'pumping rates'.
A particularly interesting question to ask is the possible long-term behavior, as this determines the state the system will eventually be in. This leads to the question of existence and uniqueness of the stationary solutions and whether they are attracting. While the existence of stationary solutions is guaranteed for a finite state space, the stationary solution need not be unique and is only attracting for continuous-time Markov chains, as discrete-time Markov chains can exhibit oscillations. For an infinite state space, stationary solutions need not exist, as the example of a symmetric random walk on $\mathbb{Z}$ shows [Bré20; Pri13; Dou+18].
In this thesis, we will focus on Markov chains with a finite state space and show in Section 2.1 that while a limiting distribution need not exist, the stationary solution of a discrete-time Markov chain can be computed by taking the time average. The number of

linearly independent stationary solutions is then given by the dimension of the eigenspace to the eigenvalue one (Equation (2.8)). For continuous-time Markov chains, all stationary solutions are attracting (compare Section 2.2), giving rise to a formula analogously to the discrete-time case (Equation (2.25)).
What this means for the state transition network is discussed in Section 2.3: For every minimal absorbing set B (see definition 6) we can construct a stationary solution $\boldsymbol{p}_\infty(\boldsymbol{p}_0 \in B)$, whose positive entries correspond to the states within the minimal absorbing set. Moreover, these vectors are linearly independent and span the linear subspace of all stationary solutions.
The differences between discrete- and continuous-time Markov chains are discussed in Section 2.4: Every transition matrix for a discrete-time Markov chain can be interpreted as the generator of a continuous-time Markov chain after subtracting the identity matrix, and for every continuous-time Markov chain there exists an embedded, discrete-time Markov chain (defined in Equation (2.47)). This makes it possible to define a 'jump unravelling' of the classical master equation (see 2.4.5).
The stationary solution itself will in general depend on the initial condition and can be separated into a probability $\mathcal{P}(B \,|\, \boldsymbol{p}_0)$ determining the amount of probability mass that is accumulated in the minimal absorbing set B (which is also the probability for a single trajectory to reach this minimal absorbing set) times the stationary distribution $\boldsymbol{p}_\infty(\boldsymbol{p}_0 \in B)$ associated to this minimal absorbing set. The full stationary solution is given by the sum over all minimal absorbing sets, resulting in a Bayes like formula

$$\boldsymbol{p}_\infty(\boldsymbol{p}_0) = \sum_{B \in \{\text{min. ab. sets}\}} \mathcal{P}(B \,|\, \boldsymbol{p}_0)\; \boldsymbol{p}_\infty(\boldsymbol{p}_0 \in B). \tag{1.1}$$

An analytical expression for both probabilities, relying only on the structure of the state transition network is derived in Section 2.5.
This result is also applicable to discrete-time Markov chains, due to the similarities discussed in 2.4.
Section 2.6 focuses on time reversible Markov chains and its connection to detailed balance and Kolmogorov's criterion, with connections to the canonical ensemble of statistical physics being discussed in 2.6.1. As a strongly connected system evolves according to the master equation, its entropy can be shown to be monotonously increasing (see Section 2.6.2), compatible with physical intuition.

Quantum master equations describe the time evolution of open quantum systems [BP02]. They are usually formulated as an initial value problem for density matrices and are -

in some sense - a generalization of classical master equations, as they describe not only the time evolution of probabilities of states (the diagonal elements of a density matrix), but also of the coherences between different states, which are the off-diagonal elements of density matrices.
An open quantum system is a quantum mechanical system in contact with an external environment. In order to derive an equation that describes the behavior of the system, we (following the arguments in [BP02]) assume that the system is markovian and that it is only weakly coupled to the environment, to arrive at the so-called Lindblad equation:

$$\begin{aligned} \partial_t \boldsymbol{\rho}(t) = \mathcal{L}(\boldsymbol{\rho}) = & \underbrace{-i[H, \boldsymbol{\rho}(t)]}_{\text{von Neumann term}} + \underbrace{\sum_{k \in I} \gamma_k \left(V_k \, \boldsymbol{\rho}(t) \, V_k^\dagger - \frac{1}{2} \left\{ V_k^\dagger V_k, \boldsymbol{\rho}(t) \right\} \right)}_{\text{dissipator term}} \\ = & -i \left(H_c \, \boldsymbol{\rho}(t) - \boldsymbol{\rho}(t) \, H_c \right) + \sum_{k \in I} \gamma_k \left(V_k \, \boldsymbol{\rho}(t) \, V_k^\dagger \right) , \\ \boldsymbol{\rho}(t=0) = & \, \boldsymbol{\rho}_0, \end{aligned} \tag{1.2}$$

where $[\cdot, \cdot]$ and $\{\cdot, \cdot\}$ denote the commutator and anticommutator, respectively, and the so-called conditional Hamiltonian H_c is defined as

$$H_c := H - \frac{i}{2} \sum_{k \in I} \gamma_k \left(V_k^\dagger \, V_k \right) =: H - \frac{i}{2} \, \Lambda, \tag{1.3}$$

$$\text{with } \hbar = 1. \tag{1.4}$$

There exist also other derivations, which are based on different assumptions, like low density or singular coupling [BP02].
The generator $\mathcal{L}$ of the Lindblad equation guarantees that the solution is trace preserving and completely positive. On the other hand, it is the only generator that has these properties. This was first shown by Gorini, Kossakowski, and Sudarshan for the finite dimensional case in 1975 [GKS76], with Lindblad extending this statement for bounded operators one year later [Lin76]. While on one hand generators of Lindblad type had already appeared in the physical literature [BP02], a similar statement for unbounded operators is still unproven [BP02].
The Lindblad equation contains the von Neumann equation, describing a closed quantum mechanical system, as a special case. The solution of the von Neumann equation is given by $\mathrm{e}^{-i\,H\,t} \, \boldsymbol{\rho}_0 \, \mathrm{e}^{i\,H\,t}$, with the self-adjoint Hamiltonian $H^\dagger = H$. Since H_c is not self-adjoint, the (conditional) time evolution operator $\mathrm{e}^{-i\,H_c\,t}$ is not unitary.

The additional terms can be interpreted as modeling the influence of the environment, in particular by inducing transitions between states of the system. The so-called Lindblad operators $\{V_k \,|\, k \in I\}$ can be interpreted as causing transitions $\boldsymbol{\rho} \to \frac{V_k \boldsymbol{\rho} V_k^\dagger}{\mathrm{Tr}[V_k \boldsymbol{\rho} V_k^\dagger]}$, where $\emptyset \neq I \subset \mathbb{N}$ is a finite index set and $\gamma_k > 0$ are positive transition rates. We will restrict ourselves to the finite-dimensional case, that is $\boldsymbol{\rho}, V_k, H \in \mathbb{C}^{N \times N}$, with $N \in \mathbb{N}, N \geq 2$.
When conducting a numerical simulation of a specific model, one often relies on unravellings of the Lindblad equation, which are ensembles of stochastic trajectories whose average yields the solution of the Lindblad equation [BP02]. While different unravellings may yield qualitative different types of dynamics, all measurable quantities depend only on the solution $\boldsymbol{\rho}(t)$, while additional properties depending on either a certain type of unravelling or a specific quantum trajectory are not accessible to experimental observations.
One important aspect in the study of the quantum master equation is its long-term behavior $\lim_{t\to\infty} \boldsymbol{\rho}(t \,|\, \boldsymbol{\rho}_0)$ and its steady states $\boldsymbol{\rho}_\infty(\boldsymbol{\rho}_0)$ (with $\boldsymbol{\rho}_\infty$ defined by $\mathcal{L}(\boldsymbol{\rho}_\infty) = 0$), as this determines the state where a quantum mechanical system will eventually be. When certain algebraic conditions on the Lindblad operators are satisfied, the steady state is unique and asymptotically stable. The most prominent theorem about such conditions was given by Spohn [Spo77]. It requires that the set of Lindblad operators form a basis, i.e. that the environment couples to all degrees of freedom, which applies only to a limited set of physical systems and cannot easily be generalized.
Most papers studying the stationary states of the Lindblad equation require additional properties of the Lindblad operators. Bua and Prosen propose a method which assumes that the underlying symmetry of the system is already known [BP12]. While being numerically cheap, finding all the symmetries of an open system can be highly non-trivial, especially when both 'strong' and 'weak' symmetries are involved (see [TM21] for a summary). Another approach was made by Trushechkina, where special kinds of Lindblad operators were considered, which allow a 'backwards' transition for every 'forward' transition (see [Tru18]). For the dimension $N = 2$ a full discussion can be found in [And+22].
The general nature of the Lindbladian in Equation (1.2) makes it difficult to compute analytical expressions for all possible steady states. In this thesis, we aim at statements about the stationary states, without making additional assumptions about the Lindblad operators. We will address the nature and number of these stationary states, how they can be determined and how they are related to the long-term behavior that results from a given initial state.
The standard way to compute the stationary state from a given initial state $\boldsymbol{\rho}_0$ would be to determine the solution $t \mapsto \boldsymbol{\rho}(t \,|\, \boldsymbol{\rho}_0)$ of the initial value problem of Equation (1.2) for all

positive times $t \geq 0$ and to evaluate the time average $\langle \boldsymbol{\rho}(t \,|\, \boldsymbol{\rho}_0) \rangle_{t \geq 0} = \lim_{T \to \infty} \frac{1}{T} \int_0^T \boldsymbol{\rho}(t \,|\, \boldsymbol{\rho}_0) \ \mathrm{d}\, t$. We propose a different method: We express the density matrix $\boldsymbol{\rho}(t)$ in terms of the average of an ensemble of stochastic quantum trajectories. Then we make use of the fact that the limit of the time average of a single quantum trajectory exists and that it may be interchanged with the ensemble limit [KM04]. We will show that both the time average and the ensemble average of a quantum trajectory can be explicitly evaluated with the theory of discrete-time Markov chains when certain conditions (in particular the number of quantum states directly after the quantum jump being finite) are satisfied.
After a short derivation of the Lindblad equation for finite dimensions (Section 3.1), we introduce the concept of unravellings of the Lindblad equation and in particular the so-called quantum-jump unravelling, which will be the focus of Chapter 3. The algorithm for the density matrix unravelling of the Lindblad equation (Section 3.3.1) introduces what is known as a stochastic quantum trajectory. An outline of our procedure to obtain a stationary solution of the Lindblad Equation (1.2) using quantum jump unravellings is given in Section 3.4. This involves computing the time average of a single quantum trajectory (Section 3.5) and the definition of the state transition network for the quantum jump unravelling in Section 3.6. After examples in Section 3.7 to demonstrate the basic idea, we recover the classical case in Section 3.7.4. After proving the existence of an analytical expression for the time-averaged state (Section 3.8), we put the building blocks together in Section 3.9 to arrive at the final formula (Equation (3.59)).

2. Markov chains and master equations for a finite state space

Markov chains are special kinds of stochastic processes that satisfy the so-called Markov property, which states that the probability for the next event depends only on the current state and not on the previous history.
Both the underlying state space as well as the time can be discrete or continuous. In the following (if not stated otherwise) we will restrict ourselves to the case of a finite state space $\Omega = \{1, \dots, |\Omega|\}$ with $|\Omega| < \infty$ and study the time average of both discrete-time, and continuous-time Markov chains.
While Markov chains as well as their stationary solutions have been excessively discussed in the existing literature [Pri13; Dou+18; Bré20], what these proofs lack are intuitive arguments and simple pictures of what 'happens' during the time evolution, that are mathematically sound at the same time. This work is meant to fill this gap. While the original ideas in [FD22] were formulated in the language of master equations, they are applicable to discrete-time Markov chains as well and are needed to determine steady states of the Lindblad equation as we will see in Chapter 3.

2.1. Discrete-time Markov chains for a finite state space

2.1.1. Defining the system

Let $(X_n)_{n\in\mathbb{N}_0}$ be a discrete-time Markov chain on Ω, where the transition probabilities are given by

$$\begin{aligned} Q_{ij} &:= q_{j\to i} := \mathcal{P}\left(X_{n+1} = i \,|\, X_n = j\right) \\ &= \mathcal{P}\left(X_{n+1} = i \,|\, X_n = j, \dots, X_0 = j_0\right) \text{ (compare [Pri13]).} \end{aligned} \tag{2.1}$$

This specifies the transition matrix $Q \in (\mathbb{R}_{\geq 0})^{|\Omega|\times|\Omega|}$, which is a column-stochastic matrix that satisfies

$$
\begin{aligned}
&Q_{ij} \geq 0 \text{ , for all } i,j \in \Omega \text{ and} \\
&\sum_{i=1}^{|\Omega|} Q_{ij} = 1 \text{ , for all } j \in \Omega.
\end{aligned}
\tag{2.2}
$$

Our system $\mathcal{S}$ is fully described by a directed, weighted graph $\mathcal{S} = (\Omega, \mathcal{E}, q)$ (which we will call a network), where the nodes are given by the set of states Ω and the edges $\mathcal{E} \subseteq \Omega \times \Omega$ (also called links) are a set of ordered pairs of states which indicate the transition between these states. The strength of a link is given by its weight function

$$
\begin{aligned}
q : \mathcal{E} &\to [0,1] \\
(j,i) &\mapsto q_{j \to i}.
\end{aligned}
\tag{2.3}
$$

When there is no transition from state j to state i, the associated transition probability vanishes, $q_{j \to i} = 0$.
To keep the notation simple, we do not distinguish between the index $n \in \{1, \ldots, |\Omega|\}$ and the state $\omega_n \in \Omega$ with index n:

$$
\begin{aligned}
\omega_n &\mathrel{\widehat{=}} n \\
\{\omega_1, \ldots, \omega_{|\Omega|}\} &\mathrel{\widehat{=}} \{1, \ldots, |\Omega|\} = \Omega.
\end{aligned}
\tag{2.4}
$$

We call $\boldsymbol{q}$ a probability vector if $\boldsymbol{q} \in (\mathbb{R}_{\geq 0})^{|\Omega|}$ and $\|\boldsymbol{q}\|_1 = 1$, that is a vector with non-negative entries that sum up to one. When $\boldsymbol{q}(n)$ is a probability vector describing the probability distribution of Ω at the time $n \in \mathbb{N}_0$, then $\boldsymbol{q}(n+1) := Q\,\boldsymbol{q}(n)$ is a probability distribution at the next time step $n+1$.

Definition 1 (Stationary solution).

A probability vector is called stationary solution of the Markov chain if its probability distribution does not change after another time step ($Q\,\boldsymbol{q} = \boldsymbol{q}$), that is, if it is an eigenvector of the transition matrix Q to the eigenvalue $\lambda = 1$.

2.1.2. Limiting behavior

The goal of this Section is to find an explicit expression for the long-term behavior of discrete-time Markov chains for a finite state space.
The limiting behavior of a Markov chain is determined by the eigenvalues of the transition matrix: When $\boldsymbol{v}$ is an eigenvector of Q to the eigenvalue $\lambda \in \mathbb{C}$ ($Q\,\boldsymbol{v} = \lambda\,\boldsymbol{v}$), we have after $n \in \mathbb{N}$ time steps:

$$Q^n \boldsymbol{v} = \lambda^n \boldsymbol{v} \xrightarrow{n\to\infty} \begin{cases} \boldsymbol{v}, \text{ if } \lambda = 1 \\ \boldsymbol{0}, \text{ if } |\lambda| < 1. \end{cases} \tag{2.5}$$

So if every transition matrix Q were diagonalizable ($Q = S\,D\,S^{-1}$ with D being diagonal with the eigenvalues on its main diagonal and the columns of S being the normalized eigenvectors) and every eigenvalue λ were either equal to one or had a modulus strictly less than one ($\lambda = 1$ or $|\lambda| < 1$), we could write any initial state as $\boldsymbol{q}_0 = \sum_{\lambda\in\sigma(Q)} \sum_{i=1}^{g_\lambda} \mu_\lambda^{(i)} \boldsymbol{v}_\lambda^{(i)}$ and compute the limiting behavior (apart from re-numbering the states) to

$$\begin{aligned} \boldsymbol{q}_\infty(\boldsymbol{q}_0) &:= \lim_{K\to\infty} \underbrace{Q^K}_{S\,D^K\,S^{-1}} \underbrace{\boldsymbol{q}_0}_{\sum_{\lambda\in\sigma(Q)} \sum_{i=1}^{g_\lambda} \mu_\lambda^{(i)} \boldsymbol{v}_\lambda^{(i)}} = \\ &= \sum_{\lambda\in\sigma(Q)} \sum_{i=1}^{g_\lambda} \mu_\lambda^{(i)}\, S \underbrace{\left(\lim_{K\to\infty} D^K\right)}_{\begin{pmatrix} \mathbb{1}_{g_{\lambda=1}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}} \underbrace{S^{-1} \boldsymbol{v}_\lambda^{(i)}}_{\boldsymbol{e}_i} = \\ &= \sum_{i=1}^{g_{\lambda=1}} \mu_\lambda^{(i)}\, \boldsymbol{v}_{\lambda=1}^{(i)}. \end{aligned} \tag{2.6}$$

Here, $\sigma(Q)$ denotes the spectrum of Q (that is the set of eigenvalues), $\boldsymbol{v}_\lambda^{(i)}$ the eigenvector of the transition matrix Q to the eigenvalue $\lambda \in \mathbb{C}$ and g_λ the geometric multiplicity of λ. However, we are faced with two problems: First, not every transition matrix is diagonalizable, as can be seen from the counter-example $Q = \frac{1}{12}\begin{pmatrix} 7 & 4 & 3 \\ 2 & 5 & 3 \\ 3 & 3 & 6 \end{pmatrix}$, which has the Jordan normal of $\begin{pmatrix} \frac{1}{4} & 1 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Second, there are transition matrices with eigenvalues possible such that $\lambda \neq 1 = |\lambda|$, as can be seen from figure (2.1). The fact that there are no eigenvalues with modulus greater than one follows from Gershgorin's circle theorem (see Section 27 and [Ger31]).

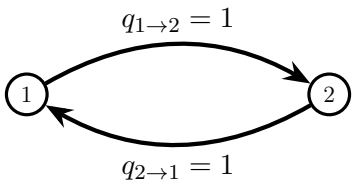


Figure 2.1.: The state transition network of a Markov chain with period equal to two: The corresponding transition matrix equals $Q = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ has the two eigenvalues $\sigma(Q) = \{-1, 1\}$. Since $Q^{2\,k+1} = Q$ and $Q^{2\,k} = \mathbb{1}_2$, the probability distribution oscillates between $\boldsymbol{q}(2\,k) = \begin{pmatrix} q_0^{(1)} \\ q_0^{(2)} \end{pmatrix}$ and $\boldsymbol{q}(2\,k+1) = \begin{pmatrix} q_0^{(2)} \\ q_0^{(1)} \end{pmatrix}$ for all natural numbers $k \in \mathbb{N}$ and the limit $\lim\limits_{k\to\infty} Q^k \boldsymbol{q}_0$ does not exist.

We will address these problems as follows: Instead of 'normal' eigenvectors, we consider a basis of generalized eigenvectors of the transition matrix, namely $\boldsymbol{h}_{\lambda,s} \in \text{kern}\,(Q - \lambda\,\mathbb{1})^s \backslash \text{kern}\,(Q - \lambda\,\mathbb{1})^{s-1}$, where s is a natural number, less or equal to the algebraic multiplicity of the corresponding eigenvalue, $s \in \{1, \ldots, a_\lambda\}$. We know from linear algebra that an ordered basis of generalized eigenvectors always exists and we can write the transition matrix as $Q = S\,J\,S^{-1}$, where J is the Jordan normal form of Q and the generalized eigenvectors are the columns of S.

Second, we focus on the time average $\lim\limits_{K\to\infty} \frac{1}{K} \sum\limits_{k=0}^{K-1} Q^k\,\boldsymbol{q}_0$ instead of the time limit $\lim\limits_{K\to\infty} Q^K\,\boldsymbol{q}_0$. In Figure 2.1, for example, the time limit does not exist, whereas the time average converges to the stationary solution of $\boldsymbol{q}_\infty = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}$.

In order to prove that the time average exists, it suffices to show that the following two conditions are satisfied:

i) For the eigenvalues $\lambda = 1$ of the transition matrix, the geometric multiplicity $g_{\lambda=1}$ coincides with the algebraic multiplicity $a_{\lambda=1}$, which means that Jordan normal

form of Q has the following form:

$$S^{-1}\,Q\,S = \begin{pmatrix} \mathbb{1}_{g_{\lambda=1}} & & & \\ & J_1 & & \\ & & \ddots & \\ & & & J_n \end{pmatrix} \tag{2.7}$$

ii) The time average for every Jordan block $J_{\lambda\neq 1}$ vanishes: $\lim\limits_{K\to\infty} \frac{1}{K} \sum\limits_{k=0}^{K-1} J_{\lambda\neq 1}^k = 0$.

Then we could replace the time limit $\lim\limits_{K\to\infty} Q^K\,\boldsymbol{q}_0$ by the time average $\lim\limits_{K\to\infty} \frac{1}{K} \sum\limits_{k=0}^{K-1} Q^k\,\boldsymbol{q}_0$ in Equation (2.6) and get:

$$\begin{aligned} \boldsymbol{q}_\infty(\boldsymbol{q}_0) &:= \lim_{K\to\infty} \frac{1}{K} \sum_{k=0}^{K-1} \underbrace{Q^k}_{S\,J^k\,S^{-1}} \underbrace{\boldsymbol{q}_0}_{\sum\limits_{\lambda\in\sigma(Q)} \sum\limits_{i=1}^{a_\lambda} \mu_\lambda^{(i)}(\boldsymbol{q}_0)\,\boldsymbol{h}_\lambda^{(i)}} = \\ &= \sum_{\lambda\in\sigma(Q)} \sum_{i=1}^{a_\lambda} \mu_\lambda^{(i)}(\boldsymbol{q}_0)\, S \underbrace{\left(\lim_{K\to\infty} \frac{1}{K} \sum_{k=0}^{K-1} J^k \right)}_{\begin{pmatrix} \mathbb{1}_{g_{\lambda=1}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}} \underbrace{S^{-1}\,\boldsymbol{h}_\lambda^{(i)}}_{\boldsymbol{e}_i} = \\ &= \sum_{i=1}^{g_{\lambda=1}} \mu_{\lambda=1}^{(i)}(\boldsymbol{q}_0)\,\boldsymbol{h}_{\lambda=1}^{(i)}. \end{aligned} \tag{2.8}$$

In the following we will prove the conditions i) and ii):

i) If $g_{\lambda=1} < a_{\lambda=1}$, there would be a generalized eigenvector of Q to the eigenvalue $\lambda = 1$ of step $s = 2$, namely

$$\boldsymbol{h}_2 := \boldsymbol{h}_{\lambda=1,s=2} \in \text{kern}\ (Q - 1\cdot\mathbb{1})^2 \backslash \text{kern}\ (Q - 1\cdot\mathbb{1}).$$

Then by definition, $\boldsymbol{h}_1 := \boldsymbol{h}_{\lambda=1,s=1} := (Q - 1\cdot\mathbb{1})\,\boldsymbol{h}_2$ is a 'normal' eigenvector of Q to the eigenvalue $\lambda = 1$ and we have $Q\,\boldsymbol{h}_2 = \boldsymbol{h}_2 + \boldsymbol{h}_1$. Then we can choose a probability vector $\boldsymbol{p}_0$, a sufficiently small number $\epsilon > 0$ and a normalization constant $\mathcal{N}$, such that $\frac{\boldsymbol{p}_0 + \epsilon\,\boldsymbol{h}_2}{\mathcal{N}}$ is also a probability vector. After $n \in \mathbb{N}$ time steps we have:

$$Q^n \left(\frac{\boldsymbol{p}_0 + \epsilon\, \boldsymbol{h}_2}{\mathcal{N}} \right) = \frac{\boldsymbol{p}_n + \epsilon\, \boldsymbol{h}_2 + \epsilon\, n\, \boldsymbol{h}_1}{\mathcal{N}}, \tag{2.9}$$

which is not bounded, in contradiction to the fact that $\frac{\boldsymbol{p}_0+\epsilon\, \boldsymbol{h}_2}{\mathcal{N}}$ is a probability vector.

Another way to show that $a_{\lambda=1} = g_{\lambda=1}$ for stochastic matrices Q, is to trace it back to the fact that $a_{\lambda=0} = g_{\lambda=0}$ for the corresponding generators $\Gamma_Q = Q - \mathbb{1}$ of the embedded continuous-time Markov chain (see Equation (2.51)), which is shown in lemma 44.

ii) For every analytical function f and any Jordan block

$$J = \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \lambda & 1 \\ & & & \lambda \end{pmatrix} \in \mathbb{C}^{D\times D}$$

of size $D \in \mathbb{N}_{\geq 2}$, we have

$$f(J_D) = \begin{pmatrix} f(\lambda) & f'(\lambda) & \dots & \frac{f^{(D-1)}(\lambda)}{(D-1)!} \\ & \ddots & \ddots & \vdots \\ & & f(\lambda) & f'(\lambda) \\ & & & f(\lambda) \end{pmatrix}. \tag{2.10}$$

When we choose a specific function

$$\begin{gathered} f(x) = \lim_{K\to\infty} \frac{1}{K} \underbrace{\sum_{k=0}^{K-1} x^k}_{\frac{1-x^K}{1-x}} = 0, \\ \text{for all } x \in \{z \in \mathbb{C} : |z| \leq 1\} \backslash \{1\}, \end{gathered} \tag{2.11}$$

we get

$$\lim_{K\to\infty} \frac{1}{K} \sum_{k=0}^{K-1} \left(J_{\lambda\neq 1}\right)^k = f(J_{\lambda\neq 1}) = \mathbf{0}^{D\times D} = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \dots & 0 \end{pmatrix} \tag{2.12}$$

Before we start looking at more explicit expressions for $\mu^{(i)}_{\lambda=1}(\boldsymbol{q}_0)$ and $\boldsymbol{h}^{(i)}_{\lambda=1}$, we first have a look at continuous-time Markov chains.

2.2. Continuous-time Markov chains and master equations for a finite state space

Given a time homogeneous, continuous-time Markov process $(X_t)_{t\geq 0}$ on the finite state space $\Omega := \{1, \dots, |\Omega|\}$, then we have for all time points $t_1 < \cdots < t_n < t_{n+1}$ and all states $j_1, \dots, j_{n-1}, i, j \in \Omega$

$$\mathcal{P}(X_{t_{n+1}} = i \,|\, X_{t_n} = j, \dots, X_{t_0} = j_0) \overset{\text{Markov}}{=\!=\!=} \mathcal{P}(X_{t_{n+1}} = i \,|\, X_{t_n} = j). \tag{2.13}$$

Time homogeneous means that the transition matrix, whose components are defined as

$$Q_{ij}(t_2, t_1) := \mathcal{P}(X_{t_2} = i \,|\, X_{t_1} = j) \text{ for } t_1 < t_2 \tag{2.14}$$

depends only on the time difference, that is

$$Q_{ij}(t_2,\, t_1) \overset{\text{time homogeneous}}{=\!=\!=} Q_{ij}(t_2 - t_1, 0) =: Q_{ij}(t_2 - t_1). \tag{2.15}$$

In addition, we assume that $t \mapsto Q_{ij}(t)$ is a differentiable function such that

$$\begin{aligned} Q_{ij}(t+\epsilon) = \mathcal{P}(X_{t+\epsilon} = i \,|\, X_t = j) = \delta_{i,j} + \underbrace{\dot{Q}_{ij}(0)}_{=:\Gamma_{ij}} \epsilon + o(\epsilon), \\ \text{for } \epsilon \to 0^+ \text{ and all } t \geq 0. \end{aligned} \tag{2.16}$$

It follows that $Q(t) := (Q_{ij}(t))_{i,j\in\Omega} \in [0,1]^{|\Omega|\times|\Omega|}$ is the time-dependent, column-stochastic transition matrix, which is completely determined by the following initial value problem:

$$\begin{aligned} \dot{Q}(t) = \underbrace{\dot{Q}(t=0)}_{\Gamma}\, Q(t) =: \Gamma\, Q(t) \text{ and} \\ Q(t=0) = \mathbb{1}_{|\Omega|}. \end{aligned} \tag{2.17}$$

A formal way to show this, is to explicitly compute the derivative

$$\begin{aligned} \frac{d}{dt}Q(t) &= \lim_{\epsilon\to 0^+} \frac{1}{\epsilon} \Big[\underbrace{Q(t+\epsilon)}_{Q(\epsilon)\,Q(t)} - Q(t) \Big] \overset{(2.13)}{=\!=\!=} \underbrace{\lim_{\epsilon\to 0^+} \frac{1}{\epsilon} \left[Q(\epsilon) - Q(0) \right]}_{\Gamma} Q(t) \\ &= \Gamma\, Q(t), \end{aligned} \tag{2.18}$$

where we used the Markov property in the second step. Another way to see this, is look at the Markov property directly, and conclude, that the only function satisfying $Q(t_1 + t_2) = Q(t_1) \cdot Q(t_2)$ (compare Equation (2.19)), is the exponential $e^{t\,A}$, with the generator A is given by the derivative $A = \frac{d}{dt} \underbrace{\mathrm{e}^{t\,A}}_{Q(t)}\Big|_{t=0} = \dot{Q}(0) = \Gamma$.

$$
\begin{aligned}
\big[Q(t_2)\,Q(t_1)\big]_{ij} &= \sum_{k\in\Omega} \underbrace{Q_{ik}(t_2)}_{\mathcal{P}(X_{t_2}=i\,|\,X_0=k)} \underbrace{Q_{kj}(t_1)}_{\mathcal{P}(X_{t_1}=k\,|\,X_0=j)} \\
&\overset{\text{Def } Q}{=\!=} \sum_{k\in\Omega} \underbrace{\mathcal{P}(X_{t_2} = i\,|\,X_0 = k)}_{\mathcal{P}(X_{t_1+t_2}=i\,|\,X_{t_1}=k)}\, \mathcal{P}(X_{t_1} = k\,|\,X_0 = j) \\
&\overset{\substack{\text{time homogeneity +}\\ \text{Markov property}}}{=\!=} \sum_{k\in\Omega} \underbrace{\Big(\mathcal{P}(X_{t_1+t_2} = i\,|\,X_{t_1} = k, X_0 = j)\Big)}_{\frac{\mathcal{P}(X_{t_1+t_2}=i, X_{t_1}=k, X_0=j)}{\mathcal{P}(X_{t_1}=k, X_0=j)}} \underbrace{\Big(\mathcal{P}(X_{t_1} = k\,|\,X_0 = j)\Big)}_{\frac{\mathcal{P}(X_{t_1}=k, X_0=j)}{\mathcal{P}(X_0=j)}} \\
&= \sum_{k\in\Omega} \underbrace{\left(\frac{\mathcal{P}(X_{t_1+t_2} = i, X_{t_1} = k, X_0 = j)}{\mathcal{P}(X_0 = j)}\right)}_{\mathcal{P}(X_{t_1+t_2}=i, X_{t_1}=k\,|\,X_0=j)} \\
&= \mathcal{P}(X_{t_1+t_2} = i\,|\,X_0 = j) = [Q(t_2 + t_1)]_{ij}.
\end{aligned}
\tag{2.19}
$$

Moreover, the following two initial value problems are equivalent:

$$
\begin{aligned}
\dot{\boldsymbol{p}}(t) &= \Gamma\,\boldsymbol{p}(t) & &\Longleftrightarrow & \dot{Q}(t) &= \Gamma\,Q(t) \\
\boldsymbol{p}(t=0) &= \boldsymbol{p}_0 & & & Q(t=0) &= \mathbb{1}_{|\Omega|},
\end{aligned}
\tag{2.20}
$$

$$
\text{with } \Gamma_{ij} = \begin{cases} \gamma_{j\to i} & \text{, for } i \neq j \\ -\sum_{k=1}^{|\Omega|} \gamma_{j\to k} & \text{, for } i = j\,. \end{cases}
$$

When the solution of the left-hand side of Equation (2.20) is given by $\boldsymbol{p}(t\,|\,\boldsymbol{p}_0)$ for every initial vector $\boldsymbol{p}_0$, we can define the transition matrix $Q(t)$ as

$$
Q(t) := \big(\boldsymbol{p}(t\,|\,\boldsymbol{p}_0 = \boldsymbol{e}_j)\big)_{j\in\Omega} = \big(\boldsymbol{p}(t\,|\,\boldsymbol{p}_0 = \boldsymbol{e}_1), \ldots, \boldsymbol{p}(t\,|\,\boldsymbol{p}_0 = \boldsymbol{e}_{|\Omega|})\big)\,,
$$

which fulfills $\dot{Q}(t) = \Gamma\,Q(t)$ and $Q(0) = \mathbb{1}_{|\Omega|}$.
Vice versa, when $Q(t)$ is given as above, then $\boldsymbol{p}(t\,|\,\boldsymbol{p}_0) := Q(t)\,\boldsymbol{p}_0$ satisfies $\dot{\boldsymbol{p}} = \Gamma\,\boldsymbol{p}$ and $\boldsymbol{p}(0) = \boldsymbol{p}_0$.

The left-hand side of Equation (2.20) is called the master equation. It is an initial value problem, describing the flow of probability in a network, where the instantaneous change of the flow of probability is the difference between in-going - and out-going probability. Thereby the right hand side of Equation (2.20) depends linearly on the transition rates, which are probabilities divided by unit time. The strength of these rates can be interpreted as capacities of pipes or pumping rates. Its component-wise notation, is of the following form:

$$\begin{aligned} \frac{d}{dt}p^{(i)}(t) &= \sum_{\substack{j=1 \\ j\neq i}}^{|\Omega|} \left(p^{(j)}(t)\,\gamma_{j\to i} - p^{(i)}(t)\,\gamma_{i\to j} \right) \\ p^{(i)}(t=0) &= p_0^{(i)} \end{aligned} \tag{2.21}$$

The matrix Γ is called the (infinitesimal) generator of the transition matrix $Q(t)$ of the network $\mathcal{S} = (\Omega, \mathcal{E}, \gamma)$ [Bré20].
Note, that from Equation (2.16) it is clear, that both $\gamma_{j\to i} := \Gamma_{ij} > 0$, for $i \neq j$ and $\sum_{i\in\Omega} \Gamma_{ij} = 0$.
While the master equation describes the probability flow for a given initial state $\boldsymbol{p}_0$, the differential equation of the solution operator $\dot{Q}(t) = \Gamma\, Q(t)$ describes 'all possible' probability flows for 'all initial states' at once.
So there are three different ways to define a time homogeneous continuous-time Markov chain on a finite state space: Via the master equation (left-hand side of Equation (2.20)), via the solution operator (right hand side of Equation (2.20)) or via first principles (see definition (2.13) and the consequence shown in Equation (2.16)). While these definitions are - in some way - equivalent, they focus on different interpretations or pictures. The first two definitions focus on what we call the ensemble interpretation, (not to be confused with the ensemble interpretation of quantum mechanics), whilst the third definition uses the interpretation as a single trajectory. The differences between these views (in particular between the ensemble picture and that of a single trajectory) will be discussed in further details in Section 2.4.6.
From this point on, the focus in this Section will be on the long-term behavior of continuous-time Markov chains via the master equation.
Similar to the discrete-time case, we have an associated state transition network, where the states correspond to the nodes and the links (indicating transitions between the states) correspond to the edges. In contrast to the discrete-time case, the weights are only required to be non-negative (in contrast to lying in the interval $[0, 1]$) and we explicitly exclude self-loops:

$$\begin{aligned} &\gamma : \mathcal{E} \to \mathbb{R}_{\geq 0} \\ &(j,i) \mapsto \gamma_{j\to i}, \text{ with } \gamma_{i\to i} = 0. \end{aligned} \tag{2.22}$$

This is due to the fact that for the continuous-time Markov chains we have transition rates instead of transition probabilities, which means that the probability for remaining in some state is always strictly positive: $\mathcal{P}(X_t = i \,|\, X_0 = i) > 0$ for all times $t \geq 0$ and all states $i \in \Omega$.
The solution is given by $\boldsymbol{p}(t \,|\, \boldsymbol{p}_0) := \mathrm{e}^{t\,\Gamma}\, \boldsymbol{p}_0$, with the initial state $\boldsymbol{p}(t=0) = \boldsymbol{p}_0 \in (\mathbb{R}_{\geq 0})^{|\Omega|}$ and the solution operator

$$\mathrm{e}^{t\,\Gamma} := \sum_{k\in\mathbb{N}_0} \frac{\Gamma^k\, t^k}{k!} = \lim_{n\to\infty} \left(\mathbb{1} + \frac{t\,\Gamma}{n}\right)^n. \tag{2.23}$$

Suppose we start with the probability vector $\boldsymbol{p}_0$ and let the system evolve with time. It can be shown (lemma 26) that the solution $\boldsymbol{p}(t \,|\, \boldsymbol{p}_0)$ is again a probability vector. So what we will see is probability 'flowing' between different states. We are interested in the following question: What does the solution $\boldsymbol{p}(t \,|\, \boldsymbol{p}_0)$ look like as $t \to \infty$ and what is its connection to the state transition network?
First, we notice the following facts about the generator Γ :

Proposition 1.

i) Since the column sum of every column of the generator matrix equals zero, $\lambda = 0$ must be an eigenvalue of Γ.

ii) From Gershgorin's circle theorem (see Section 27 and [Ger31]), it follows that every eigenvalue $\lambda \neq 0$ of the generator of the transition matrix has a real part that is strictly less than zero, $\mathrm{Re}\,[\sigma(\Gamma)\backslash\{0\}] \subset \mathbb{R}_{<0}$.

iii) The geometric multiplicity for the eigenvalue $\lambda = 0$ agrees with the algebraic multiplicity, $g_{\lambda=0} = a_{\lambda=0}$, since otherwise one could construct special initial probability vectors, whose dynamics would not be bounded.

iv) Every stationary solution of the master equation lies in the kernel of the generator of the transition matrix and vice versa.

v) The column sum of every matrix power of the generator matrix Γ equals zero, whereas the column sum of the solution operator $Q(t)$ equals one.

For a more detailed proof, see lemma 28 in the appendix.
The question is, what we can learn from these facts about the time evolution $\boldsymbol{p}(t \,|\, \boldsymbol{p}_0) = \mathrm{e}^{t\,\Gamma}\, \boldsymbol{p}_0$, especially for the limit $t \to \infty$? When writing the initial probability vector as a

linear combination of generalized eigenvectors of Γ, namely $\boldsymbol{p}_0 = \sum_{\lambda\in\sigma(\Gamma)} \sum_{i=1}^{a_\lambda} \mu_\lambda^{(i)}(\boldsymbol{p}_0)\, \boldsymbol{h}_\lambda^{(i)}$, we know that the time evolution of any generalized eigenvector $\boldsymbol{h}_{\lambda\neq 0,m}$ to an eigenvalue $\lambda \neq 0$ will tend to zero, while the eigenvectors for the eigenvalue $\lambda = 0$ are steady states:

$$\mathrm{e}^{t\,\Gamma}\, \boldsymbol{h}_{\lambda,m} = \mathrm{e}^{\lambda t} \sum_{k=1}^{m-1} \frac{t^k}{k!}\, \boldsymbol{h}_{\lambda,m-k} \xrightarrow{t\to\infty} \begin{cases} \boldsymbol{h}_{\lambda=0,m=1} = \boldsymbol{v}_{\lambda=0} & \text{, if } \lambda = 0 \\ \boldsymbol{0} & \text{, if } \lambda \neq 0. \end{cases} \tag{2.24}$$

This means that the limiting state $\boldsymbol{p}_\infty(\boldsymbol{p}_0) := \lim_{t\to\infty} \mathrm{e}^{t\,\Gamma}\, \boldsymbol{p}_0$ exists and is a linear combination of stationary states $\boldsymbol{v}_{\lambda=0} \in \mathrm{kern}\ (\Gamma)$:

$$\boldsymbol{p}_\infty(\boldsymbol{p}_0) := \lim_{t\to\infty} \mathrm{e}^{t\,\Gamma}\, \boldsymbol{p}_0 = \sum_{i=1}^{g_{\lambda=0}} \mu_{\lambda=0}^{(i)}(\boldsymbol{p}_0)\, \boldsymbol{v}_{\lambda=0}^{(i)}, \tag{2.25}$$

where $\boldsymbol{v}_{\lambda=0}^{(i)}$ are eigenvectors of the generator matrix to the eigenvalue $\lambda = 0$ and the coefficients possibly depend on the initial condition.
So far, we have studied only the generator matrix Γ, which also encodes the corresponding state transition network, with the states and possible transitions. But what does it mean in the languages of directed graphs, when probability 'flows' in a network. In order to address this question, we will need to introduce certain terminology.

2.3. Networks in the context of master equations

While equation (2.25) is the algebraic solution of the steady state of the master equation, it tells us little about relation of the eigenvectors to the states of the system or the dynamics of a trajectory. The following Section provides an intuitive picture using state transition networks (directed, weighted graphs), while an interpretation as well as an analytical expression for both the coefficient $\mu_{\lambda=0}^{(i)}(\boldsymbol{p}_0)$ and the eigenvector $\boldsymbol{v}_{\lambda=0}^{(i)}$ is given in Section 2.5.

2.3.1. Definition and preparatory considerations

Definition 2. Walks and paths
If state $b \in \Omega$ can be reached from state $a \in \Omega$, $a \neq b$ via a finite series of transitions, there is a walk from a to b, which we indicate by $a \rightsquigarrow b$, and we say that state b is reachable from state a.

Formally, this means that we call an $n+1$-tupel $\omega \in \Omega^{n+1}$ a walk in the network $\mathcal{S}$ of length $|\omega| = n \in \mathbb{N}$ if it has a positive weight γ_ω, defined as $\gamma_\omega := \prod_{k=1}^{|\omega|} \gamma_{\omega_k \to \omega_{k+1}}$, which is just the product over all the weights along that walk.
A path is a special walk, where all states (and therefore all links) are pairwise different.

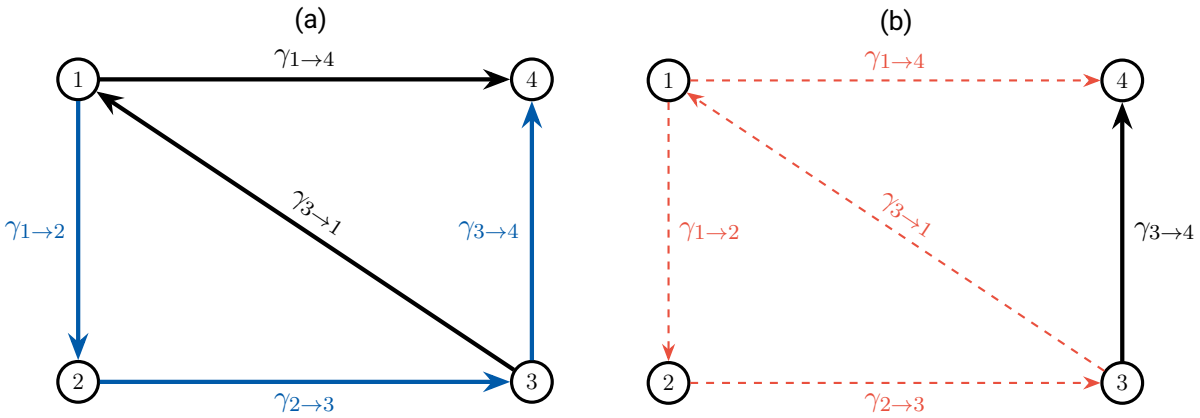


Figure 2.2.: Example and counter example of a path in a network: The blue links $w_1 = (1,2,3,4)$ in Figure 2.2a form a *path* with the corresponding weight $\gamma_{w_1} = \gamma_{1\to2}\,\gamma_{2\to3}\,\gamma_{3\to4}$, whereas the dashed red lines $w_2 = (1,2,3,1,4)$ in Figure 2.2b only form a *walk* and not a *path*, since it contains the state number 1 multiple times.

We denote with

- $\mathcal{R}(^{\to}a) := \{b \in \Omega \,|\, b \rightsquigarrow a\}$ the set of states from where a path to a exists and with
- $\mathcal{R}(a^{\to}) := \{b \in \Omega \,|\, a \rightsquigarrow b\}$ the set of states to which a path from a exists.

Definition 3 (Weak, unilateral, and strong connectedness).

We call the network $\mathcal{S} = (\Omega, \mathcal{E}, \gamma)$

i) weakly connected if the corresponding undirected graph of $\mathcal{S}$ is connected;

ii) unilaterally connected if for all a, $b \in \Omega$, b is reachable from a OR a is reachable from b;

iii) strongly connected if for all a, $b \in \Omega$, b is reachable from a and a is reachable from b, $a \rightsquigarrow b$ AND $b \rightsquigarrow a$.

Figure 2.3 illustrates the difference.

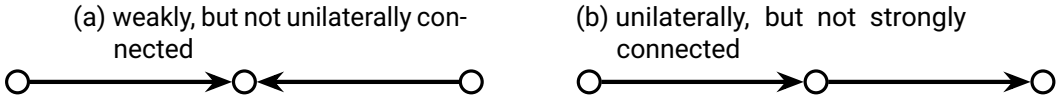


Figure 2.3.: The differences between weak, unilateral and strong connectivity.

Definition 4 (Subnetworks).

We call a network $\mathcal{S}_B = (\Omega_B, \mathcal{E}_B)$ a subnetwork of $\mathcal{S} = (\Omega, \mathcal{E})$ if it contains some of its states and all the original links between those states, that is:

$$\begin{aligned} \mathcal{S}_B \subseteq \mathcal{S} \iff & \Omega_B \subseteq \Omega \text{ and} \\ & \mathcal{E}_B := \{(i,j) \in \mathcal{E} \,:\, i,j \in \Omega_B\}. \end{aligned} \tag{2.26}$$

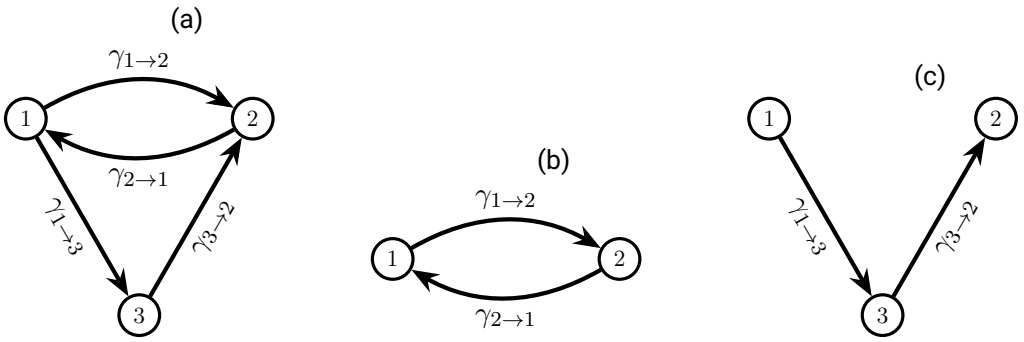


Figure 2.4.: Illustrating the concept of subnetworks: With Figure 2.4a being the original network, Figure 2.4b is a subnetwork, according to definition 4, while Figure 2.4c is not.

Definition 5 (Absorbing subsets).

We call a subnetwork $\mathcal{S}_B = (B, \mathcal{E}_B) \subseteq \mathcal{S}$ absorbing if there are no edges pointing out of B, that is if $\gamma_{b\to b^c} = 0$ for all $b \in B$ and $b^c \in B^C$.

In the context of master equations, this means that probability cannot flow out of B, which means that the probability mass focused in B can only increase:

$$\begin{aligned}
\frac{d}{dt}\left(\sum_{b\in B} p_b(t)\right) &= \sum_{b\in B} \underbrace{\dot{p}_b(t)}_{\sum_{j\in\Omega} p_j\,\gamma_{j\to b} - p_b\gamma_{b\to j}} \overset{\Omega = B\dot\cup B^C}{=\!=\!=} \\
&= \sum_{b\in B}\left(\sum_{\beta\in B} p_\beta\,\gamma_{\beta\to b} + \sum_{\alpha\in B^C} p_\alpha\,\gamma_{\alpha\to b}\right) \\
&- \sum_{b\in B}\left(\sum_{\beta\in B} p_b\,\gamma_{b\to\beta} + \sum_{\alpha\in B^C} p_b\,\underbrace{\gamma_{b\to\alpha}}_{0}\right) = \\
&= \sum_{\alpha\in B^C} p_\alpha \sum_{b\in B}\gamma_{\alpha\to b} + \underbrace{\sum_{b\in B}\sum_{\beta\in B} p_\beta\,\gamma_{\beta\to b} - \sum_{b\in B}\sum_{\beta\in B} p_b\,\gamma_{b\to\beta}}_{0} \geq 0
\end{aligned} \tag{2.27}$$

Note that the intersection of absorbing sets is again absorbing. This follows directly from the definition. Later, we will use the fact that $\mathcal{R}({}^{\to}a)^C$ and $\mathcal{R}(a^{\to})$ are absorbing subsets. The two statements are shown as follows:

- If $\mathcal{R}({}^{\to}a)^C$ were not absorbing, there would be a state $c \in \mathcal{R}({}^{\to}a)^C$ and a state $b \in \mathcal{R}({}^{\to}a)$ such that $c \rightsquigarrow b$. However, this would imply $c \rightsquigarrow b \rightsquigarrow a$, which is a contradiction.
- If $\mathcal{R}(a^{\to})$ were not absorbing, there would be a state $b \in \mathcal{R}(a^{\to})$ and a state $c \in \mathcal{R}(a^{\to})^C$ such that $b \rightsquigarrow c$. However, this would imply $a \rightsquigarrow b \rightsquigarrow c$, which is a contradiction.

So absorbing sets can be seen as 'attractors', towards which the probability mass flows to. But within such an absorbing set, there can be a sub-structure, namely again absorbing sets which are strictly smaller than the original. When repeating this process, we arrive (since the state space is finite) at absorbing sets, which are smallest, in the sense of set inclusion. This leads to the following definition.

Definition 6 (Minimal absorbing sets).

An absorbing subnetwork $\mathcal{S}_B \subseteq \mathcal{S}$ is called minimal if for all absorbing subnetworks $\mathcal{S}_C \subseteq \mathcal{S}$ with $\mathcal{S}_C \subseteq \mathcal{S}_B$, we have $\mathcal{S}_B = \mathcal{S}_C$. In particular, there can be more than one minimal absorbing subnetwork.

We call $B \subseteq \Omega$ a minimal absorbing set, whenever $(B, \mathcal{E}_B) \subseteq \mathcal{S}$ is a minimal absorbing subnetwork.
Every minimal absorbing set B is strongly connected. To see this, assume that there are two states $i, j \in B$ with $i \nrightarrow j$. Then $i \in \left(\mathcal{R}(^{\rightarrow}j)^C \cap B\right)$ and $j \notin \left(\mathcal{R}(^{\rightarrow}j)^C \cap B\right)$. This implies that $\left(\mathcal{R}(^{\rightarrow}j)^C \cap B\right)$ is an intersection of two minimal absorbing sets, which is non-empty ($i \in \left(\mathcal{R}(^{\rightarrow}j)^C \cap B\right)$) and strictly less than B ($j \notin \left(\mathcal{R}(^{\rightarrow}j)^C \cap B\right)$), in contradiction to the assumption that B is minimal.
For every state $\omega \in \Omega$ the set $\mathcal{R}(\omega^{\rightarrow})$ is absorbing, so it must contain a minimal absorbing set. This means, that for every state $\omega \in \Omega$ there is a path to some minimal absorbing set, that is there exists a minimal absorbing set $B \subseteq \Omega$ and a state $b \in B$ such that $\omega \rightsquigarrow b$.

Definition 7 (Directed graph associated to a matrix and adjacency matrix).

For any complex matrix $A \in \mathbb{C}^{N \times N}$ with $N \in \mathbb{N}_{\geq 2}$ the associated directed graph $\mathcal{S}_A$ of A, is the graph with N states, such that two states j and i are connected if and only if $A_{ij} > 0$:

$$\begin{aligned} \mathcal{S}_A = (\Omega_A, \mathcal{E}_A) \quad & \Omega_A = \{1, \ldots, N\} \\ & \mathcal{E}_A = \{(j,i) \in \{1, \ldots, N\}^2 \,:\, A_{ij} > 0\}. \end{aligned} \tag{2.28}$$

When A has only non-negative entries ($A_{ij} \geq 0$ for all $i, j \in \{1, \ldots, N\}$), then the i-j-th component of the matrix power A^n for $n \in \mathbb{N}_{\geq 1}$ is strictly positive if and only if there is a path of length n from state j to state i.
An important matrix for a directed graph is the adjacency matrix, whose i-j-th entry equals one, whenever there is a link from j to i and zero otherwise:

$$A_{ij} = \begin{cases} 1 & \text{, if } \gamma_{j \to i} > 0 \text{ and} \\ 0 & \text{, else.} \end{cases} \tag{2.29}$$

In contrast to Γ, the adjacency matrix A has zeros on the main diagonal and only tells qualitatively whether two links j and i are directly connected. It contains no quantitative information about the strength of the links.

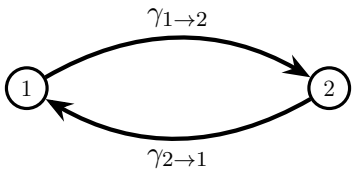


Figure 2.5.: Illustrating the difference between the adjacency matrix $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and the generator matrix $\Gamma = \begin{pmatrix} -\gamma_{1\to 2} & \gamma_{2\to 1} \\ \gamma_{1\to 2} & -\gamma_{2\to 1} \end{pmatrix}$.

Figure 2.5 shows an example of a strongly connected network. The corresponding adjacency matrix is $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and the generator is $\Gamma = \begin{pmatrix} -\gamma_{1\to 2} & \gamma_{2\to 1} \\ \gamma_{1\to 2} & -\gamma_{2\to 1} \end{pmatrix}$. Unlike the adjacency matrix A, all the column sums of the generator Γ equal zero.
The associated graph of the adjacency matrix is the graph itself. In this case, the i-j-th entry of its n-th matrix power is exactly the number of paths of length n from j to i.

Definition 8 (Direct sum of two networks).

The direct sum of two networks $\mathcal{S}_1 = (\Omega_1, \mathcal{E}_1)$ and $\mathcal{S}_2 = (\Omega_2, \mathcal{E}_2)$ is given by $\mathcal{S}_1 \oplus \mathcal{S}_2 := \{\Omega_1 \times \Omega_2, \mathcal{E}_1 \times \mathcal{E}_2\}$. The adjacency matrix of the combined network is given by the direct sum of the two adjacency matrices:

$$A_{\mathcal{S}_1 \oplus \mathcal{S}_2} = A_{\mathcal{S}_1} \oplus A_{\mathcal{S}_2} = \begin{pmatrix} A_{\mathcal{S}_1} & 0 \\ 0 & A_{\mathcal{S}_2} \end{pmatrix}. \tag{2.30}$$

Then every network is the direct sum of its weakly connected components.

Definition 9 (Diagonal dominance of column vectors).

Let $A \in \mathbb{C}^{N\times N}$ be a complex matrix.

- For $j \in \{1, \ldots, N\}$, we call the j-th column of A strictly diagonal dominant (SDD) if $|A_{jj}| > \sum_{\substack{i=1, \\ i\neq j}}^{N} |A_{ij}|$.
- If every column of a matrix is SSD, then by Gershgorin circle theorem [HW06], this matrix is non-singular.

- The definition of a weakly diagonal dominant (WDD) column of a matrix is the same as the previous one, but with a '$\geq$' sign instead of a '$>$' sign.
- We call the columns of a matrix A weakly chained diagonal dominant (WCDD) if every column is WDD and for all columns $j \in \{1, \dots, N\}$, there exists a SDD column $j_{\text{SSD}} \in \{1, \dots, N\}$ of A, an index $n \in \{1, \dots, N\}$ and a path $j = j_1 \to \cdots \to j_n = j_{\text{SSD}}$ to the state with this SSD column j_{SSD} in the directed graph associated to A (see definition 7) , that is $\prod_{k=0}^{n-1} A_{j_{k+1},j_k} \neq 0$ for some states $j_k \in \{1, \dots, N\}$ and $k \in \{1, \dots, n\}$.

It can be shown that matrices with WCDD columns (rows) are non-singular (see Section 29 in the Appendix for a proof).
The reason why we introduced definition 9, is that in the following Section 2.3.2 we prove that a certain matrix (called Γ_{B_0}) is invertible. While SDD matrices are always invertible (this follows directly from Gershgorin's circle theorem 27), WDD matices need not be invertible, as can be seen from the matrix $\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$. The condition WCDD is 'in between' the other two conditions, in the sense that it is both not as restricting an assumption as SDD, but at the same time powerful enough to guarantee invertibility.

2.3.2. Structure of the generator matrix Γ

Below, we will use the matrix Γ in a specific form that can be obtained by re-numbering the states.
Let $\mathcal{B} = \{B \subseteq \Omega \,:\, B \text{ is a minimal absorbing set } \}$ be the set of minimal absorbing sets and define $B_0 := \Omega \backslash \left(\bigcup_{B \in \mathcal{B}} B \right)$. We number the states as follows: The first $|B_0|$ states are those not contained in minimal absorbing networks. Then we count the states which lie in minimal absorbing networks block-wise, that is

$$
\begin{aligned}
B_0 &:= \{1, \dots, |B_0|\} = \Omega \backslash \left(\bigcup_{i=1}^{|\mathcal{B}|} B_i \right) \\
B_n &= \left\{ \sum_{k=0}^{n-1} |B_k| + l \,:\, l \in \{1, \dots, |B_n|\} \right\} \\
&= \left\{ \sum_{k=0}^{n-1} |B_k| + 1, \dots, \sum_{k=0}^{n-1} |B_k| + |B_n| \right\}, \text{ for } n \in \{1, \dots, |\mathcal{B}|\}.
\end{aligned}
\tag{2.31}
$$

This means that every set B_i for $i \in \{0, 1, \ldots, |\mathcal{B}|\}$ consists of successive natural numbers. After this re-numbering of the states, we can write Γ in the following form

$$\Gamma = \begin{pmatrix} \Gamma_{B_0} & \mathbf{0}^{|B_0|\times|B_1|} & \cdots & \mathbf{0}^{|B_0|\times|B_{|\mathcal{B}|}|} \\ \Gamma_{B_0\to B_1} & \Gamma_{B_1} & & \mathbf{0}^{|B_1|\times|B_{|\mathcal{B}|}|} \\ \vdots & & \ddots & \\ \Gamma_{B_0\to B_{|\mathcal{B}|}} & \mathbf{0}^{|B_{|\mathcal{B}|}|\times|B_1|} & & \Gamma_{B_{|\mathcal{B}|}} \end{pmatrix} \text{ with the matrices} \tag{2.32}$$

$$\Gamma_{B_0} \in \mathbb{R}^{|B_0|\times|B_0|},\ \Gamma_{B_i} \in \mathbb{R}^{|B_i|\times|B_i|}$$
$$\text{and } \Gamma_{B_0\to B_i} \in \mathbb{R}^{|B_i|\times|B_0|},\ \text{for } i \in \{1, \ldots, |\mathcal{B}|\}.$$

If there is no matrix Γ_{B_0} ($|B_0| = 0$) and there is only one minimal absorbing set ($|\mathcal{B}| = 1$), then Γ is called irreducible [HW06; HJ12], otherwise it is called reducible.
Figure 2.6 gives an illustrating example.

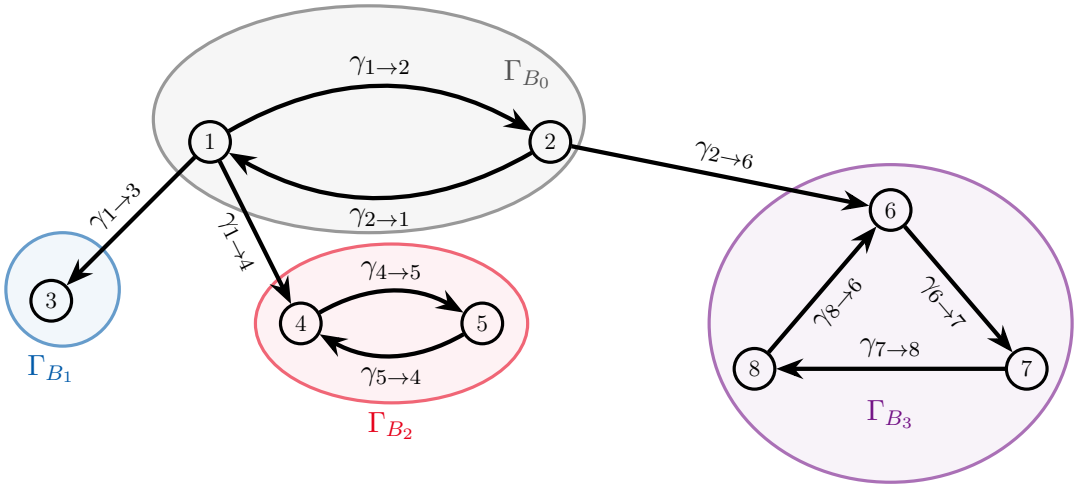


Figure 2.6.: Example of a network / directed graph with the minimal absorbing sets being $B_1 = \{3\}$, $B_2 = \{4, 5\}$ and $B_3 = \{6, 7, 8\}$.

The matrices of the absorbing subnetworks and the full matrix Γ for this example are

$$\Gamma_{B_1} = 0,\ \Gamma_{B_2} = \begin{pmatrix} -\gamma_{4\to5} & \gamma_{5\to4} \\ \gamma_{4\to5} & -\gamma_{5\to4} \end{pmatrix},\ \Gamma_{B_3} = \begin{pmatrix} -\gamma_{6\to7} & 0 & \gamma_{8\to6} \\ \gamma_{6\to7} & -\gamma_{7\to8} & 0 \\ 0 & \gamma_{7\to8} & -\gamma_{8\to6} \end{pmatrix}$$

and

$$\Gamma = \begin{pmatrix} \Gamma_{B_0} & 0 & 0 & 0 \\ \Gamma_{B_0 \to B_1} & \Gamma_{B_1} & 0 & 0 \\ \Gamma_{B_0 \to B_2} & 0 & \Gamma_{B_2} & 0 \\ \Gamma_{B_0 \to B_3} & 0 & 0 & \Gamma_{B_3} \end{pmatrix} =$$

$$= \begin{pmatrix} -\gamma_{1\to2} - \gamma_{1\to3} - \gamma_{1\to4} & \gamma_{2\to1} & 0 & 0 & 0 & 0 & 0 & 0 \\ \gamma_{1\to2} & -\gamma_{2\to1} - \gamma_{2\to6} & 0 & 0 & 0 & 0 & 0 & 0 \\ \gamma_{1\to3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \gamma_{1\to4} & 0 & 0 & -\gamma_{4\to5} & \gamma_{5\to4} & 0 & 0 & 0 \\ 0 & 0 & 0 & \gamma_{4\to5} & -\gamma_{5\to4} & 0 & 0 & 0 \\ 0 & \gamma_{2\to6} & 0 & 0 & 0 & -\gamma_{6\to7} & 0 & \gamma_{8\to6} \\ 0 & 0 & 0 & 0 & 0 & \gamma_{6\to7} & -\gamma_{7\to8} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \gamma_{7\to8} & -\gamma_{8\to6} \end{pmatrix}$$

Lemma 10 (If the matrix Γ_{B_0} exists, it is invertible).

In order to show that the matrix Γ_{B_0} is invertible, it suffices to show that Γ_{B_0} is WCDD (compare the definition 9 of WCDD and the proof of lemma 29, that WCDD matrices are indeed non-singular).
We assume that Γ is of the form (2.32) and $\mathcal{B} = \{B_1, \ldots, B_{|\mathcal{B}|}\}$ is the set of minimal absorbing sets.

1) Γ_{B_0} is WDD, since

$$\sum_{\substack{i=1,\\ i\neq j}}^{|B_0|} \underbrace{|\,(\Gamma_{B_0})_{ij}\,|}_{|\Gamma_{ij}|} \leq \sum_{\substack{i=1,\\ i\neq j}}^{|\Omega|} |\,\Gamma_{ij}\,| = |\,\Gamma_{jj}\,| = |\,(\Gamma_{B_0})_{jj}\,|.$$

2) Let $b_0 \in B_0$ be an arbitrary state in B_0. Then there are two cases:

   i) If $\gamma_{b_0 \to b} > 0$ for some $b \in \Omega \backslash B_0$, then the b_0-th row is SDD, since

$$\sum_{\substack{i=1,\\ i\neq b_0}}^{|B_0|} \underbrace{|\,(\Gamma_{B_0})_{i,\,b_0}\,|}_{|\Gamma_{i,\,b_0}|} < \sum_{\substack{i=1,\\ i\neq b_0}}^{|B_0|} |\Gamma_{i,\,b_0}| + \underbrace{|\Gamma_{b,\,b_0}|}_{\gamma_{b_0\to b}} \leq \sum_{\substack{i=1,\\ i\neq b_0}}^{|\Omega|} |\Gamma_{i,\,b_0}| = |\Gamma_{b_0,\,b_0}| = |\,(\Gamma_{B_0})_{b_0,\,b_0}\,|.$$

   ii) If the b_0-th column is not SSD ($\gamma_{b_0 \to b} = 0$ for all $b \in \Omega \backslash B_0$), we know from Section 6 that there exists a path from the state b_0 to some minimal absorbing set $B \in \mathcal{B}$, that is $b_0 \to \ldots \to b_0' \to b$ for some state $b \in \Omega \backslash B_0$ and $b_0' \in B_0$.

But this again means that if the b_0-column is not SDD, there is a path from b_0 to the SDD column b_0' (where $\gamma_{b_0' \to b} > 0$, for some $b \in \Omega \backslash B_0$).

Hence, we conclude that Γ_{B_0} is WCDD and therefore invertible.

Definition 11 (Relaxing networks).

We call a network $\mathcal{S} = (\Omega, \mathcal{E})$ relaxing if there exists a unique stationary state $\boldsymbol{p}_\infty$ such that for all initial conditions $\boldsymbol{p}_0$ the dynamics of the master equation converges to this stationary state, that is $\lim_{t\to\infty} \boldsymbol{p}(t \,|\, \boldsymbol{p}_0) = \boldsymbol{p}_\infty$. This means that the eigenspace of Γ corresponding to the eigenvalue $\lambda = 0$ is one-dimensional, $\dim(\ker(\Gamma)) = 1$.

2.3.3. A sufficient criterion for a relaxing network

In the following Section, we are looking for a sufficient criterion for a network to be relaxing. We start with the following observation:

Observation 12.

Let $\boldsymbol{v}$ be an eigenvector of Γ to a real eigenvalue $\lambda \in \mathbb{R}$, that is $\Gamma \boldsymbol{v} = \lambda \boldsymbol{v}$. Then we have

$$\begin{aligned} \mathrm{e}^{\lambda t}\|\boldsymbol{v}\|_1 = \|\mathrm{e}^{\lambda t}\boldsymbol{v}\|_1 = \|\mathrm{e}^{t\Gamma}\boldsymbol{v}\|_1 = \sum_{i=1}^{|\Omega|} \left| \sum_{j=1}^{|\Omega|} \left(\mathrm{e}^{t\Gamma}\right)_{ij} v_j \right| \leq \\ \overset{(*)}{\leq} \sum_{i,j=1}^{|\Omega|} \Bigg| \underbrace{\left(\mathrm{e}^{t\Gamma}\right)_{ij}}_{\geq 0} \Bigg| \cdot |v_j| = \sum_{j=1}^{|\Omega|} |v_j| \underbrace{\left(\sum_{i=1}^{|\Omega|} \left(\mathrm{e}^{t\Gamma}\right)_{ij} \right)}_{=1} = \|\boldsymbol{v}\|_1. \end{aligned} \tag{2.33}$$

The identity $\sum_{i=1}^{|\Omega|} \left(\mathrm{e}^{t\Gamma}\right)_{ij} = 1$ is shown in lemma 28 in the appendix.
Now, let $\boldsymbol{v} \in \ker(\Gamma)$, which means $\lambda = 0$ and we have equality in the above estimation. But on the other hand, we have equality in (*) if and only if

$$\left(\left(\mathrm{e}^{t\Gamma}\right)_{i,1} v_1, \ldots, \left(\mathrm{e}^{t\Gamma}\right)_{i,|\Omega|} v_{|\Omega|} \right) \in (\mathbb{R}_{\geq 0})^{|\Omega|} \cup (\mathbb{R}_{\leq 0})^{|\Omega|} \text{ for all } i \in \{1, \ldots, |\Omega|\}, \tag{2.34}$$

that is, when all vectors on the left-hand side of Equation (2.34) have either only non-negative or non-positive entries.

If we were able to show that $\boldsymbol{v} \in (\mathbb{R}_{\geq 0})^{|\Omega|} \cup (\mathbb{R}_{\leq 0})^{|\Omega|}$ for all $\boldsymbol{v} \in \text{kern}\,(\Gamma)$, then we could conclude that the network were relaxing, since any vector space with only non-negative entries or non-positive entries must be one-dimensional (see lemma 33 in the appendix for a formal proof).
Clearly, a sufficient criterion is, when the solution operator $\mathrm{e}^{t\,\Gamma}$ has only strictly positive entries, $(\mathrm{e}^{t\,\Gamma})_{ij} > 0$ for all $i, j \in \{1, \ldots, |\Omega|\}$.
However, this need not be the case for an arbitrary network, where the solution operator can have vanishing entries (see Figure 2.7). Having only non-negative entries is not sufficient, as the remark 32 in the appendix shows.

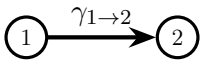


Figure 2.7.: Example of a network with generator $\Gamma = \begin{pmatrix} -\gamma_{1\to 2} & 0 \\ \gamma_{1\to 2} & 0 \end{pmatrix}$ with the solution operator, whose entries are not strictly positive: $\mathrm{e}^{t\,\Gamma} = \begin{pmatrix} \mathrm{e}^{-t\,\gamma_{1\to 2}} & 0 \\ 1 - \mathrm{e}^{-t\,\gamma_{1\to 2}} & 1 \end{pmatrix}$

This sufficient criterion is met, when the corresponding network is strongly connected, where a proof is given in lemma 30 in the appendix.
This leads to the following theorem:

Theorem 13 (Consequences of a strongly connected network).

When the network $\mathcal{S} = (\Omega, \mathcal{E})$ is strongly connected, then it is relaxing, and the stationary state $\boldsymbol{p}_\infty$ has only strictly positive entries, $\boldsymbol{p}_\infty \in (\mathbb{R}_{>0})^{|\Omega|}$. In that case, the network is called ergodic.

Proof. We know from lemma 30 that a strongly connected network implies strictly positive entries of the solution operator: $\mathrm{e}^{t\,\Gamma} \in (\mathbb{R}_{>0})^{|\Omega|}$ for all $t > 0$.
Following the arguments made in observation 12, this means that the network is relaxing ($\dim(\text{kern}\,(\Gamma - 0 \cdot \mathbb{1})) = 1$).
So let $\boldsymbol{p}_\infty \in \text{kern}\,(\Gamma) \cap (\mathbb{R}_{\geq 0})^{|\Omega|}$ be the unique steady state, with non-negative entries and $\|\boldsymbol{p}_\infty\|_1 = 1$. Then $\boldsymbol{p}_\infty$ has only strictly positive components since we have equality in (2.35) if and only if all components of $\boldsymbol{p}_\infty$ are zero:

$$p_\infty^{(i)} = \frac{\sum\limits_{\substack{j=1,\\ j\neq i}}^{|\Omega|} p_\infty^{(j)}\, \gamma_{j\to i}}{\sum\limits_{\substack{j=1,\\ j\neq i}}^{|\Omega|} \gamma_{i\to j}} \geq 0\,. \tag{2.35}$$

□

2.3.4. The relation between the dimension of the kernel of the generator matrix and the number of minimal absorbing sets

Theorem 14 ($g_{\lambda=0} = |\mathcal{B}|$).

The number of minimal absorbing sets $|\mathcal{B}|$ equals the dimension $g_{\lambda=0}$ of the kernel of the generator matrix Γ.

Proof. Let $\mathcal{B}$ be the set of minimal absorbing sets and let Γ be of the form (2.32), that is

$$\Gamma = \begin{pmatrix} \Gamma_{B_0} & \mathbf{0}^{|B_0|\times|B_1|} & \dots & \mathbf{0}^{|B_0|\times|B_{|\mathcal{B}|}|} \\ \Gamma_{B_0\to B_1} & \Gamma_{B_1} & & \mathbf{0}^{|B_1|\times|B_{|\mathcal{B}|}|} \\ \vdots & & \ddots & \\ \Gamma_{B_0\to B_{|\mathcal{B}|}} & \mathbf{0}^{|B_{|\mathcal{B}|}|\times|B_1|} & & \Gamma_{B_{|\mathcal{B}|}} \end{pmatrix} \tag{2.36}$$

- First, we show that for every minimal absorbing set, we can construct a steady state of Γ and these steady states are linearly independent: Let $\boldsymbol{p}_{B_i} \in \operatorname{kern}(\Gamma_{B_i}) \cap (\mathbb{R}_{>0})^{|B_i|}$ with $\|\boldsymbol{p}_{B_i}\|_1 = 1$ for all $i \in \{1,\dots,|\mathcal{B}|\}$. We know from Section 6 that all B_i are strongly connected and from Theorem 13 that all $\boldsymbol{p}_{B_i}$ are well defined and uniquely determined.

  Further, define

$$\boldsymbol{p}_\infty\,(\boldsymbol{p}_0 \in B_i) := \begin{pmatrix} \underbrace{0,\dots,0}_{\sum_{k=0}^{i-1}|B_k|}, \boldsymbol{p}_{B_i}, \underbrace{0,\dots,0}_{\sum_{k=i+1}^{|\mathcal{B}|}|B_k|} \end{pmatrix} \tag{2.37}$$

  and note that the set of vectors $\{\boldsymbol{p}_\infty(\boldsymbol{p}_0 \in B_i) : i \in \{1,\dots,|\mathcal{B}|\}\}$ are linearly independent (even pairwise orthogonal).

The vector $\boldsymbol{p}_\infty(\boldsymbol{p}_0 \in B)$ is indeed the stationary probability vector the system approaches, assuming that the initial probability distribution was restricted to the minimal absorbing set B, that is $\sum_{b\in B} p_0^{(b)} = 1$.

Then we have

$$\Gamma\, \boldsymbol{p}_\infty\,(\boldsymbol{p}_0 \in B_i) = \left(\underbrace{0,\ldots,0}_{\sum_{k=0}^{i-1}|B_k|}, \underbrace{\Gamma_{B_i}\,\boldsymbol{p}_{B_i}}_{0}, \underbrace{0,\ldots,0}_{\sum_{k=i+1}^{|\mathcal{B}|}|B_k|} \right) \overset{\boldsymbol{p}_{B_i}\in \text{kern }(\Gamma_{B_i})}{=\!=\!=\!=\!=\!=\!=\!=\!=\!=} \mathbf{0}^{|\Omega|}. \tag{2.38}$$

and hence span $(\{\boldsymbol{p}_\infty(\boldsymbol{p}_0 \in B_i) : i \in \{1,\ldots,|\mathcal{B}|\}\}) \subseteq \text{kern }(\Gamma)$.

- On the other hand, we show that every arbitrary element $\boldsymbol{v} \in \text{kern }(\Gamma)$ of the kernel of Γ, lies in the span of the stationary states constructed above:

  We start, by writing $\boldsymbol{v}$ as

  $$\boldsymbol{v} = (\boldsymbol{v}_0, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_{|\mathcal{B}|}) \in \mathbb{R}^{|B_0|} \times \mathbb{R}^{|B_1|} \times \cdots \times \mathbb{R}^{|B_n|} \text{ with } \boldsymbol{v}_i \in \mathbb{R}^{|B_i|}. \tag{2.39}$$

  Then we know that

  $$\mathbf{0}^{|\Omega|} \overset{\boldsymbol{v}\in \text{kern }(\Gamma)}{=\!=\!=\!=\!=\!=\!=\!=\!=\!=} \Gamma\,\boldsymbol{v} = \begin{pmatrix} \Gamma_{B_0}\,\boldsymbol{v}_0 \\ \Gamma_{B_0\to B_1}\,\boldsymbol{v}_0 + \Gamma_{B_1}\,\boldsymbol{v}_1 \\ \vdots \\ \Gamma_{B_0\to B_{|\mathcal{B}|}}\,\boldsymbol{v}_0 + \Gamma_{B_{|\mathcal{B}|}}\,\boldsymbol{v}_{|\mathcal{B}|} \end{pmatrix} \overset{\substack{\Gamma_{B_0} \text{ is invertible} \\ \Rightarrow \boldsymbol{v}_0 = 0}}{=\!=\!=\!=\!=\!=\!=\!=\!=\!=} \begin{pmatrix} \mathbf{0}^{|B_0|} \\ \Gamma_{B_1}\,\boldsymbol{v}_1 \\ \vdots \\ \Gamma_{B_{|\mathcal{B}|}}\,\boldsymbol{v}_{|\mathcal{B}|} \end{pmatrix}. \tag{2.40}$$

  The last equality holds because Γ_{B_0} is invertible, hence $\boldsymbol{v}_0 = \mathbf{0}^{|B_0|}$. Moreover, since $\dim \text{kern }(\Gamma_{B_i}) = 1$ for all $i \in \{1,\ldots,|\mathcal{B}|\}$, we conclude that

  $$\boldsymbol{v}_i \in \text{kern }(\Gamma_{B_i}) = \text{span }\big(\boldsymbol{p}_\infty(\boldsymbol{p}_0 \in B_i) : i \in \{1,\ldots,|\mathcal{B}|\}\big), \tag{2.41}$$

  that is $\boldsymbol{v}_i = \lambda_i\,\boldsymbol{p}_\infty(\boldsymbol{p}_0 \in B_i)$ for some $\lambda \in \mathbb{R}$.

  It then follows that

$$\boldsymbol{v} = \left(\underbrace{0, \ldots, 0}_{|B_0|}, \lambda_1 \boldsymbol{p}_{B_1}, \ldots, \lambda_{|\mathcal{B}|} \boldsymbol{p}_{B_{|\mathcal{B}|}} \right)^T =$$

$$= \sum_{i=1}^{|\mathcal{B}|} \lambda_i \underbrace{\left(\overbrace{0, \ldots, 0}^{\sum_{k=0}^{i-1} |B_k|}, \boldsymbol{p}_{B_i}, \overbrace{0, \ldots, 0}^{\sum_{k=i+1}^{|\mathcal{B}|} |B_k|} \right)^T}_{\boldsymbol{p}_\infty(\boldsymbol{p}_0 \in B_i)} = \sum_{i=1}^{|\mathcal{B}|} \lambda_i \, \boldsymbol{p}_\infty(\boldsymbol{p}_0 \in B_i), \text{ and hence}$$

$$\text{kern } (\Gamma) \subseteq \text{span } \Big(\{ \boldsymbol{p}_\infty(\boldsymbol{p}_0 \in B_i) \, : \, i \in \{1, \ldots, |\mathcal{B}|\} \} \Big). \tag{2.42}$$

□

This means that we can construct a basis of steady states from the set of minimal absorbing sets, with every basis vector corresponding to exactly one minimal absorbing set.
For the example given in Figure 2.6, the vectors $\boldsymbol{p}_{B_i}$ are given by

$$\boldsymbol{p}_{B_1} = 1, \, \boldsymbol{p}_{B_2} = \frac{1}{\gamma_{4\to5} + \gamma_{5\to4}} \begin{pmatrix} \gamma_{5\to4} \\ \gamma_{4\to5} \end{pmatrix},$$

$$\boldsymbol{p}_{B_3} = \frac{1}{\gamma_{7\to8}\,\gamma_{8\to6} + \gamma_{8\to6}\,\gamma_{6\to7} + \gamma_{6\to7}\,\gamma_{7\to8}} \begin{pmatrix} \gamma_{7\to8}\,\gamma_{8\to6} \\ \gamma_{8\to6}\,\gamma_{6\to7} \\ \gamma_{6\to7}\,\gamma_{7\to8} \end{pmatrix},$$

whereas the the basis vectors $\boldsymbol{p}_\infty(\boldsymbol{p}_0 \in B_i)$ are

$$\boldsymbol{p}_\infty\left(\boldsymbol{p}_0 \in B_1\right) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \boldsymbol{p}_\infty\left(\boldsymbol{p}_0 \in B_2\right) = \frac{1}{\gamma_{4\to5} + \gamma_{5\to4}} \begin{pmatrix} 0 \\ 0 \\ 0 \\ \gamma_{5\to4} \\ \gamma_{4\to5} \\ 0 \\ 0 \\ 0 \end{pmatrix} \text{ and}$$

$$\boldsymbol{p}_\infty\left(\boldsymbol{p}_0 \in B_3\right) = \frac{1}{\gamma_{7\to8}\,\gamma_{8\to6} + \gamma_{8\to6}\,\gamma_{6\to7} + \gamma_{6\to7}\,\gamma_{7\to8}} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \gamma_{7\to8}\,\gamma_{8\to6} \\ \gamma_{8\to6}\,\gamma_{6\to7} \\ \gamma_{6\to7}\,\gamma_{7\to8} \end{pmatrix}.$$

Remark 15.

- In theorem 13 we have seen that a strongly connected network is a sufficient, but not a necessary condition. With the help of theorem 14, we can say that the master equation is relaxing if and only if the corresponding state transition network has exactly one minimal absorbing set.
- In Section 2.2, we have seen what the stationary solution would look like, namely

$$\boldsymbol{p}_\infty(\boldsymbol{p}_0) := \lim_{t\to\infty} \mathrm{e}^{t\,\Gamma}\,\boldsymbol{p}_0 \overset{(2.25)}{=\!=} \sum_{i=1}^{g_{\lambda=0}} \mu^{(i)}_{\lambda=0}(\boldsymbol{p}_0)\,\boldsymbol{v}^{(i)}_{\lambda=0}, \tag{2.43}$$

  Now we now that the geometric multiplicity $g_{\lambda=0}$ of the eigenvalue $\lambda = 0$ is just the number of minimal absorbing sets $|\mathcal{B}|$, so we can replace the sum by a sum over all minimal absorbing sets: $\sum_{i=1}^{g_{\lambda=0}} \to \sum_{B\in\mathcal{B}}$. Further, we know that every eigenvector $\boldsymbol{v}_{\lambda=0}$ corresponds to exactly one minimal absorbing set B, in the sense that its strictly positive entries are associated to the states in this minimal absorbing set:

$$\boldsymbol{p}_\infty(\boldsymbol{p}_0 \in B) = \left(\underbrace{0, \ldots, 0}_{\sum_{k=0}^{i-1} |B_k|}, \boldsymbol{p}_B, \underbrace{0, \ldots, 0}_{\sum_{k=i+1}^{|\mathcal{B}|} |B_k|} \right). \tag{2.44}$$

2.4. Connection between continuous- and discrete- time Markov chains

We started in Section 2.1 with the study of discrete-time Markov chains and found that the stationary solution, given by its time average, is a linear combination of eigenvectors to the eigenvalue $\lambda = 1$, even though it need not be attracting.
In Section 2.2 we considered continuous-time Markov chain and found with Equation (2.25) a similar formula, whose connection with the structure of the state transition network was shown in Section 2.3. In this Section we want to explore both the difference of, and the connection between continuous- and discrete-time Markov chains.

2.4.1. From continuous- to discrete-time Markov chains

Lemma 16 (The waiting time distribution).

The waiting time of a state $j \in \Omega$ is the time, the system remains at that state until the next jump. It is a random variable, which is exponentially distributed (see Section A.9 in the appendix for a proof) with parameter $\gamma_{j\to} := -\Gamma_{j,j} = \sum_{k\in\Omega} \gamma_{j\to k}$, that is $\tau_{j\to} \sim \text{Exp}\,(\gamma_{j\to})$.
When we assume that we already know that the waiting time for a specific transition $j \to k$ is exponentially distributed with parameter $\gamma_{j\to k}$ (*), the proof is much simpler:

$$\begin{aligned} \mathcal{P}(t_{n+1} - t_n > t \,|\, X_{t_n} = j) = \mathcal{P}\left(\underbrace{\{\tau_{j\to k} > t\} \text{ for all } k \in \Omega}_{\bigcap_{k\in\Omega} \{\tau_{j\to k} > t\}} \right) = \\ \overset{\substack{\text{the random variables } \{\tau_{j\to k} : k\in\Omega\} \\ \text{are mutually independent}}}{=\!=\!=\!=\!=\!=\!=\!=\!=\!=\!=\!=} \prod_{k\in\Omega} \underbrace{\mathcal{P}\left(\tau_{j\to k} > t\right)}_{\mathrm{e}^{-t\,\gamma_{j\to k}}} = \\ \overset{(*)}{=\!=} \mathrm{e}^{-t \sum_{k\in\Omega} \gamma_{j\to k}} = \mathrm{e}^{-t\,\gamma_{j\to}}. \end{aligned} \tag{2.45}$$

Lemma 17 (Transition probability).

Let the system be in state $j \in \Omega$ with $\gamma_{j\to} \neq 0$. Then the probability for the next state to be $i \in \Omega\backslash\{j\}$ is given by $\frac{\gamma_{j\to i}}{\gamma_{j\to}}$

Proof.

$$\begin{aligned}
\mathcal{P}\big(X_{t_{n+1}} = i \,|\, X_{t_n} = j\big) &= \int_{\mathbb{R}_{>0}} \gamma_{j\to i}\, \mathrm{e}^{-t\,\gamma_{j\to i}}\ \underbrace{\mathcal{P}\left(\bigcap_{k\in\Omega\backslash\{i\}} \{\tau_{j\to k} > t\}\right)}_{\prod_{k\in\Omega\backslash\{i\}} \mathcal{P}(\{\tau_{j\to k}>t\})}\ \mathrm{d}t = \\
&\overset{\text{independence}}{=\!=} \int_{\mathbb{R}_{>0}} \gamma_{j\to i} \underbrace{\left(\mathrm{e}^{-t\,\gamma_{j\to i}} \prod_{k\in\Omega\backslash\{i\}} \overbrace{\mathcal{P}(\tau_{j\to k} > t)}^{\mathrm{e}^{-t\,\gamma_{j\to k}}}\right)}_{\mathrm{e}^{-t\,\gamma_{j\to}}}\ \mathrm{d}t \\
&= \int_{\mathbb{R}_{>0}} \gamma_{j\to i}\, \mathrm{e}^{-t\,\gamma_{j\to}}\ \mathrm{d}t = \frac{\gamma_{j\to i}}{\gamma_{j\to}} = \frac{\Gamma_{ij}}{-\Gamma_{jj}}.
\end{aligned} \tag{2.46}$$

□

This means that for every continuous-time Markov chain there is an embedded discrete-time Markov chain, whose transition matrix Q_Γ is given by

$$\begin{aligned}
(Q_\Gamma)_{ij} := \mathcal{P}\left(X_{t_{n+1}} = i \,|\, X_{t_n} = j\right) &= \begin{cases} \delta_{ij}, & \text{if } \Gamma_{jj} = 0 \\ \delta_{ij} + \frac{\Gamma_{ij}}{(-\Gamma_{jj})}, & \text{if } \Gamma_{jj} \neq 0 \end{cases} = \\
&= \delta_{ij} - \Gamma_{ij}\,(\Gamma_{jj})^\circ \text{ or in matrix notation} \\
Q_\Gamma &= \mathbb{1} - \Gamma\ [\mathrm{diag}\ (\Gamma)]^\circ,
\end{aligned} \tag{2.47}$$

where the symbol $\circ$ denoted the Moore-Penrose pseudoinverse. For a diagonal matrix D, it is defined as follows:

$$\begin{aligned}
(D^\circ)_{ij} &= \delta_{i,j}\,(D_{ii})^\circ \text{ with} \\
(D_{ii})^\circ &:= \begin{cases} 0, \text{ if } D_{ii} = 0 \\ \frac{1}{D_{ii}}, \text{ else} \end{cases}
\end{aligned} \tag{2.48}$$

We note that the columns of the transition matrix of the associated discrete-time Markov chain are given by

$$\left((Q_\Gamma)_{ij}\right)_{i\in\Omega} = \left((\mathbb{1} - \Gamma\left[\operatorname{diag}\left(\Gamma\right)\right]^\circ)_{ij}\right)_{i\in\Omega} = \begin{cases} \boldsymbol{e}_j & \text{, if } \gamma_{j\to}=0 \\ \left(\frac{\gamma_{j\to i}}{\gamma_{j\to}}\right)_{i\in\Omega} & \text{, else.} \end{cases} \tag{2.49}$$

An example is given in Figure 2.8

$$D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & \pm 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \Longrightarrow D^\circ = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \pm 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Figure 2.8.: Example of a matrix and the associated Moore-Penrose pseudoinverse.

The matrix Q_Γ is column-stochastic, which means we can interpret it as the associated transition matrix of the embedded discrete-time Markov chain, which we call embedded discrete-time Markov chain associated to the continuous-time Markov chain with the generator Γ (see Figure 2.9 for an illustration).
This is known as the property of competing exponentials [Bré20].

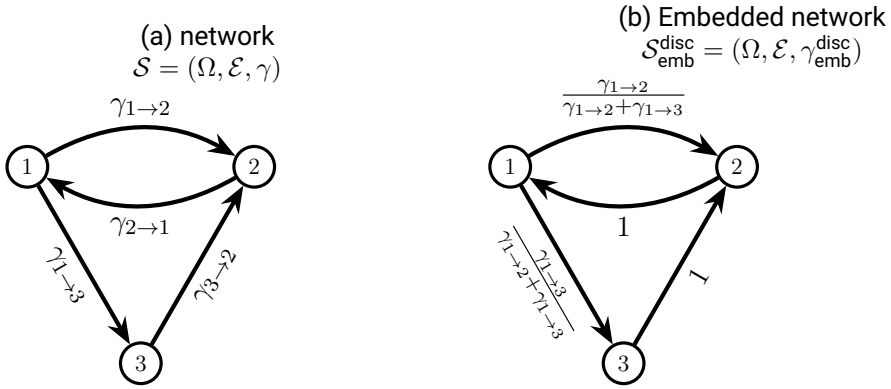


Figure 2.9.: Illustrating the difference between a continuous-time Markov chain with the generator $\Gamma = \begin{pmatrix} -\gamma_{1\to2} - \gamma_{1\to3} & \gamma_{2\to1} & 0 \\ \gamma_{1\to2} & -\gamma_{2\to1} & \gamma_{3\to2} \\ \gamma_{1\to3} & 0 & -\gamma_{3\to2} \end{pmatrix}$ with stationary solution $\boldsymbol{p}_* = \frac{1}{\mathcal{Z}_p} \begin{pmatrix} \gamma_{3\to2}\,\gamma_{2\to1} \\ (\gamma_{1\to2} + \gamma_{1\to3})\,\gamma_{3\to2} \\ \gamma_{2\to1}\,\gamma_{1\to3} \end{pmatrix}$ and the transition matrix of the associated discrete-time Markov chain: $Q_\Gamma = \begin{pmatrix} 0 & 1 & 0 \\ \frac{\gamma_{1\to2}}{\gamma_{1\to2}+\gamma_{1\to3}} & 0 & 1 \\ \frac{\gamma_{1\to3}}{\gamma_{1\to2}+\gamma_{1\to3}} & 0 & 0 \end{pmatrix}$ with stationary solution $\boldsymbol{q}_* = \frac{1}{\mathcal{Z}_q} \begin{pmatrix} 1 \\ 1 \\ \frac{\gamma_{1\to3}}{\gamma_{1\to2}+\gamma_{1\to3}} \end{pmatrix}$, where $\mathcal{Z}_p$ and $\mathcal{Z}_q$ are normalization factors.

We further note, that for the minimal absorbing set $B \in \mathcal{B}$ the sequence $\left(\sum_{b\in B} (Q_\Gamma)^n_{b,j} \right)_{n\in\mathbb{N}}$ is convergent, since it is both bounded and monotonously increasing:

$$\begin{aligned} \sum_{b\in B} \underbrace{\left(Q_\Gamma^{n+1}\right)_{b,j}}_{\sum_{k\in\Omega} (Q_\Gamma)_{b,k} \left(Q_\Gamma^n\right)_{k,j}} &\overset{B\subseteq\Omega}{\geq} \sum_{k\in B} \left(Q_\Gamma^n\right)_{k,j} \underbrace{\sum_{b\in B} (Q_\Gamma)_{b,k}}_{1} \\ &\overset{B \text{ is absorbing}}{=\!=\!=} \sum_{k\in B} \left(Q_\Gamma^n\right)_{k,j} . \end{aligned} \tag{2.50}$$

This fact will be relevant in Equation (2.80) in Section 2.5.2, where we look for an analytical expression for the probability of a single trajectory to be captured by a minimal absorbing set.

2.4.2. From discrete- to continuous-time Markov chains

Given a discrete-time Markov chain $\mathcal{S} = (\Omega, \mathcal{E}, q)$ with the corresponding transition matrix Q, we can define a continuous-time Markov chain $\mathcal{S}^{\text{cont}} = (\Omega, \mathcal{E}_q^{\text{cont}}, \gamma_q^{\text{cont}})$ with the generator

$$\Gamma_Q := Q - \mathbb{1}_{|\Omega|}. \tag{2.51}$$

Note, that in contrast to discrete-time Markov chain, state transition networks of continuous-time Markov chains must not have self-loops.
Figure 2.10 demonstrates this procedure:

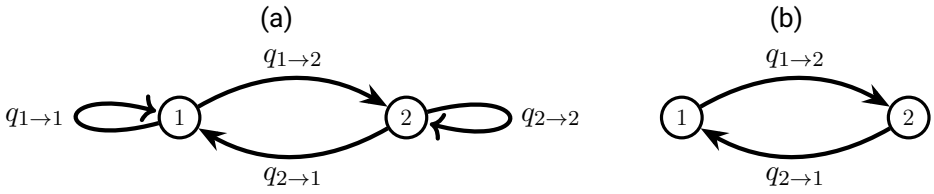


Figure 2.10.: Illustration of a discrete-time Markov chain, given by the transition matrix $Q = \begin{pmatrix} 1 - q_{1\to2} & q_{2\to1} \\ q_{1\to2} & 1 - q_{2\to1} \end{pmatrix}$ and the associated continuous-time Markov chain given by its generator $\Gamma_Q = \begin{pmatrix} -q_{1\to2} & q_{2\to1} \\ q_{1\to2} & -q_{2\to1} \end{pmatrix}$.

Equation (2.51) also implies that every generalized eigenvector $\boldsymbol{h}_{\lambda,s}(Q)$ of the eigenvalue $\lambda \in \sigma(Q)$ of the transition matrix Q is at the same time a generalized eigenvector $\boldsymbol{h}_{\lambda-1,s}(\Gamma_Q)$ to the eigenvalue $\lambda - 1 \in \sigma(\Gamma_Q)$ of the generator Γ:

$$\begin{aligned} \boldsymbol{h}_{\lambda,s}(Q) \in \text{kern } (Q - \lambda\mathbb{1})^s \backslash \text{kern } (Q - \lambda\mathbb{1})^{s-1} \mathrel{\widehat{=}} \\ \boldsymbol{h}_{\lambda-1,s}(\Gamma_Q) \in \text{kern } (\Gamma_Q - (\lambda-1)\mathbb{1})^s \backslash \text{kern } (\Gamma_Q - (\lambda-1)\mathbb{1})^{s-1}. \end{aligned} \tag{2.52}$$

When interested in the long-term behavior $\boldsymbol{q}_\infty$ of a discrete-time Markov chain with initial distribution $\boldsymbol{q}_0$ and transition matrix Q, one can instead compute the long-term

behavior of the corresponding continuous-time Markov chain with the generator matrix $\Gamma_Q = Q - \mathbb{1}_{|\Omega|}$: $\boldsymbol{q}_\infty(\boldsymbol{q}_0, Q) = \boldsymbol{p}_\infty(\boldsymbol{q}_0, Q - \mathbb{1}_{|\Omega|})$.

2.4.3. 'Switching' between continuous - and discrete-time Markov chains

We have seen how to switch from continuous-time Markov chains to discrete-time Markov chains ($\Gamma \to Q$) and vice versa ($Q \to \Gamma$). We now look at what happens if one keeps repeating this process:
Starting with the generator of a continuous-time Markov chain Γ, then $\Gamma_{Q_\Gamma} \neq \Gamma$ is possible, but $Q_\Gamma = Q_{\Gamma_{Q_\Gamma}}$:

$$
\begin{aligned}
Q_\Gamma &\overset{\text{Def}}{=\!=} \mathbb{1} - \Gamma\ [\text{diag}\ (\Gamma)]^\circ \\
\Gamma_{Q_\Gamma} &\overset{\text{Def}}{=\!=} Q_\Gamma - \mathbb{1} = -\Gamma\ [\text{diag}\ (\Gamma)]^\circ \\
Q_{\Gamma_{Q_\Gamma}} &\overset{\text{Def}}{=\!=} \mathbb{1} - \underbrace{\Gamma_{Q_\Gamma}}_{-\Gamma\,[\text{diag}\ (\Gamma)]^\circ} \left[\text{diag}\ \left(\underbrace{\Gamma_{Q_\Gamma}}_{-\Gamma\,[\text{diag}\ (\Gamma)]^\circ}\right)\right]^\circ = \\
&= \mathbb{1} - \Gamma\ [\text{diag}\ (\Gamma)]^\circ \left[\underbrace{\text{diag}\ \left(\Gamma\ [\text{diag}\ (\Gamma)]^\circ\right)}_{\text{diag}\ (\Gamma)\,[\text{diag}\ (\Gamma)]^\circ}\right]^\circ = \\
&\overset{(2.54a)}{=\!=} \mathbb{1} - \Gamma\ [\text{diag}\ (\Gamma)]^\circ\ \underbrace{\left[\text{diag}\ (\Gamma)\ [\text{diag}\ (\Gamma)]^\circ\right]^\circ}_{\text{diag}\ (\Gamma)\,[\text{diag}\ (\Gamma)]^\circ} = \\
&\overset{(2.54b)}{=\!=} \mathbb{1} - \Gamma\ \underbrace{[\text{diag}\ (\Gamma)]^\circ\ \text{diag}\ (\Gamma)\ [\text{diag}\ (\Gamma)]^\circ}_{[\text{diag}\ (\Gamma)]^\circ} = \\
&\overset{(2.54c)}{=\!=} \mathbb{1} - \Gamma\ [\text{diag}\ (\Gamma)]^\circ = Q_\Gamma,
\end{aligned}
\tag{2.53}
$$

where we used the fact that for a diagonal matrix D we have:

$$
\text{diag}\ (A\,D) = \text{diag}\ (A)\,\text{diag}\ (D)\ \text{and} \tag{2.54a}
$$

$$
(D\,D^\circ)^\circ = (D\,D^\circ) \tag{2.54b}
$$

$$
D^\circ\,D\,D^\circ = D^\circ. \tag{2.54c}
$$

Equation (2.53) immediately gives us $Q_{\Gamma_{Q_{\Gamma_Q}}} = Q_{\Gamma_Q}$. When depicting a transition from a continuous-time Markov chain to a discrete-time Markov chain via a dashed arrow, we

have the following diagram:

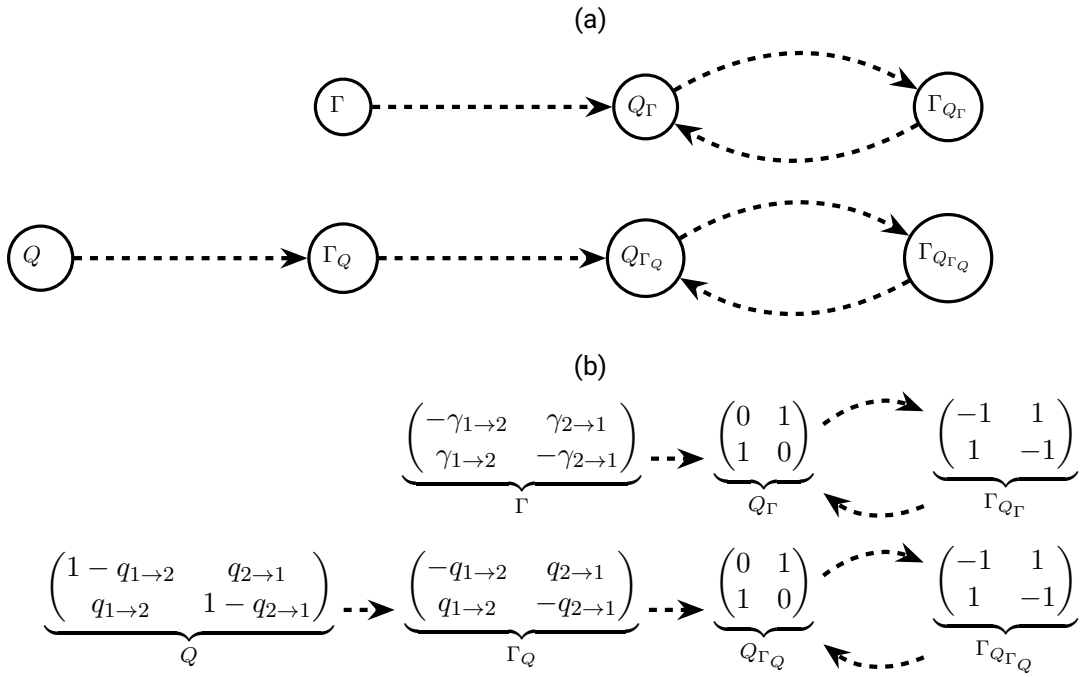


Figure 2.11.: Illustration of the transitions from a (continuous/discrete)-time Markov chain towards a (discrete/continuous)-time Markov chain and vice versa. While Figure 2.11a shows the general transformation of transition matrices, Figure 2.11b shows that it is indeed possible to have $\Gamma_{Q_\Gamma} \neq \Gamma$, $Q_{\Gamma_Q} \neq Q$ and $\Gamma_{Q_{\Gamma_Q}} \neq \Gamma_Q$.

2.4.4. The connection between stationary solutions of continuous- and embedded discrete-time Markov chains

Now that we have seen how continuous- and discrete-time Markov chains are connected, we have a look at the relation between their stationary solutions.
Let $\Gamma := \Gamma_B$ be the generator of a minimal absorbing set $B \subseteq \Omega$ of a continuous-time Markov chain, $\boldsymbol{p}_\infty$ its stationary solution ($\Gamma\,\boldsymbol{p}_\infty = \mathbf{0}$) and $Q_\Gamma = \mathbb{1} - [\mathrm{diag}\,(\Gamma)]^\circ\,\Gamma$ the transition matrix of the embedded discrete-time Markov chain, corresponding to the generator Γ. Then

$$\boldsymbol{q}_\infty := \begin{cases} 1 & \text{, if } |B| = 1 \\ \frac{-\operatorname{diag}(\Gamma)\,\boldsymbol{p}_\infty}{\|\operatorname{diag}(\Gamma)\,\boldsymbol{p}_\infty\|_1} & \text{, if } |B| \geq 2 \end{cases} \tag{2.55}$$

is the stationary solution of the embedded discrete-time Markov chain Q_Γ (note that $\big(-(\operatorname{diag}(\Gamma)\,\boldsymbol{p}_\infty)_i\big)_{i\in\Omega} = \big(\underbrace{-\Gamma_{ii}}_{\geq 0}\,\underbrace{p_\infty^{(i)}}_{>0}\big)_{i\in\Omega} = \mathbf{0} \iff \Gamma_{ii} = 0$ for all $i \in \Omega$, so this is well defined):

$$\begin{aligned} Q_\Gamma\,\boldsymbol{q}_\infty &= \big(\mathbb{1} - \Gamma\,[\operatorname{diag}(\Gamma)]^\circ\big)\,\underbrace{[\operatorname{diag}(\Gamma)]\,\boldsymbol{p}_\infty}_{\boldsymbol{q}_\infty} = \boldsymbol{q}_\infty + \underbrace{\Gamma\,[\operatorname{diag}(\Gamma)]^\circ \operatorname{diag}(\Gamma)\,\boldsymbol{p}_\infty}_{\Gamma\,\boldsymbol{p}_\infty = \mathbf{0}} \\ &= \boldsymbol{q}_\infty, \end{aligned} \tag{2.56}$$

where we used the fact that

$$\begin{aligned} (\Gamma\,[\operatorname{diag}(\Gamma)]^\circ \operatorname{diag}(\Gamma)\,\boldsymbol{p}_\infty)_i &= \sum_{j\in\Omega} \Gamma_{ij}\,\underbrace{(\Gamma_{jj})^\circ\,(\Gamma_{jj})}_{1-\delta_{0,\Gamma_{jj}}}\,p_\infty(j) = \\ &= \underbrace{\left(\sum_{j\in\Omega} \Gamma_{ij}\,p_\infty(j)\right)}_{(\Gamma\,\boldsymbol{p}_\infty)_i = 0} + \sum_{\substack{j\in\Omega \\ \Gamma_{jj}=0}} \Gamma_{ij}\,p_\infty(j) \\ &\overset{\substack{\Gamma_{jj}=0 \Rightarrow \\ \Gamma_{ij}=0\,\forall\, i\in\Omega}}{=\!=\!=\!=\!=\!=\!=\!=} 0. \end{aligned} \tag{2.57}$$

When on the other hand, the stationary solution of the embedded, discrete-time Markov chain Q_Γ is given by $\boldsymbol{q}_\infty$ (that is $Q_\Gamma\,\boldsymbol{q}_\infty = \boldsymbol{q}_\infty$), then

$$\boldsymbol{p}_\infty := \begin{cases} 1 & \text{, if } |B| = 1 \\ \frac{-[\operatorname{diag}(\Gamma)]^\circ\,\boldsymbol{q}_\infty}{\|[\operatorname{diag}(\Gamma)]^\circ\,\boldsymbol{q}_\infty\|_1} & \text{, if } |B| \geq 2 \end{cases} \tag{2.58}$$

is the stationary solution of the original Markov chain (by a similar argument as made above, this is well defined):

$$\begin{aligned} \Gamma\,\boldsymbol{p}_\infty &\overset{\text{Def } \boldsymbol{p}_\infty}{=\!=\!=} \Gamma\,[\operatorname{diag}(\Gamma)]^\circ\,\boldsymbol{q}_\infty + \overbrace{\underbrace{(Q_\Gamma)}_{\mathbb{1}-\Gamma\,[\operatorname{diag}(\Gamma)]^\circ}\,\boldsymbol{q}_\infty}^{\mathbf{0}} - \boldsymbol{q}_\infty = \\ &= \mathbf{0}. \end{aligned} \tag{2.59}$$

2.4.5. Jump unravelling of the master equations

In analogy to the so called quantum jump unravelling of quantum master equation (which we will consider in Section 3.3), we define a jump unravelling of the classical master equation. It is a way to 'extract' a single trajectory $\omega \in \mathcal{U}$ out of the set of all possible trajectories $\mathcal{U}$.

1) Choose a starting state: $X(t=0,\omega) = j \in \Omega$ and a final stopping time $T_f := T_{\text{final}}$

2) Choose the waiting time according to

$$\mathcal{P}(t_{n+1} - t_n > t \,|\, X(t_n) = j) \overset{\text{lemma 16}}{=\!=} \mathrm{e}^{-t\,\gamma_{j\to}} = \mathrm{e}^{t\,\Gamma_{j,j}} \tag{2.60}$$

3a) If $\gamma_{j\to} = 0$ the waiting time is infinite, that is $\mathcal{P}(t_{n+1} = \infty) = 1$ and hence the system will remain in that state forever: $X(t,\omega) = j$ for all $t \geq t_n$.

3b) If $\gamma_{j\to} > 0$ the waiting time is finite, and the probability distribution for the next state is given by:

$$\mathcal{P}\big(X(t_{n+1}) = i \,|\, X(t_n = j)\big) \overset{\text{lemma 17}}{=\!=} \frac{\gamma_{j\to i}}{\gamma_{j\to}} = \frac{\gamma_{j\to i}}{-\Gamma_{jj}} = \frac{\gamma_{j\to i}}{\sum\limits_{k\in\Omega} \gamma_{j\to k}}. \tag{2.61}$$

We now repeat the steps 2) and 3), until the final time is reached, $t_{n+1} > T_f$.

Now that we have computed a single trajectory, the question remains, 'how likely' such a trajectory is, that is, what is the probability density $\mathcal{P}(\omega)$? Clearly, when conducting a numerical simulation, one could compute $p^{(i)}(t \,|\, \boldsymbol{p}_0 = \boldsymbol{e}_j)$ by computing multiple trajectories and taking the average:

$$p^{(i)}(t \,|\, \boldsymbol{p}_0 = \boldsymbol{e}_j) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} 1_{\{X(t,\omega_n)=i \,|\, X_0=j\}} \overset{N\gg 1}{\approx} \frac{1}{N} \sum_{n=1}^{N} 1_{\{X(t,\omega_n)=i \,|\, X_0=j\}}. \tag{2.62}$$

An alternative way is shown in [BP02]: We have $p^{(i)}(t \,|\, \boldsymbol{p}_0 = \boldsymbol{e}_j) = \sum\limits_{n\in\mathbb{N}} p^{(i)}(n,t \,|\, \boldsymbol{p}_0 = \boldsymbol{e}_j)$, where $\big(p^{(i)}(n,t \,|\, \boldsymbol{p}_0 = \boldsymbol{e}_j)\big)_{n\in\mathbb{N}}$ are the probabilities of reaching state $i \in \Omega$ from state $j \in \Omega$ after a time $t \geq 0$ and exactly $n \in \mathbb{N}$ jumps and can be computed iteratively:

$$
\begin{aligned}
p^{(i)}(n=0,t\,|\,\boldsymbol{p}_0=\boldsymbol{e}_j) &= \underbrace{[\mathcal{P}(t_1-t_0>t)]}_{\mathrm{e}^{-t\,\gamma_{j\to}}}\cdot\delta_{i,j} = \delta_{i,j}\,\mathrm{e}^{-t\,\gamma_{j\to}} \\
p^{(i)}(n+1,t\,|\,\boldsymbol{p}_0=\boldsymbol{e}_j) &= \int_0^t \mathrm{d}\tau \sum_{\substack{l\in\Omega\\ l\neq i}} \mathrm{e}^{-(t-\tau)\,\gamma_{i\to}}\,\gamma_{l\to i}\,p^{(l)}(n,\tau\,|\,\boldsymbol{p}_0=\boldsymbol{e}_j).
\end{aligned}
\tag{2.63}
$$

2.4.6. Two different views: Ensemble versus single trajectory

Let us take another look at the master equation. We have:

$$
\begin{aligned}
\boldsymbol{p}(t\,|\,\boldsymbol{p}_0) := \mathrm{e}^{t\,\Gamma}\underbrace{\boldsymbol{p}_0}_{\sum_{j\in\Omega} p_0^{(j)}\boldsymbol{e}_j} &= \sum_{j\in\Omega} p_0^{(j)}\underbrace{\mathrm{e}^{t\,\Gamma}\boldsymbol{e}_j}_{\boldsymbol{p}(t\,|\,\boldsymbol{p}_0=\boldsymbol{e}_j)} \\
= \sum_{j\in\Omega} p_0^{(j)}\underbrace{\boldsymbol{p}(t\,|\,\boldsymbol{p}_0=\boldsymbol{e}_j)}_{\sum_{i\in\Omega}\boldsymbol{e}_i\,p^{(i)}(t\,|\,\boldsymbol{p}_0=\boldsymbol{e}_j)} &= \sum_{i,j\in\Omega}\boldsymbol{e}_i\,p^{(i)}(t\,|\,\boldsymbol{p}_0=\boldsymbol{e}_j)\,p_0^{(j)},
\end{aligned}
\tag{2.64}
$$

so it suffices to compute

$$
p^{(i)}(t\,|\,\boldsymbol{p}_0=\boldsymbol{e}_j) = \int_{\mathcal{U}} \mathrm{d}\,\mathcal{P}(\omega)\,\mathcal{P}\left(X(t,\omega)=i\,|\,X(t=0,\omega)=j\right).
$$

There are two ways to interpret $p^{(i)}(t\,|\,\boldsymbol{p}_0=\boldsymbol{e}_j)$: We can either see it as the fraction of a 'probability mass' which has started at state j and is now (after some time $t>0$) accumulated at state i, or as the probability for a single realization to be at time t in state i, provided that it started at state j.
The connection between these two interpretations, is to interpret the 'probability mass' as an (uncountable) ensemble of trajectories (compare Figure 2.12 for an illustration).
Note that a single trajectory need not converge, neither for continuous-, nor for discrete-time Markov chains. In fact, it converges if and only if the minimal absorbing set the trajectory is eventually captured in, consists of exactly one state.

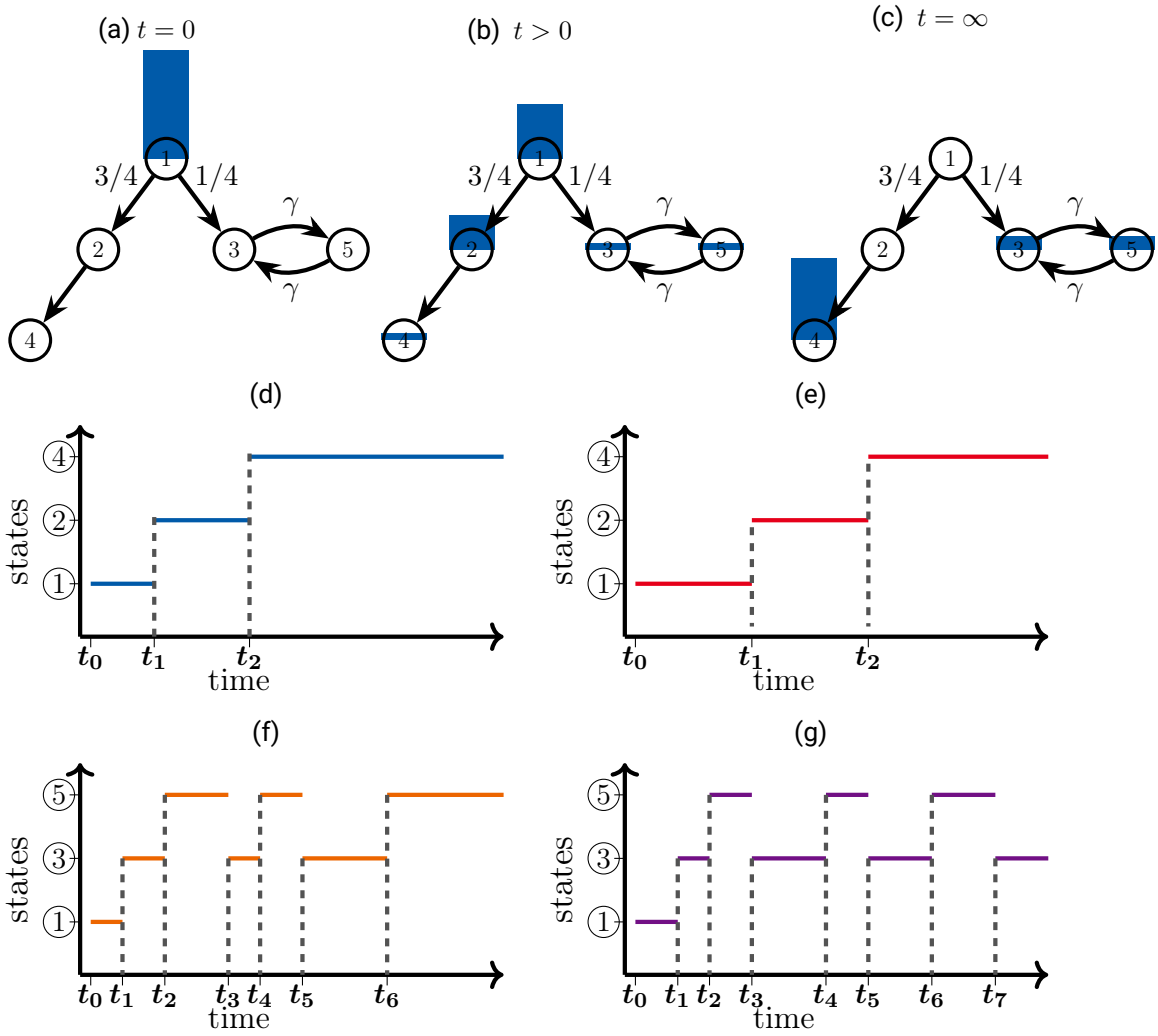


Figure 2.12.: Illustration of the two different interpretations of the time evolution of the master equation: The Figures 2.12a to 2.12c show the 'ensemble interpretation', by illustrating the flow of the 'probability mass' in the state transition network. The behavior is similar to that of a fluid, where the states can be interpreted as 'containers', the links as 'pipelines' and the link strength as the capacity of the pipeline or the pumping rate.
On the other hand, the figures 2.12d to 2.12f shows possible trajectories of a single 'particle', hopping between different states. Averaging over all possible trajectories gives the solution of the master equation: $p^{(i)}(t \,|\, \boldsymbol{p}_0 = \boldsymbol{e}_j) = \int_{\mathcal{U}} 1_{\{X(t,\omega)=i \,|\, X_0=j\}} \,\mathrm{d}\,\mathcal{P}(\omega)$.

2.4.7. The time average of a single trajectory

In Chapter 3 we need to compute the time average of a single quantum trajectory. In favor of a better comparison with the quantum case, we will compute the time average of a single trajectory of a continuous-time Markov chain.
We define:

$$\begin{aligned} \text{the time average as } \langle \bullet \rangle_{t\geq 0} &:= \lim_{T\to\infty} \frac{1}{T} \int_0^T \mathrm{d}t \, \bullet \text{ and the} \\ \text{ensemble average as } \langle \bullet \rangle_{\omega\in\mathcal{U}} &:= \int_{\mathcal{U}} \mathrm{d}\mathcal{P}(\omega) \bullet \\ p_\infty^{(i)}(\omega \,|\, \boldsymbol{p}_0 = \boldsymbol{e}_j) &:= \langle 1_{\{X(t,\omega)=i, X(t=0,\omega)=j\}} \rangle_{t\geq 0} \end{aligned} \tag{2.65}$$

We know that a single trajectory is eventually being 'captured' by some minimal absorbing set $B \in \mathcal{B}$ (see Section A.8 in the appendix). Depending on whether the number of jumps in a trajectory is finite or infinite, its time average differs. In order to compute it, we have to distinguish these two cases:

i) finitely many jumps
In this case, the trajectory will eventually be in some minimal absorbing set $B = \{i\}$, which includes exactly one state:

$$\begin{aligned} p_\infty^{(i)}(\omega \,|\, \boldsymbol{p}_0 = \boldsymbol{e}_j) \overset{X(t,\omega)=i \forall t\geq t_0}{=\!=\!=\!=\!=\!=\!=\!=\!=} & \underbrace{\lim_{T\to\infty} \frac{1}{T} \int_0^{t_0} 1_{\{X(t,\omega)=i, X_0=j\}} \, \mathrm{d}t}_{0} \\ & + \lim_{T\to\infty} \underbrace{\frac{1}{T} \int_{t_0}^{T} 1_{\{X(t,\omega)=i, X_0=j\}} \, \mathrm{d}t}_{\frac{T-t_0}{T}} = 1. \end{aligned} \tag{2.66}$$

ii) infinitely many jumps

For a fixed trajectory $\omega \in \mathcal{U}$ with infinitely many jumps and a fixed state $s \in \Omega$ appearing in the trajectory we denote by $\alpha_s(k)$ the position in the trajectory, where the corresponding state s appears for the k-th time, by $J(t)$ the number of jumps before time $T > 0$ and by $J_s(T)$ the number of times the trajectory has visited state $s \in \Omega$, so

$$
\begin{aligned}
\alpha_s(k) &:= \alpha_s(k,\omega) \\
&:= \min_{M\in\mathbb{N}} \big\{ |\, \{X(t_n,\omega) \,:\, X(t_n,\omega) = s \text{ and } n \le M\} \,| = k \big\} \\
J(T) &:= J(\omega,T) = |\{n \in \mathbb{N} \,:\, t_n \le T\}| \\
J_s(T) &:= J_s(\omega,T) = |\{n \in \{1,\dots,J(T)\} \,:\, X(t_n) = s\}|
\end{aligned}
\tag{2.67}
$$

Note that we define $\alpha_s(k)$ only for those states that actually appear in the trajectory. In particular we have $X(t_{\alpha_s(k)},\omega) = s$ for all $k \in \mathbb{N}$ and for all states s appearing in the trajectory. Moreover, we can rearrange the following sum:

$$
\begin{aligned}
T &= \frac{1}{J(T)} \sum_{n=0}^{J(T)-1} \tau_n \\
&= \sum_{B\in\mathcal{B}} \sum_{b\in B} \frac{J_b(T)}{J(T)} \, \frac{1}{J_b(T)} \sum_{k=0}^{J_b(T)-1} \tau\big(\alpha_b(k)\big) \text{ for all trajectories } \omega \in \mathcal{U},
\end{aligned}
\tag{2.68}
$$

where $\tau_n := \tau(\omega,n) := t(\omega,n+1) - t(\omega,n) = t_{n+1} - t_n$ is the waiting time defined in lemma 16. This rearrangement will be used in Equation (2.72) to compute the time average of a single trajectory for a classical, continuous-time Markov chain.

For all states the waiting times are independent, identically distributed, so by the law of large numbers we have:

$$
\frac{1}{K} \sum_{k=0}^{K-1} \tau(\alpha_b(k)), \xrightarrow{K\to\infty} \overline{\tau_b},
\tag{2.69}
$$

where $\overline{\tau_b}$ is the average waiting time for the state $b \in \Omega$.

Now we are in a position to compute the time-averaged probability, which is the time-averaged fraction of time the trajectory spends in a specific state. Let $i, j \in \Omega$ be two states. Since every trajectory is eventually 'captured' by some minimal absorbing set (compare Section A.8 in the appendix), we have for a transient state $i \in B_0$

$$
p_\infty^{(i)}(\omega \,|\, \boldsymbol{p}_0 = \boldsymbol{e}_j) = \langle 1_{\{X(t,\omega)=i, X_0=j\}} \rangle_{t\ge 0} \xrightarrow{\text{if } i\in B_0} 0,
\tag{2.70}
$$

so the time average vanishes, whenever the final state is not contained in some minimal absorbing set.

On the other hand, when a trajectory is already contained in some minimal absorbing set $B \in \mathcal{B}$, then for every state $s \in B$ the fraction $\frac{J_s(T)}{J(T)}$ converges to the stationary probability $q_\infty^{(s)}(\boldsymbol{p}_0 \in B)$ associated to this minimal absorbing set:

$$\frac{J_s(T)}{J(T)} = \frac{1}{J(T)} \sum_{k=0}^{J(T)-1} \delta_{X(t_n),s} \xrightarrow{T\to\infty} q_\infty^{(s)}(\boldsymbol{p}_0 \in B). \tag{2.71}$$

Then we can compute the time average of a stochastic trajectory ω starting at the state $j \in \Omega$ and assume that this trajectory will eventually be captured by the minimal absorbing set $B \in \mathcal{B}$. This time average can be interpreted as the fraction of time this trajectory spends in the state $i \in \Omega$:

$$
\begin{aligned}
p_\infty^{(i)}(\omega \,|\, \boldsymbol{p}_0 = \boldsymbol{e}_j,\, \omega \in B \text{ eventually }) &= \langle 1_{\{X(t,\omega)=i, X_0=j, \omega\in B \text{ eventually }\}} \rangle_{t\geq 0} \\
&= \lim_{T\to\infty} \frac{1}{T} \underbrace{\int_0^T \mathrm{d}t\, 1_{\{X(t,\omega)=i,\, X_0=j\}}}_{\sum_{n=0}^{J(T)-1} \int_{t_n}^{t_{n+1}} \mathrm{d}t\, 1_{\{X(t,\omega)=i,\, X_0=j\}}} \cdot 1_{\{\omega\in B \text{ eventually }\}} \\
&= \lim_{T\to\infty} \frac{1}{T} \sum_{s\in\Omega} \sum_{k=0}^{J_s(T)-1} \underbrace{\int_0^{\tau(\alpha_s(k))} \mathrm{d}t\, 1_{\{X(t,\omega)=i,\, X_0=j\}}}_{\delta_{s,i}\,\tau(\alpha_s(k))} \cdot 1_{\{\omega\in B \text{ eventually }\}} \\
&= \lim_{T\to\infty} \frac{J(T)}{T} \frac{J_i(T)}{J(T)} \frac{1}{J_i(T)} \sum_{k=0}^{J_i(T)-1} \tau(\alpha_s(k)) \cdot 1_{\{\omega\in B \text{ eventually }\}} \\
&= \lim_{T\to\infty} \frac{\frac{J_i(T)}{J(T)} \frac{1}{J_i(T)} \sum_{k=0}^{J_i(T)-1} \tau(\alpha_i(k))}{\frac{T}{J(T)}} \cdot \frac{1_{\{\omega\in B \text{ eventually }\}}}{1_{\{\omega\in \bigcup_{B'\in\mathcal{B}} B' \text{ eventually }\}}} \\
&= \lim_{T\to\infty} \frac{1_{\{\omega\in B \text{ eventually }\}} \frac{J_i(T)}{J(T)} \frac{1}{J_i(T)} \sum_{k=0}^{J_i(T)-1} \tau(\alpha_i(k))}{\sum_{B'\in\mathcal{B}} 1_{\{\omega\in B' \text{ eventually }\}} \sum_{b\in B} \frac{J_b(T)}{J(T)} \frac{1}{J_b(T)} \sum_{k=0}^{J_b(T)-1} \tau(\alpha_b(k))} \\
&\overset{|\Omega|<\infty}{=\!=} \frac{\overbrace{\left(\lim_{T\to\infty} \frac{J_i(T)}{J(T)} \cdot 1_{\{\omega\in B \text{ eventually }\}}\right)}^{q_\infty^{(i)}(\boldsymbol{p}_0\in B)} \overbrace{\left(\lim_{T\to\infty} \frac{1}{J_i(T)} \sum_{k=0}^{J_i(T)-1} \tau(\alpha_i(k))\right)}^{\overline{\tau_i}}}{\sum_{B'\in\mathcal{B}} \sum_{b'\in B'} \underbrace{\left(\lim_{T\to\infty} 1_{\{\omega\in B' \text{ eventually }\}} \frac{J_{b'}(T)}{J(T)}\right)}_{q_\infty^{(b')}(\boldsymbol{p}_0\in B')} \underbrace{\left(\lim_{T\to\infty} \frac{1}{J_{b'}(T)} \sum_{k=0}^{J_{b'}(T)-1} \tau(\alpha_{b'}(k))\right)}_{\overline{\tau_{b'}}}} \\
&= \frac{q_\infty^{(i)}(\boldsymbol{p}_0 \in B)\, \overline{\tau_i}}{\sum_{B'\in\mathcal{B}} \sum_{b'\in B'} q_\infty^{(b')}(\boldsymbol{p}_0 \in B')\, \overline{\tau_{b'}}} \\
&=: \frac{q_\infty^{(i)}(\boldsymbol{p}_0 \in B)\, \overline{\tau_i}}{\mathcal{N}} = p_\infty^{(i)}(\boldsymbol{p}_0 \in B),
\end{aligned}
\tag{2.72}
$$

where $\mathcal{N} := \sum_{B' \in \mathcal{B}} \sum_{b' \in B'} q_\infty^{(b')}(\boldsymbol{p}_0 \in B')\, \overline{\tau_{b'}}$ is a normalization constant.

We note that this expression (in agreement with Equation (2.58)), is independent of the initial state $\boldsymbol{p}_0 = \boldsymbol{e}_j$ and does not depend on all the states of the trajectory $\omega \in \mathcal{U}$, but only on the minimal absorbing set $B \in \mathcal{B}$ it is eventually captured in.

For the general case we have:

$$\begin{aligned} p_\infty^{(i)}(\omega \,|\, \boldsymbol{p}_0 = \boldsymbol{e}_j) &= \langle 1_{\{X(t,\omega)=i,\, X_0=j,\, \omega \in \bigcup_{B\in\mathcal{B}} B \text{ eventually }\}} \rangle_{t\geq 0} \\ &= \sum_{B\in\mathcal{B}} \underbrace{1_{\{\omega \in B \text{ eventually }, X_0=j\}}}_{1_{\mathcal{U}(j\rightsquigarrow B)}(\omega)} \cdot \underbrace{\langle 1_{\{X(t,\omega)=i,\, X_0=j,\, \omega\in B \text{ eventually }\}} \rangle_{t\geq 0}}_{p_\infty^{(i)}(\boldsymbol{p}_0 \in B)} \\ &= \sum_{B\in\mathcal{B}} 1_{\mathcal{U}(j\rightsquigarrow B)}(\omega)\, \frac{q_\infty^{(i)}(\boldsymbol{p}_0 \in B)\, \overline{\tau_i}}{\sum_{B'\in\mathcal{B}} \sum_{b'\in B'} q_\infty^{(b')}(\boldsymbol{p}_0 \in B')\, \overline{\tau_{b'}}}, \end{aligned} \tag{2.73}$$

where $\mathcal{U}(j \rightsquigarrow B)$ is the set of all trajectories that started at state j and are eventually in the minimal absorbing set $B \in \mathcal{B}$, with $\mathcal{U}(j \rightsquigarrow) := \bigcup_{B\in\mathcal{B}} \mathcal{U}(j \rightsquigarrow B)$ and $\mathcal{P}(\mathcal{U}(j \rightsquigarrow)) = 1$.

When interpreting $p_\infty^{(i)}(\omega)$ as the average fraction of time the trajectory ω spends in state $i \in \Omega$, then we can conclude from Equation (2.72) that the average time spent in a state is proportional to the fraction of visits $q_\infty^{(i)}(\omega)$ to that state times the average waiting time $\overline{\tau_i}$ in that state. These considerations will be important in Section 3.4.

This can be used to show that the time average and the ensemble average commute for any trajectory, starting at some state $j \in \Omega$:

$$
\begin{aligned}
\underbrace{\langle\langle 1_{\{X(t,\omega)=i,\, X_0=j\}}\rangle_{t\geq 0}\rangle_{\omega\in\mathcal{U}}}_{p_\infty^{(i)}(\omega\,|\,\boldsymbol{p}_0=\boldsymbol{e}_j)} &= \int_{\mathcal{U}(j\rightsquigarrow)} \underbrace{p_\infty^{(i)}(\omega\,|\,\boldsymbol{p}_0=\boldsymbol{e}_j)}_{\sum_{B\in\mathcal{B}} 1_{\mathcal{U}(j\rightsquigarrow B)}(\omega)\,\frac{q_\infty^{(i)}(\boldsymbol{p}_0\in B)\,\overline{\tau_i}}{\mathcal{N}}} \mathrm{d}\,\mathcal{P}(\omega) \\
&= \sum_{B\in\mathcal{B}} \frac{q_\infty^{(i)}(\boldsymbol{p}_0\in B)\,\overline{\tau_i}}{\mathcal{N}} \underbrace{\int_{\mathcal{U}(j\rightsquigarrow B)} \mathrm{d}\,\mathcal{P}(\omega)}_{\mathcal{P}\left(\mathcal{U}(j\rightsquigarrow B)\right)} = \\
&= \sum_{B\in\mathcal{B}} \underbrace{\mathcal{P}\big(\mathcal{U}(j\rightsquigarrow B)\big)}_{\mathcal{P}(j\rightsquigarrow B)} \frac{q_\infty^{(i)}(\boldsymbol{p}_0\in B)\,\overline{\tau_i}}{\mathcal{N}} = \qquad (2.74) \\
&\overset{\text{lemma 31}}{=\!=\!=} \lim_{t\to\infty} \left(\mathrm{e}^{t\,\Gamma}\right)_{ij} \\
&= \lim_{t\to\infty} \mathcal{P}(X(t)=i\,|\,X_0=j) = \\
&= \underbrace{\langle \mathcal{P}(X(t)=i\,|\,X_0=j)\rangle_{t\geq 0}}_{\langle 1_{\{X(t,\omega)=i,\, X_0=j\}}\rangle_{\omega\in\mathcal{U}}} = \\
&= \langle\langle 1_{\{X(t,\omega)=i,\, X_0=j\}}\rangle_{\omega\in\mathcal{U}}\rangle_{t\geq 0}.
\end{aligned}
$$

2.5. Analytical expression for the stationary solution of the master equation

2.5.1. The steady state of a minimal absorbing set

In the following, we derive an analytical expression for the stationary state of the master equation in the case that the state transition network is strongly connected. In this case, the kernel of the generator is one-dimensional (see: Section 13) and $\boldsymbol{p}_\infty \in \operatorname{kern}(\Gamma) \cap (\mathbb{R}_{>0})^{|\Omega|}$ has strictly positive entries and is uniquely determined by $\|\boldsymbol{p}_\infty\|_1 = 1$. This fact is known as the 'Markov chain tree theorem'. Various proofs can be found in [MG13; CK78; Wil22; Hil66; LR86]. The original statement was first formulated in 1948 by Tutte [Tut48].

We denote with $[A]_{ij}$ the first minor of the matrix A, that is the determinant of the matrix that results from a matrix $A \in \mathbb{C}^{N\times N}$ by deleting row number i and column number j (For a general definition of minors, see definition 41). With this notation, the adjugate of the matrix A is defined as

$$\operatorname{adj}(A) := \left((-1)^{j+i} \left[A \right]_{j\,i} \right)_{i,j \in \{1,\ldots,N\}} = \begin{pmatrix} (-1)^{1+1} \left[A \right]_{1\,1} & \ldots & (-1)^{N+1} \left[A \right]_{N\,1} \\ \vdots & & \vdots \\ (-1)^{1+N} \left[A \right]_{1\,N} & \ldots & (-1)^{N+N} \left[A \right]_{N\,N} \end{pmatrix}, \tag{2.75}$$

with the property $A \operatorname{adj}(A) = \det(A)\, \mathbb{1}_N = \operatorname{adj}(A)\, A$.
Since the generator Γ is singular, we have

$$0^{|\Omega| \times |\Omega|} = \underbrace{\det(\Gamma)}_{0}\, \mathbb{1}_{|\Omega|} = \Gamma \operatorname{adj}(\Gamma) = \Gamma \left(\underbrace{\operatorname{adj}(\Gamma)\, \boldsymbol{e}_1}_{\in \operatorname{kern}(\Gamma)}, \ldots, \underbrace{\operatorname{adj}(\Gamma)\, \boldsymbol{e}_{|\Omega|}}_{\in \operatorname{kern}(\Gamma)} \right). \tag{2.76}$$

This means that all columns of the adjugate must lie in the kernel of Γ ($\operatorname{adj}(\Gamma)\, \boldsymbol{e}_i \in \operatorname{kern}(\Gamma)$ for all $i \in \Omega$) and are therefore proportional to the unique stationary solution $\operatorname{adj}(\Gamma)\, \boldsymbol{e}_i \propto \boldsymbol{p}_\infty$. Hence we have for all $i \in \Omega$:

$$
\begin{aligned}
\boldsymbol{p}_\infty &\overset{(2.76)}{\propto} \operatorname{adj}(\Gamma)\,\boldsymbol{e}_i = \Big(\underbrace{\operatorname{adj}(\Gamma)_{ki}}_{(-1)^{k+i}\,[\Gamma]_{k,\,i}} \Big)_{k\in\{1,\dots,|\Omega|\}} = \\
&\overset{(2.75)}{=\!=} (-1)^{k+i} \Big(\underbrace{[\Gamma]_{k\,i}}_{(-1)^{k+i}\,[\Gamma]_{k,\,k}} \Big)_{k\in\{1,\dots,|\Omega|\}} = \\
&\overset{\text{lemma 43}}{=\!=} \begin{pmatrix} [\Gamma]_{1,\,1} \\ \vdots \\ [\Gamma]_{|\Omega|,|\Omega|} \end{pmatrix} \overset{(A.26)}{=\!=} \begin{pmatrix} \sum_{T\,\in\,\mathcal{T}(\to 1,\mathcal{S})} \gamma_T \\ \vdots \\ \sum_{T\,\in\,\mathcal{T}(\to |\Omega|,\mathcal{S})} \gamma_T \end{pmatrix}, \\
\Longrightarrow \boldsymbol{p}_\infty &= \frac{1}{\mathcal{Z}_\Omega} \begin{pmatrix} \sum_{T\,\in\,\mathcal{T}(\to 1,\mathcal{S})} \gamma_T \\ \vdots \\ \sum_{T\,\in\,\mathcal{T}(\to |\Omega|,\mathcal{S})} \gamma_T \end{pmatrix}
\end{aligned}
\tag{2.77}
$$

with the normalization factor

$$\mathcal{Z}_\Omega := \sum_{j\in\Omega} \sum_{T\in\,\mathcal{T}(\to j,\mathcal{S})} \gamma_T \text{ and}$$

$$\mathcal{T}(\to j,\mathcal{S}) := \{\text{in-trees in the network } \mathcal{S} \text{ rooted at the state } j \in \Omega\},$$

compare definition 36.

We used the definition of the adjugate, that is expressing its entries via first minors ($\operatorname{adj}(\Gamma)_{ij} = (-1)^{i+j}\,[\Gamma]_{j,i}$) and the fact that for the generator matrix Γ, every first minor $[\Gamma]_{j,i}$ is related to the first principal minors, via $[\Gamma]_{j,i} = (-1)^{i+j}[\Gamma]_{i,i}$. The first principal minors of Γ can be computed via the sums of the weights of all possible in-trees, as shown in theorem 42.

We now want to talk briefly about the computational complexity needed to compute all stationary solutions $\boldsymbol{p}_\infty(\boldsymbol{p}_0 \in B)$. It is possible to search numerically for strongly connected components in linear time [GM78]. When merging these strongly connected components into a single state via graph condensation [GM78] and following a single

path to its end, one arrives at a minimal absorbing set. This requires $\mathcal{O}\big(|\Omega| + |\mathcal{E}|\big)$ time. Computing the stationary solution $\boldsymbol{p}_B$ for a minimal absorbing set $B \in \mathcal{B}$ can be done in $\mathcal{O}\big(|\mathcal{E}_B| \cdot N_B\big)$, where N_B is the total number of in-trees of the subnetwork $\mathcal{S}_B := \{B, \{(i,j) \in \mathcal{E} \,:\, i,j \in B\}\}$. This results in a total time of $\mathcal{O}\big(|\Omega| + |\mathcal{E}| + \sum_{B \in \mathcal{B}} |\mathcal{E}_B| \cdot N_B\big)$.

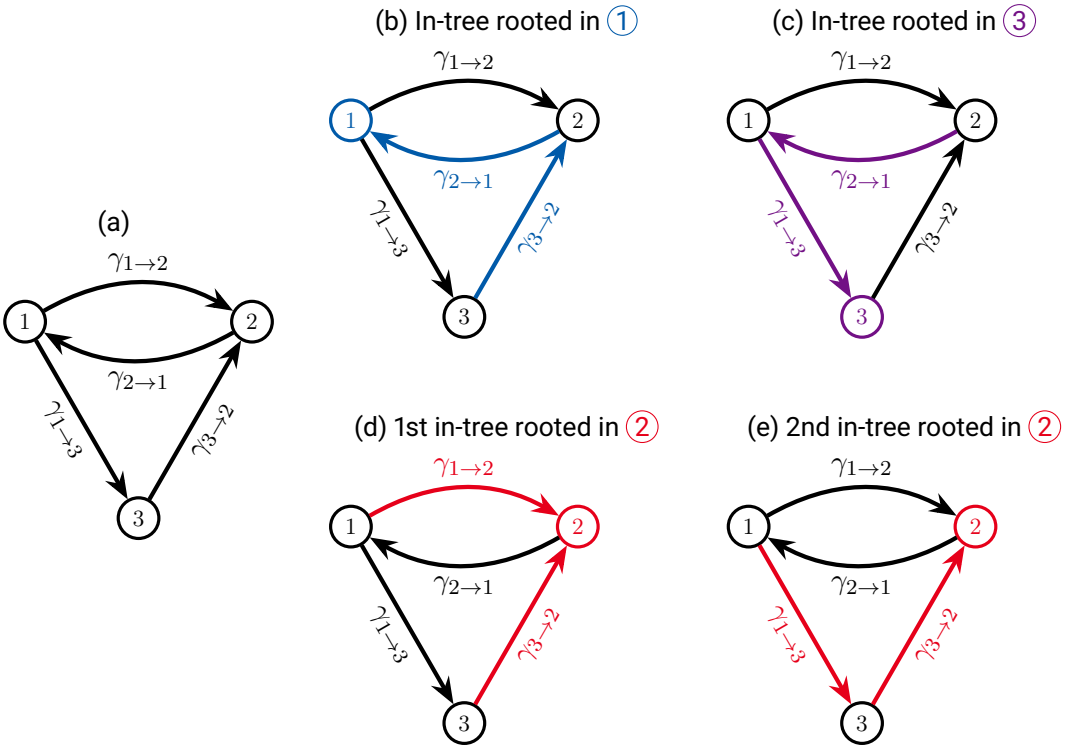


Figure 2.13.: Example of a strongly connected network (Figure 2.13a) together with the corresponding in-trees rooted in state number 1 (Figure 2.13b) and 3 (Figure 2.13c), and the two in-trees rooted in state number 2 (Figure 2.13d and 2.13e). The kernel of Γ is the span of a vector whose i-th component is the sum over all in-trees rooted in state number i of the product of the rates of all edges that constitute that particular in-tree. In this example we have:

$$\text{kern}\,(\Gamma) = \text{span}\left\{\begin{pmatrix} \gamma_{3\to2} \cdot \gamma_{2\to1} \\ \gamma_{1\to2} \cdot \gamma_{3\to2} + \gamma_{1\to3} \cdot \gamma_{3\to2} \\ \gamma_{2\to1} \cdot \gamma_{1\to3} \end{pmatrix}\right\}.$$

2.5.2. Explicit expression for the probabilities $\mu_B(\boldsymbol{p}_0)$

We note that the coefficients μ_B in front of the vectors $\boldsymbol{p}_\infty(\boldsymbol{p}_0 \in B)$ are non-negative and sum up to one:

$$\begin{aligned}
\mu_B(\boldsymbol{p}_0) = \mu_B(\boldsymbol{p}_0) \cdot \underbrace{\sum_{b\in B} (\boldsymbol{p}_\infty(\boldsymbol{p}_0 \in B))^{(b)}}_{1} = \sum_{b\in B} \underbrace{\mu_B(\boldsymbol{p}_0) \cdot (\boldsymbol{p}_\infty(\boldsymbol{p}_0 \in B))^{(b)}}_{(\boldsymbol{p}_\infty(\boldsymbol{p}_0))^{(b)}} \geq 0 \\
1 = \sum_{j\in\Omega} \left(\underbrace{\boldsymbol{p}_\infty(\boldsymbol{p}_0)}_{\sum_{B\in\mathcal{B}} \mu_B(\boldsymbol{p}_0)\cdot \boldsymbol{p}_\infty(\boldsymbol{p}_0\in B)} \right)^{(j)} = \sum_{B\in\mathcal{B}} \mu_B(\boldsymbol{p}_0) \cdot \underbrace{\sum_{j\in\Omega} (\boldsymbol{p}_\infty(\boldsymbol{p}_0 \in B))^{(j)}}_{1} = \\
= \sum_{B\in\mathcal{B}} \mu_B(\boldsymbol{p}_0).
\end{aligned} \tag{2.78}$$

This means, we can interpret $(\mu_B(\boldsymbol{p}_0))_{B\in\mathcal{B}}$ as a probability distribution, with $\mu_B(\boldsymbol{p}_0)$ being the fraction of probability 'mass' contained in the minimal absorbing set $B \in \mathcal{B}$, or alternatively, as probability, that a single trajectory is eventually being captured by the minimal absorbing set $B \in \mathcal{B}$, so

$$\mu_B(\boldsymbol{p}_0) = \mathcal{P}(B \,|\, \boldsymbol{p}_0) =: \mathcal{P}(\boldsymbol{p}_0 \rightsquigarrow B \text{ eventually }). \tag{2.79}$$

This means that these coefficients $(\mu_B)_{B\in\mathcal{B}}$, appearing in a basis decomposition of the initial state (compare Equation (2.25)), have an interpretation as a probability. In order to find an analytical expression, we consider:

$$\begin{aligned}
\mu_B(\boldsymbol{e}_j) &= \mathcal{P}(B \,|\, \boldsymbol{p}_0 = \boldsymbol{e}_j) = \mathcal{P}(j \rightsquigarrow B \text{ eventually }) = \\
&= \lim_{t\to\infty} \underbrace{\sum_{b\in B} \left(\mathrm{e}^{t\,\Gamma}\right)_{b,j}}_{\mathcal{P}(X_t\in B \,|\, X_0=j)} = \lim_{k\to\infty} \underbrace{\mathcal{P}(X_{t_k} \in B \,|\, X_0 = j)}_{\sum_{b\in B} (Q_\Gamma)^k_{b,j}} = \\
&= \lim_{k\to\infty} \sum_{b\in B} (Q_\Gamma)^k_{b,j} = \sum_{\omega\in \bigcup_{n\in\mathbb{N}} \mathcal{W}\left(\mathcal{S}^{\text{disc}}_{\text{emb}}, j \xrightarrow{n} B\right)} q_\omega,
\end{aligned} \tag{2.80}$$

where the sum goes over all walks in the network $\mathcal{S}^{\text{disc}}_{\text{emb}}$ associated to the embedded, discrete-time Markov chain (defined in Section 2.4.1) from the state j to any state $b \in B$.

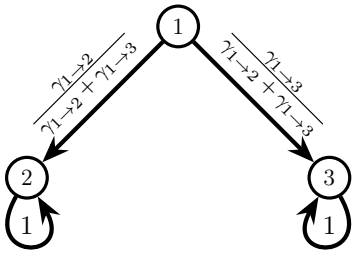
Note, that we have seen, that the expression $\lim_{k\to\infty} \sum_{b\in B} (Q_\Gamma)^k_{b,j}$ indeed converges, since it is monotonously increasing and bounded (compare Equation (2.50)).
When $\omega = (j = a_0, \ldots, \underbrace{a_L}_{\in B})$ is a walk of Length $L \in \mathbb{N}$ from state j to the minimal absorbing set B, with $(a_0, \ldots, a_L) \in \Omega^L$, then the corresponding weight of this walk is given by

$$q_\omega := \prod_{j=0}^{L-1} \underbrace{q_{a_j \to a_{j+1}}}_{\frac{\gamma_{a_j \to a_{j+1}}}{\gamma_{a_j \to}}}. \tag{2.81}$$

When we have $j \in B$, then $\mu_B(j) = 1$.
Figure 2.14 illustrates what the expression $\mu_B(j)$ means in terms of weight of walks in the embedded network $\mathcal{S}_{\text{emb}}$. Whereas the maximum length of the walks to the minimal absorbing sets is bounded in the two figures 2.14a and 2.14b, there are walks of arbitrary length possible in Figure 2.14c. For example, the probability for a trajectory starting in state 1 to eventually reach the minimal absorbing set $\{3\}$ is given by :

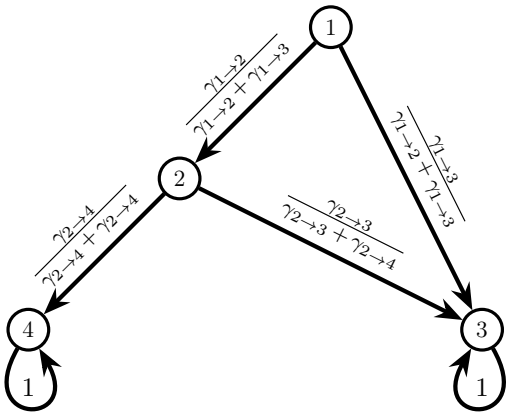
$$\begin{aligned} \mathcal{P}(1 \rightsquigarrow \{3\}) &= \frac{\gamma_{1\to 3}}{\gamma_{1\to 2} + \gamma_{1\to 3}} \cdot \underbrace{\sum_{n\geq 0} \left(\frac{\gamma_{1\to 2}}{\gamma_{1\to 2} + \gamma_{1\to 3}} \cdot \frac{\gamma_{2\to 1}}{\gamma_{2\to 1} + \gamma_{2\to 4}} \right)^n}_{\frac{1}{1 - \frac{\gamma_{1\to 2}}{\gamma_{1\to 2}+\gamma_{1\to 3}} \cdot \frac{\gamma_{2\to 1}}{\gamma_{2\to 1}+\gamma_{2\to 4}}}} \\ &= \frac{\gamma_{1\to 2}\,\gamma_{2\to 1} + \gamma_{1\to 3}\,\gamma_{2\to 4}}{\gamma_{1\to 2}\,\gamma_{2\to 1} + \gamma_{1\to 3}\,\gamma_{2\to 4} + \gamma_{1\to 2}\,\gamma_{2\to 4}}. \end{aligned} \tag{2.82}$$

With the initial state being given by $\boldsymbol{p}_0 = \sum_{j\in\Omega} p_0^{(j)}\, \boldsymbol{e}_j$, we have:



(a) $\cdot\mathcal{P}(1 \rightsquigarrow \{2\}) = \frac{\gamma_{1\to2}}{\gamma_{1\to2}+\gamma_{1\to3}}$,

$\cdot\mathcal{P}(1 \rightsquigarrow \{3\}) = \frac{\gamma_{1\to3}}{\gamma_{1\to2}+\gamma_{1\to3}}$

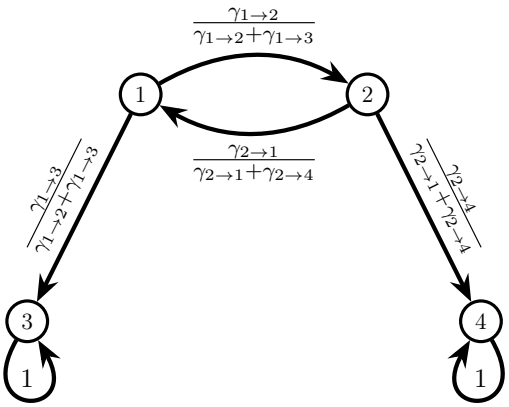


(b) $\cdot\mathcal{P}(1 \rightsquigarrow \{3\}) = \frac{\gamma_{1\to3}}{\gamma_{1\to2}+\gamma_{1\to3}} + \frac{\gamma_{1\to2}}{\gamma_{1\to2}+\gamma_{1\to3}} \cdot \frac{\gamma_{2\to3}}{\gamma_{2\to3}+\gamma_{2\to4}}$,

$\cdot\mathcal{P}(1 \rightsquigarrow \{4\}) = \frac{\gamma_{1\to2}}{\gamma_{1\to2}+\gamma_{1\to3}} \cdot \frac{\gamma_{2\to4}}{\gamma_{2\to3}+\gamma_{2\to4}}$,

$\cdot\mathcal{P}(2 \rightsquigarrow \{3\}) = \frac{\gamma_{2\to3}}{\gamma_{2\to3}+\gamma_{2\to4}}$,

$\cdot\mathcal{P}(2 \rightsquigarrow \{4\}) = \frac{\gamma_{2\to4}}{\gamma_{2\to3}+\gamma_{2\to4}}$



(c) $\cdot\mathcal{P}(1 \rightsquigarrow \{3\}) = \frac{\gamma_{1\to3}\,\gamma_{2\to1} + \gamma_{1\to3}\,\gamma_{2\to4}}{\gamma_{1\to3}\,\gamma_{2\to1} + \gamma_{1\to3}\,\gamma_{2\to4} + \gamma_{1\to2}\,\gamma_{2\to4}}$,

$\cdot\mathcal{P}(1 \rightsquigarrow \{4\}) = \frac{\gamma_{1\to2}\,\gamma_{2\to4}}{\gamma_{1\to3}\,\gamma_{2\to1} + \gamma_{1\to3}\,\gamma_{2\to4} + \gamma_{1\to2}\,\gamma_{2\to4}}$,

$\cdot\mathcal{P}(2 \rightsquigarrow \{4\}) = \frac{\gamma_{1\to2}\,\gamma_{2\to4} + \gamma_{1\to3}\,\gamma_{2\to4}}{\gamma_{1\to3}\,\gamma_{2\to1} + \gamma_{1\to3}\,\gamma_{2\to4} + \gamma_{1\to2}\,\gamma_{2\to4}}$,

$\cdot\mathcal{P}(2 \rightsquigarrow \{3\}) = \frac{\gamma_{1\to3}\,\gamma_{2\to1}}{\gamma_{1\to3}\,\gamma_{2\to1} + \gamma_{1\to3}\,\gamma_{2\to4} + \gamma_{1\to2}\,\gamma_{2\to4}}$

Figure 2.14.: Illustrating the analytical expression of the coefficient $\mu_B(j) = \mathcal{P}(j \rightsquigarrow B)$ for three examples of embedded networks $\mathcal{S}^{\text{disc}}_{\text{emb}}$.

$$
\begin{aligned}
\sum_{B\in\mathcal{B}} \mu_B(\boldsymbol{p}_0)\, \boldsymbol{p}_\infty(\boldsymbol{p}_0 \in B) = \boldsymbol{p}_\infty(\boldsymbol{p}_0) &= \lim_{t\to\infty} \left(\mathrm{e}^{t\,\Gamma}\right)\, \boldsymbol{p}_0 \\
&\overset{\boldsymbol{p}_0 = \sum_{j\in\Omega} p_0^{(j)}\, \boldsymbol{e}_j}{=\!=\!=\!=\!=\!=\!=\!=\!=\!=} \sum_{j\in\Omega} p_0^{(j)} \left(\underbrace{\lim_{t\to\infty} \mathrm{e}^{t\,\Gamma}\, \boldsymbol{e}_j}_{\sum_{B\in\mathcal{B}} \mu_B(\boldsymbol{e}_j)\, \boldsymbol{p}_\infty(\boldsymbol{p}_0\in B)} \right) = \\
&= \sum_{B\in\mathcal{B}} \underbrace{\left(\sum_{j\in\Omega} \mu_B(\boldsymbol{e}_j)\, p_0^{(j)} \right)}_{\mu_B(\boldsymbol{p}_0)} \boldsymbol{p}_\infty(\boldsymbol{p}_0 \in B) \text{ and} \qquad (2.83) \\
\implies \mu_B(\boldsymbol{p}_0) &= \sum_{j\in\Omega} p_0^{(j)}\, \mu_B(\boldsymbol{e}_j) \\
&\overset{(2.80)}{=\!=\!=} \sum_{j\in\Omega} p_0^{(j)} \sum_{\omega\in \bigcup_{n\in\mathbb{N}} \mathcal{W}\left(\mathcal{S}_{\mathrm{emb}}^{\mathrm{disc}},\, j \xrightarrow{n} B\right)} q_\omega .
\end{aligned}
$$

2.6. Time reversible Markov chains, detailed balance and Kolmogorov's criterion

In the following Section we will explore conditions under which a Markov chain is time reversible. This will lead us to the concept of detailed balance and Kolmogorov's criterion.

Definition 18 (Stationary and time reversible Markov chains).

A probability vector (both for discrete times, as well as for continuous times) on a discrete state space Ω is called

- stationary with respect to $\boldsymbol{p}_* := (\mathcal{P}(X_{t_k} = i))_{i\in\Omega} \in (\mathbb{R}_{\geq 0})^{|\Omega|}$ if

$$
\mathcal{P}\Big(X_{t_j} = i_j \,:\, j \in \{1,\ldots,n\}\Big) = \mathcal{P}\Big(X_{t+t_j} = i_j \,:\, j \in \{1,\ldots,n\}\Big) \text{ and} \qquad (2.84)
$$

- time reversible with respect to $\boldsymbol{p}_* := (\mathcal{P}(X_{t_k} = i))_{i\in\Omega} \in (\mathbb{R}_{\geq 0})^{|\Omega|}$ if

$$\mathcal{P}\Big(X_{t_j} = i_j \,:\, j \in \{1,\ldots,n\}\Big) = \mathcal{P}\Big(X_{t-t_j} = i_j \,:\, j \in \{1,\ldots,n\}\Big), \tag{2.85}$$

for all times $t_1 < \ldots < t_n$, $t \in \mathbb{R}$ and all states $i_1, \ldots, i_n \in \Omega$ with $n \in \mathbb{N}$.

A reversible Markov chain is always in a stationary state:

$$\begin{aligned}
&\mathcal{P}\Big(X_{t_j} = i_j \,:\, j \in \{1,\ldots,n\}\Big) \overset{\text{reversible}}{=\!=} \mathcal{P}\Big(X_{t-t_j} = i_j \,:\, j \in \{1,\ldots,n\}\Big) \\
&\quad \overset{t=0}{=\!=} \mathcal{P}\Big(X_{-t_j} = i_j \,:\, j \in \{1,\ldots,n\}\Big) \overset{\text{reversible}}{=\!=} \mathcal{P}\Big(X_{t-(-t_j)} = i_j \,:\, j \in \{1,\ldots,n\}\Big) \\
&\quad = \mathcal{P}\Big(X_{t+t_j} = i_j \,:\, j \in \{1,\ldots,n\}\Big).
\end{aligned} \tag{2.86}$$

What reversibility in a Markov chain means, is that a movie of its time evolution (once a stationary state is reached) looks equivalently, when it is played backwards.
We note that this does not mean we can invert the process in some way, say:

$$\begin{aligned}
&\boldsymbol{p}_1 := \boldsymbol{p}(t \,|\, \boldsymbol{p}_0) \text{ and} \\
&\boldsymbol{p}_0 = \boldsymbol{p}(t \,|\, \boldsymbol{p}_1).
\end{aligned} \tag{2.87}$$

We also note that the 'time-inversion operator' $\mathsf{e}^{-t\,\Gamma}$ does not necessarily map probability vectors to probability vectors. We only have

$$\mathsf{e}^{-t_1\,\Gamma} : \{\mathsf{e}^{t\,\Gamma}\,\boldsymbol{p}_0 \,:\, t \geq t_1\} \to \{\mathsf{e}^{(t-t_1)\,\Gamma}\,\boldsymbol{p}_0 \,:\, t \geq t_1\}. \tag{2.88}$$

Definition 19 (Detailed balance).

We recall the definition of the stationary solution of a Markov chain. A Markov chain is defined as

$$\begin{aligned}
&\dot{\boldsymbol{p}}(t) = \Gamma\,\boldsymbol{p}(t), \text{ for the continuous-time case and} \\
&\Delta\boldsymbol{q}_n := \boldsymbol{q}_{n+1} - \boldsymbol{q}_n = (Q - \mathbb{1})\,\boldsymbol{q}_n, \text{ for the discrete-time case.}
\end{aligned} \tag{2.89}$$

A probability vector is called stationary if the right hand side of Equation (2.89) vanishes:

$$\begin{aligned}
\Gamma\,\boldsymbol{p}_* = \boldsymbol{0} \iff \text{for all } i \in \Omega : 0 &= \sum_{j\in\Omega}\Big(p_*^{(j)}\,\gamma_{j\to i} - p_*^{(i)}\,\gamma_{i\to j}\Big) \\
(Q - \mathbb{1})\,\boldsymbol{q}_* = \boldsymbol{0} \iff \text{for all } i \in \Omega : 0 &= \sum_{j\in\Omega}\Big(q_*^{(j)}\,q_{j\to i} - q_*^{(i)}\,q_{i\to j}\Big).
\end{aligned} \tag{2.90}$$

When in Equation (2.90) every summand vanishes (as apposed to the whole sum), we say that the stationary solution exhibits detailed balance.
Note that a statement about detailed balance is only non-trivial if both states i and j lie in the same minimal absorbing set.

Theorem 20. A Markov chain is time reversible with respect to $\boldsymbol{p}_*$ if and only if, $\boldsymbol{p}_*$ satisfies the detailed balance condition.

Proof of the discrete-time version :

‘ $\Longrightarrow$ ’ Suppose, the Markov chain is reversible. With $p_*^{(i)} := \mathcal{P}(X_n = i)$ we have:

$$\begin{aligned} p_*^{(i)} \cdot q_{i\to j} &= \underbrace{\mathcal{P}(X_{n+1}=j \,|\, X_n = i) \cdot \mathcal{P}(X_n = i)}_{\mathcal{P}(X_{n+1}=j,\, X_n=i)} = \mathcal{P}(X_{n+1}=j,\, X_n=i) \\ &\overset{\text{reversible}}{=\!=} \mathcal{P}(X_{n+1}=i,\, X_n=j) = \underbrace{\mathcal{P}(X_n=j)}_{p_*^{(j)}} \cdot \underbrace{\mathcal{P}(X_{n+1}=i \,|\, X_n=j)}_{q_{j\to i}} \\ &= p_*^{(j)} \cdot q_{j\to i}. \end{aligned} \tag{2.91}$$

This means that the Markov chain exhibits detailed balance.

‘ $\Longleftarrow$ ’ When the condition for detailed balance is fulfilled for every two states we can define the ‘starting probability’ as $\mathcal{P}(X_1 = i_1) := p_*^{(i_1)}$ and compute the transition probability according to:

$$\begin{aligned} \mathcal{P}(X_1 = i_1, \ldots, X_n = i_n) &= \prod_{k=1}^{n-1} \underbrace{\mathcal{P}(X_{k+1} = i_{k+1} \,|\, X_k = i_k)}_{q_{i_k\to i_{k+1}}} \cdot \underbrace{\mathcal{P}(X_1 = i_1)}_{p_*^{(i_1)}} \\ &= p_*^{(i_1)} \prod_{k=1}^{n-1} q_{i_k\to i_{k+1}} \overset{\substack{\text{detailed balance} \\ \text{successively}}}{=\!=} \underbrace{p_*^{(i_n)}}_{\mathcal{P}(X_1=i_n)} \cdot \prod_{k=1}^{n-1} \underbrace{q_{i_{k+1}\to i_k}}_{\mathcal{P}(X_{k+1}=i_k \,|\, X_k=i_{k+1})} \\ &= \mathcal{P}(X_1 = i_n, \ldots, X_n = i_1). \end{aligned} \tag{2.92}$$

This means that the Markov chain is reversible.

□

Proof of the continuous-time version :

'$\Longrightarrow$' Suppose, the Markov chain is reversible. With $p_*^{(j)} := \mathcal{P}(X_t = j)$ we have:

$$\begin{aligned} p_*^{(j)} \cdot \gamma_{j\to i} &= \mathcal{P}(X_t = j) \cdot \lim_{\epsilon\to 0^+} \frac{\mathcal{P}(X_{t+\epsilon} = i \,|\, X_t = j)}{\epsilon} \\ &= \lim_{\epsilon\to 0^+} \frac{\mathcal{P}(X_{t+\epsilon} = i,\, X_t = j)}{\epsilon} \\ &\overset{\text{reversible}}{=\!=\!=} \lim_{\epsilon\to 0^+} \frac{\mathcal{P}(X_{t+\epsilon} = i,\, X_t = j)}{\epsilon} \\ &= \underbrace{\mathcal{P}(X_t = i)}_{p_*^{(i)}} \cdot \underbrace{\lim_{\epsilon\to 0^+} \frac{\mathcal{P}(X_{t+\epsilon} = j \,|\, X_t = i)}{\epsilon}}_{\gamma_{i\to j}} \\ &= p_*^{(i)} \cdot \gamma_{i\to j}. \end{aligned} \tag{2.93}$$

This means that the Markov chain exhibits detailed balance.

'$\Longleftarrow$' When the condition for detailed balance is fulfilled for every two states we can define the 'starting probability' as $\mathcal{P}(X_t = i_1) := p_*^{(i_1)}$ and compute the probability density function

$$\rho(\cdot, T) : (\Omega \times \mathbb{R}_{\geq 0})^n \to \mathbb{R}_{\geq 0}$$

for the transition probability for a fixed time $T > 0$ according to:

$$
\begin{aligned}
\rho\left(\overset{\tau_1}{i_1} \to \cdots \to \overset{\tau_n}{i_n},\, T\right) &= p_*^{(i_1)} \left(\prod_{k=1}^{n-1} \gamma_{i_k\to}\, \mathrm{e}^{-\tau_k\, \gamma_{i_k\to}}\, \frac{\gamma_{i_k\to i_{k+1}}}{\gamma_{i_k\to}}\right) \mathrm{e}^{-\tau_n\, \gamma_{i_n\to}} \\
&\overset{\text{detailed balance}}{=\!=\!=} p_*^{(i_n)} \left(\prod_{k=1}^{n-1} \gamma_{i_{k+1}\to i_k}\right) \mathrm{e}^{-\sum_{k=1}^{n} \tau_k\, \gamma_{k\to}} \\
&= p_*^{(i_n)} \left(\prod_{k=1}^{n-1} \gamma_{i_{k+1}\to}\, \mathrm{e}^{-\tau_{k+1}\, \gamma_{i_{k+1}\to}}\, \frac{\gamma_{i_{k+1}\to i_k}}{\gamma_{i_{k+1}\to}}\right) \mathrm{e}^{-\tau_1\, \gamma_{i_1\to}} \\
&= \rho\left(\overset{\tau_n}{i_n} \to \cdots \to \overset{\tau_1}{i_1},\, T\right).
\end{aligned}
$$

By integrating over all possible waiting times, we get for a given path

$$
\begin{aligned}
\mathcal{P}\Big(i_1 \to \cdots \to i_n,\, T\Big) &= \int_0^T \mathrm{d}\,\tau_1 \cdots \int_0^T \mathrm{d}\,\tau_n\, 1_{\left\{\sum_{k=1}^{n} \tau_k = T\right\}}\, \rho\left(\overset{\tau_1}{i_1} \to \cdots \to \overset{\tau_n}{i_n},\, T\right) \\
&\overset{\substack{\text{exchanging} \\ \text{integrals}}}{=\!=\!=} \int_0^T \mathrm{d}\,\tau_n \cdots \int_0^T \mathrm{d}\,\tau_1\, 1_{\left\{\sum_{k=1}^{n} \tau_k = T\right\}}\, \rho\left(\overset{\tau_n}{i_n} \to \cdots \to \overset{\tau_1}{i_1},\, T\right) \\
&= \mathcal{P}\left(i_n \to \cdots \to i_1,\, T\right),
\end{aligned}
\tag{2.94}
$$

where the notation $(i_1 \to \cdots \to i_n)$ means that the system moves sequentially through the states i_1 to i_n, while the notation $\left(\overset{\tau_1}{i_1} \to \cdots \to \overset{\tau_n}{i_n}\right)$ additionally requires the system to stay in the state $i_k \in \Omega$ for the time $\tau_k > 0$.

From Equation (2.94) we can deduce that the probability for one sequence of transitions is the same as the reverse sequence, hence the Markov chain is reversible.

□

Theorem 21 (Kolmogorov's criterion). A Markov chain exhibits detailed balance with respect to $\boldsymbol{p}_*$ if and only if there are no net 'circular flows' within a minimal absorbing set, that is for any finite number of states $i_1, \ldots, i_n \in B \in \mathcal{B}$, we have

$$\begin{aligned}
&\prod_{k=1}^{n-1} q_{i_k \to i_{k+1}} \cdot q_{i_n \to i_1} = \prod_{k=1}^{n-1} q_{i_{k+1} \to i_k} \cdot q_{i_1 \to i_n} \text{ for the discrete-times and} \\
&\prod_{k=1}^{n-1} \gamma_{i_k \to i_{k+1}} \cdot \gamma_{i_n \to i_1} = \prod_{k=1}^{n-1} \gamma_{i_{k+1} \to i_k} \cdot \gamma_{i_1 \to i_n} \text{ for the continuous-times,}
\end{aligned} \tag{KC}$$

where $B \in \mathcal{B}$ is a minimal absorbing set and $n \in \mathbb{N}$ a natural number.

proof of the discrete-time version.

'$\Longrightarrow$' Suppose the condition for detailed balance holds. Let $i_1, \ldots, i_n \in B$ be states in a minimal absorbing set, such that there is a closed path $i_1 \to \ldots, \to i_n$. Then we have:

$$\begin{aligned}
p_*^{(i_1)}\, q_{i_1 \to i_2} &= p_*^{(i_2)}\, q_{i_2 \to i_1} \\
\vdots \qquad &\qquad \vdots \\
p_*^{(i_n)}\, q_{i_n \to i_1} &= p_*^{(i_1)}\, q_{i_1 \to i_n}
\end{aligned} \tag{2.95}$$

After multiplying these equations and keeping in mind that (since all states are in some minimal absorbing set) all stationary probabilities are strictly positive, we conclude that Kolmogorov's criterion is fulfilled:

$$\prod_{k=1}^{n-1} q_{i_k \to i_{k+1}}\, q_{i_n \to i_1} = \prod_{k=1}^{n-1} q_{i_{k+1} \to i_k}\, q_{i_1 \to i_n}.$$

'$\Longleftarrow$' Suppose, that Kolmogorov's criterion holds true. To check for detailed balance, we choose a minimal absorbing set $B \in \mathcal{B}$ with $|B| \geq 3$ (otherwise, detailed balance is always fulfilled) and three, pairwise different states within it: i, j, $x \in B$.

Since B is strongly connected, every two states are mutually reachable and we denote with $q(i \rightsquigarrow j)$ the weight of the path from state i to state j.

We define:

$$\begin{aligned}
p_*^{(k)} &:= \frac{1}{\mathcal{Z}} \frac{q(x \rightsquigarrow k)}{q(k \rightsquigarrow x)} \text{ for all states } k \in \Omega \backslash \{x\} \text{ and} \\
p_*^{(x)} &:= \frac{1}{\mathcal{Z}},
\end{aligned} \tag{2.96}$$

where $\frac{1}{\mathcal{Z}}$ is a normalization constant, such that $\left(p_*^{(i)}\right)_{i\in\Omega}$ is a probability vector. It suffices to show that this probability vector satisfies the detailed balance condition and is therefore stationary.

Note, that definition (2.96) is independent of the choice of the specific path, since for two paths $i \overset{(1)}{\rightsquigarrow} x$ and $i \overset{(2)}{\rightsquigarrow} x$, as well as the reverse paths $x \overset{(1)}{\rightsquigarrow} i$ and $x \overset{(2)}{\rightsquigarrow} i$ we have:

$$\begin{aligned} q(i \overset{(1)}{\rightsquigarrow} x) \cdot q(x \overset{(2)}{\rightsquigarrow} i) &\overset{\text{Kolmogorov's criterion}}{=\!=\!=} q(i \overset{(2)}{\rightsquigarrow} x) \cdot q(x \overset{(1)}{\rightsquigarrow} i) \\ \Longrightarrow \frac{q(x \overset{(1)}{\rightsquigarrow} i)}{q(i \overset{(1)}{\rightsquigarrow} x)} &= \frac{q(x \overset{(2)}{\rightsquigarrow} i)}{q(i \overset{(2)}{\rightsquigarrow} x)}. \end{aligned} \tag{2.97}$$

Let us now consider a closed path in the network: $j \rightsquigarrow x \rightsquigarrow i \to j$ and its reversed path. By Kolmogorov's criterion, both paths have the same weight, that is:

$$\begin{aligned} q(j \rightsquigarrow x) \cdot q(x \rightsquigarrow i) \cdot q_{i\to j} &\overset{\text{Kolmogorov's criterion}}{=\!=\!=} q_{j\to i} \cdot q(i \rightsquigarrow x) \cdot q(x \rightsquigarrow j) \\ \Longrightarrow \frac{q(x \rightsquigarrow i)}{q(i \rightsquigarrow x)} \cdot q_{i\to j} &= \frac{q(x \rightsquigarrow j)}{q(j \to x)} \cdot q_{j\to i} \\ \Longrightarrow \underbrace{\frac{1}{\mathcal{Z}} \frac{q(x \rightsquigarrow i)}{q(i \rightsquigarrow x)}}_{p_*^{(i)}} q_{i\to j} &= \underbrace{\frac{1}{\mathcal{Z}} \frac{q(x \rightsquigarrow j)}{q(j \rightsquigarrow x)}}_{p_*^{(j)}} q_{j\to i} \text{ , for all } i,j \in B\backslash\{x\} \text{ and} \\ q_{x\to i}\, q(i \rightsquigarrow x) &\overset{\text{Kolmogorov's criterion}}{=\!=\!=} q_{i\to x}\, q(x \rightsquigarrow i) \\ \Longrightarrow \underbrace{p_*^{(x)}}_{\frac{1}{\mathcal{Z}}} q_{x\to i} &= \underbrace{\frac{1}{\mathcal{Z}} \frac{q(x \rightsquigarrow i)}{q(i \rightsquigarrow x)}}_{p_*^{(i)}} q_{i\to x}. \end{aligned} \tag{2.98}$$

This means that the condition for detailed balance is fulfilled.

□

proof of the continuous-time version.
Replacing the transition probability $q_{i\to j}$ between two states by the corresponding transition rate $\gamma_{i\to j}$ in the above proof for the discrete-time case yields the desired result. □

Theorem 22 (Equivalent characterization of detailed balance).

i) version with no transient states:

For a Markov chain, there exists a positive value $\beta > 0$ (representing the inverse temperature) and an 'energy-vector' $\boldsymbol{\epsilon} \in \mathbb{R}^{|\Omega|}$ such that for all states $i, j \in \Omega$ we have

$$\begin{aligned} \mathrm{e}^{-\beta\,\epsilon_i}\,\gamma_{i\to j} &= \mathrm{e}^{-\beta\,\epsilon_j}\,\gamma_{j\to i} \text{ for continuous times and} \\ \mathrm{e}^{-\beta\,\epsilon_i}\,q_{i\to j} &= \mathrm{e}^{-\beta\,\epsilon_j}\,q_{j\to i} \text{ for discrete times.} \end{aligned} \tag{2.99}$$

if and only if the Markov chain has a stationary solution, which exhibits detailed balance and there are no transient states.

ii) general version:

There exists a stationary solution for the Markov chain, which exhibits detailed balance if and only if there exists a positive value $\beta > 0$ (representing the inverse temperature) and an 'energy-vector' $\boldsymbol{\epsilon} \in (\mathbb{R} \cup \{\infty\})^{|\Omega|} \setminus \left\{\{\infty\}^{|\Omega|}\right\}$ (where some, but not all energies can be infinite) such that for all states $i, j \in \Omega$ we have

$$\begin{aligned} \mathrm{e}^{-\beta\,\epsilon_i}\,\gamma_{i\to j} &= \mathrm{e}^{-\beta\,\epsilon_j}\,\gamma_{j\to i} \text{ for continuous times and} \\ \mathrm{e}^{-\beta\,\epsilon_i}\,q_{i\to j} &= \mathrm{e}^{-\beta\,\epsilon_j}\,q_{j\to i} \text{ for discrete times.} \end{aligned} \tag{2.100}$$

Proof. We restrict ourselves to the continuous-time cases, the proof for discrete times is analogous, when replacing the transition rate by the transition probability.

i) version with no transient state:

If the condition holds, we can define a stationary probability as $p_*^{(i)} := \frac{\mathrm{e}^{-\beta\,\epsilon_i}}{\mathcal{Z}}$, with $\mathcal{Z} := \sum_{j\in\Omega} \mathrm{e}^{-\beta\,\epsilon_j}$. The condition also implies, that a transition vanishes if and only if the corresponding reverse transition vanishes, ruling out all transient states.
Vice versa if we have detailed balance and no transient states, we can set $\mathrm{e}^{-\beta\,\epsilon_i} := p_*^{(i)}$, which is well defined, since all stationary probabilities in minimal absorbing sets are strictly positive (compare theorem 13)

ii) general version:

First, we notice that the condition, that not all energies are infinite is crucial, since otherwise the condition would always be fulfilled. Moreover, the definition $p_*^{(i)} := \frac{\mathrm{e}^{-\beta\,\epsilon_i}}{\mathcal{Z}}$ is well defined, since the partition sum is strictly positive.
Vice versa, the definition $\mathrm{e}^{-\beta\,\epsilon_i} := p_*^{(i)}$ now makes sense even for transient states $i \in B_0$, where the stationary probability vanishes:
$p_*^{(i)} = 0 \Longrightarrow \epsilon_i = \infty$.

We note the existence of a stationary solution with detailed balance does not mean that every stationary solution exhibits detailed balance. A counter example is given in Figure 2.15 below.

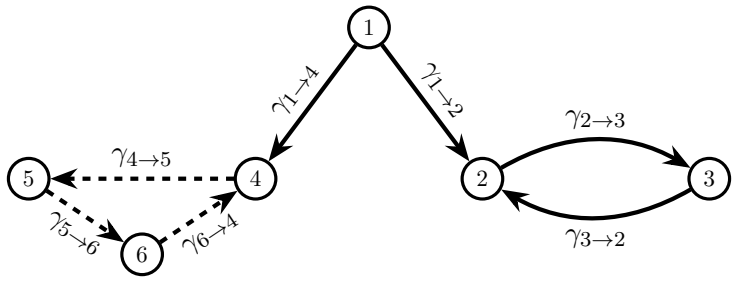


Figure 2.15.: Example of a network, whose set of stationary solutions is given by $\left\{ \alpha \begin{pmatrix} 0 \\ \boldsymbol{p}_{B_1} \\ \mathbf{0}^{(3)} \end{pmatrix} + (1-\alpha) \begin{pmatrix} 0 \\ \mathbf{0}^{(2)} \\ \boldsymbol{p}_{B_2} \end{pmatrix} : \alpha \in [0,1] \right\}$, with
$\boldsymbol{p}_{B_1} = \frac{1}{\gamma_{2\to3}+\gamma_{3\to2}} \begin{pmatrix} \gamma_{3\to2} \\ \gamma_{2\to3} \end{pmatrix}$ and
$\boldsymbol{p}_{B_2} = \frac{1}{\gamma_{5\to6}\,\gamma_{6\to4}+\gamma_{6\to4}\,\gamma_{4\to5}+\gamma_{4\to5}\,\gamma_{5\to6}} \begin{pmatrix} \gamma_{5\to6}\,\gamma_{6\to4} \\ \gamma_{6\to4}\,\gamma_{4\to5} \\ \gamma_{4\to5}\,\gamma_{5\to6} \end{pmatrix}$, but only one of them, namely $\begin{pmatrix} 0 \\ \boldsymbol{p}_{B_1} \\ \mathbf{0}^{(3)} \end{pmatrix}$ exhibits detailed balance.
This is due to the fact that as long as there is a non-vanishing probability mass $\epsilon > 0$ contained in the minimal absorbing set $B_2 = \{4, 5, 6\}$, this probability will 'flow in the circle' $4 \to 5 \to 6 \to 4$ indicated by the dashed links and thus violating Kolmogorov's criterion.

□

2.6.1. The canonical ensemble

When the condition of theorem 22 is fulfilled for all states in a minimal absorbing set $B \in \mathcal{B}$, that is for all $i, j \in B$ there exists a positive number $\beta > 0$ and a real vector $\boldsymbol{\epsilon} \in \mathbb{R}^{|B|}$ such that

$$\mathrm{e}^{-\beta\,\epsilon_i}\,\gamma_{i\to j} = \mathrm{e}^{-\beta\,\epsilon_j}\,\gamma_{j\to i},$$

then the stationary solution of that minimal absorbing set is given by the canonical ensemble, namely

$$\begin{aligned} \boldsymbol{p}_B &= \frac{\mathrm{e}^{-\beta\,\boldsymbol{\epsilon}}}{\mathcal{Z}} \text{ that is} \\ \boldsymbol{p}_{B,b} &= \frac{\mathrm{e}^{-\beta\,\epsilon_b}}{\mathcal{Z}} \text{ with } \mathcal{Z} = \sum_{b\in B} \mathrm{e}^{-\beta\,\epsilon_b} \end{aligned} \tag{2.101}$$

In the special case when all energies are the same (that is all transition rates are identical), the stationary solution is a uniform probability distribution:

$$\boldsymbol{p}_B = \frac{1}{|B|}\Big(\underbrace{1,\ldots,1}_{|B|}\Big). \tag{2.102}$$

Another way to interpret this is the high temperature limit $\beta \to 0^+$.

Proof. Since we know that the entries of the stationary solution are strictly positive, it suffices to show, that $\frac{p_\infty^{(i)}}{p_\infty^{(j)}} = \frac{\mathrm{e}^{-\beta\,\epsilon_i}}{\mathrm{e}^{-\beta\,\epsilon_j}}$.
The idea is to turn an in-tree rooted at state $i \in \Omega$ into an in-tree rooted at some other state $j \in \Omega$ by reversing some of the links on the path $(j \rightsquigarrow i) := j = \alpha(0) \to \cdots \to \alpha(L) = i$, for $L \in \mathbb{N}$, while keeping track of the corresponding Boltzmann factors:

$$\gamma(j \rightsquigarrow i) := \prod_{l=0}^{L-1} \frac{\gamma_{\alpha(l)\to\alpha(l+1)}}{\gamma_{\alpha(l+1)\to\alpha(l)}\, \mathrm{e}^{-\beta\left(\epsilon_{\alpha(l+1)}-\epsilon_{\alpha(l)}\right)}}$$
$$= \prod_{l=0}^{L-1} \gamma_{\alpha(l+1)\to\alpha(l)} \cdot \underbrace{\prod_{l=0}^{L-1} \mathrm{e}^{-\beta\left(\epsilon_{\alpha(l+1)}-\epsilon_{\alpha(l)}\right)}}_{\mathrm{e}^{-\beta\sum_{l=0}^{L-1}\left(\epsilon_{\alpha(l+1)}-\epsilon_{\alpha(l)}\right)}} = \tag{2.103}$$
$$\overset{\substack{\alpha(L)=i \\ \alpha(0)=j}}{=\!=\!=} \gamma(i \rightsquigarrow j) \cdot \epsilon^{-\beta\,(\epsilon_i-\epsilon_j)}.$$

This means, that for all in-trees $T_{\to i}$ rooted at state $i \in \Omega$ and for all in-trees rooted at some other state $j \in \Omega$, we have $\gamma_{T\to i} = \gamma_{T\to j} \cdot \epsilon^{-\beta\,(\epsilon_j-\epsilon_i)}$, which means for the i-th component of the stationary probability vector:

$$p_\infty^{(i)} = \frac{1}{\mathcal{Z}} \sum_{T_{\to i}\in\mathcal{T}(\to i,\mathcal{S})} \underbrace{\gamma_{T\to i}}_{\gamma_{T\to j}\cdot\epsilon^{-\beta\,(\epsilon_i-\epsilon_j)}} = \epsilon^{-\beta\,(\epsilon_i-\epsilon_j)} \cdot \underbrace{\frac{1}{\mathcal{Z}} \sum_{T_{\to i}\in\mathcal{T}(\to j,\mathcal{S})} \gamma_{T\to j}}_{p_\infty(j)} \tag{2.104}$$
$$\Longrightarrow \frac{p_\infty^{(i)}}{p_\infty^{(j)}} = \frac{\mathrm{e}^{-\beta\,\epsilon_i}}{\mathrm{e}^{-\beta\,\epsilon_j}}.$$

This concludes the proof. □

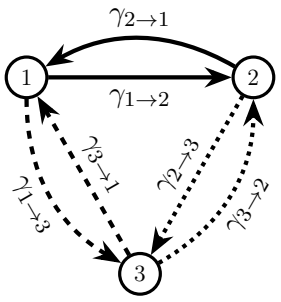


Figure 2.16.: Illustrating the reversal of links of in-trees in a network, corresponding to the canonical ensemble. Reversing some the links results in an in-tree, whose weight has additional Boltzmann factors (compare Equation (2.105)).

$$
\begin{aligned}
\sum_{T_{\to 1}\in\mathcal{T}(\to 1,\mathcal{S})} \gamma_{T_{\to 1}} = \; & \underbrace{\gamma_{2\to 3}}_{\gamma_{3\to 2}\,\mathrm{e}^{-\beta\,(\epsilon_3-\epsilon_2)}} \overbrace{\gamma_{1\to 3}\,\mathrm{e}^{-\beta\,(\epsilon_1-\epsilon_3)}}^{\gamma_{3\to 1}} \\
& + \gamma_{3\to 2} \underbrace{\gamma_{2\to 1}}_{\gamma_{1\to 2}\,\mathrm{e}^{-\beta\,(\epsilon_1-\epsilon_2)}} \\
& + \gamma_{3\to 1} \underbrace{\gamma_{2\to 1}}_{\gamma_{1\to 2}\,\mathrm{e}^{-\beta\,(\epsilon_1-\epsilon_2)}} \\
& = \mathrm{e}^{-\beta\,(\epsilon_1-\epsilon_2)} \left(\sum_{T_{\to 2}\in\mathcal{T}(\to 2,\mathcal{S})} \gamma_{T_{\to 2}} \right) \\
& = \mathrm{e}^{-\beta\,(\epsilon_1-\epsilon_2)} \Big(\gamma_{1\to 3} \underbrace{\gamma_{3\to 2}}_{\gamma_{2\to 3}\,\mathrm{e}^{-\beta\,(\epsilon_2-\epsilon_3)}} \\
& + \underbrace{\gamma_{3\to 2}}_{\gamma_{2\to 3}\,\mathrm{e}^{-\beta\,(\epsilon_2-\epsilon_3)}} \gamma_{1\to 2} \\
& + \overbrace{\gamma_{1\to 3}\,\mathrm{e}^{-\beta\,(\epsilon_1-\epsilon_3)}}^{\gamma_{3\to 1}} \underbrace{\gamma_{1\to 2}}_{\gamma_{2\to 1}\,\mathrm{e}^{-\beta\,(\epsilon_2-\epsilon_1)}} \Big) \\
& = \mathrm{e}^{-\beta\,(\epsilon_1-\epsilon_3)} \left(\sum_{T_{\to 3}\in\mathcal{T}(\to 3,\mathcal{S})} \gamma_{T_{\to 3}} \right).
\end{aligned}
\tag{2.105}
$$

2.6.2. The maximization of entropy

Given a master equation of a strongly connected network as defined in Equation (2.20). The goal of this subsection is to define a time-dependent entropy function and to show that this function is increasing. We follow the arguments given by Kelly [Kel11].
We suppress the dependence of the initial state $\boldsymbol{p}_0$, in the sense that

$$
\begin{aligned}
\boldsymbol{p}(t) &= \boldsymbol{p}(t \,|\, \boldsymbol{p}_0) \text{ and} \\
\boldsymbol{p}_* &:= \boldsymbol{p}_\infty(\boldsymbol{p}_0).
\end{aligned}
$$

We first note, that for a stationary solution $\boldsymbol{p}_*$ we have

$$\begin{aligned} \boldsymbol{p}_* &= \mathrm{e}^{\tau\,\Gamma}\,\boldsymbol{p}_* \text{ and component-wise} \\ p_*^{(i)} &= \sum_{j\in\Omega} \left(\mathrm{e}^{\tau\,\Gamma}\right)_{ij} p_*^{(j)} \end{aligned} \tag{2.106}$$

Hence the following numbers are non-negative and add up to one:

$$\begin{aligned} \frac{\left(\mathrm{e}^{\tau\,\Gamma}\right)_{ij} p_*^{(j)}}{p_*^{(i)}} &\geq 0 \\ \sum_{j\in\Omega} \frac{\left(\mathrm{e}^{\tau\,\Gamma}\right)_{ij} p_*^{(j)}}{p_*^{(i)}} &= 1. \end{aligned} \tag{2.107}$$

Now let $h : \mathbb{R} \to \mathbb{R}$ be any concave function. Then the function defined as

$$H(t) := \sum_{i\in\Omega} h\left(\frac{p^{(i)}(t)}{p_*^{(i)}}\right) \tag{2.108}$$

is strictly increasing:

Proof.

$$\begin{aligned} H(t+\tau) &= \sum_{i\in\Omega} p_*^{(i)}\, h\Big(\frac{\overbrace{p^{(i)}(t+\tau)}^{\sum_{j\in\Omega}\left(\mathrm{e}^{\tau\,\Gamma}\right)_{ij} p^{(j)}(t)}}{p_*^{(i)}}\Big) \\ &= \sum_{i\in\Omega} p_*^{(i)}\, h\left(\sum_{j\in\Omega} \frac{\left(\mathrm{e}^{\tau\,\Gamma}\right)_{ij} p_*^{(j)}}{p_*^{(i)}} \cdot \frac{p^{(j)}(t)}{p_*^{(j)}}\right) \\ &\overset{\text{h concave}}{\underset{(2.107)}{>}} \sum_{i\in\Omega} p_*^{(i)} \sum_{j\in\Omega} \frac{\left(\mathrm{e}^{\tau\,\Gamma}\right)_{ij} p_*^{(j)}}{p_*^{(i)}}\, h\left(\frac{p^{(j)}(t)}{p_*^{(j)}}\right) \\ &= \sum_{j\in\Omega} p_*^{(j)}\, h\left(\frac{p^{(j)}(t)}{p_*^{(j)}}\right) \cdot \underbrace{\sum_{i\in\Omega} \left(\mathrm{e}^{\tau\,\Gamma}\right)_{ij}}_{1} \\ &= H(t). \end{aligned} \tag{2.109}$$

□

When we choose $h(x) := -x \ln(x)$ as a special case of a concave function, then the function

$$H_S(t) = -\sum_{i\in\Omega} p^{(i)}(t) \ln\left(\frac{p^{(i)}(t)}{p_*^{(i)}}\right) =: -\mathrm{KL}\big((p^{(i)}(t))_{i\in\Omega} \,\|\, (p_*^{(i)})_{i\in\Omega}\big) \tag{2.110}$$

becomes the negative of the Kullback-Leibler divergence (also called relative entropy). This relative entropy is convergent (since it is increasing and bounded from above), non-negative and it vanishes if and only if the two probability distributions coincide, that is when the stationary state is reached $\boldsymbol{p}(t) = \boldsymbol{p}_*$ (which is never the case for a finite time if $\boldsymbol{p}_0 \neq \boldsymbol{p}_*$).
This means from a physical point of view that a strongly connected system approaching equilibrium is closely related to the fact that its relative entropy is maximized, where the maximum of entropy of the system corresponds to its equilibrium state.

3. The Lindblad equation

3.1. Derivation of the Lindblad equation for finite dimensions

In this Section we give a short derivation of the Lindblad equation following the arguments from [BP02].
Given the initial state of a quantum system plus environment is a product state $\boldsymbol{\rho}(t = 0) = \boldsymbol{\rho}_S(t = 0) \otimes \boldsymbol{\rho}_B$, where $\boldsymbol{\rho}_S \in \mathbb{C}^{N\times N}$ is the state of the system and $\boldsymbol{\rho}_B \in \mathbb{C}^{D_B\times D_B}$ is the equilibrium state of the bath, which is constant over time.
The quantum state at a time $t > 0$ is given by first computing the time evolution of the combined system (system + bath) and then taking the average over all the states representing the environment (this is sometime called 'tracing' over the bath, due to the fact that a quantum mechanical average of an observable is computed by taking a trace of the density matrix times the observable). This results in the so-called Kraus operator representation:

$$\begin{aligned}\boldsymbol{\rho}_S(t) &= \mathrm{Tr}_B\left[\mathcal{U}(t)\left(\boldsymbol{\rho}_S(0)\otimes\boldsymbol{\rho}_B\right)\mathcal{U}(t)^\dagger\right] = \\ &\overset{\text{Kraus}}{=\!=} \sum_{i,j=1}^{N^2} c_{i,j}(t)\, F_i\, \boldsymbol{\rho}_S(0)\, (F_j)^\dagger,\end{aligned} \tag{3.1}$$

where $(F_i)_{i\in\{1,\dots,N^2\}}$ is an orthonormal basis of $\mathbb{C}^{N\times N}$ with $F_{N^2} := \frac{1}{\sqrt{N}}\mathbb{1}_N$ and $(c_{ij})_{i,j\in\{1,\dots,N^2\}}$ is a positive-definite coefficient matrix over the complex numbers. In the following we will suppress the index 'S' of the density matrix of the system.
The generator can then be computed via

$$\begin{aligned}\mathcal{L}(\boldsymbol{\rho}(t)) &= \frac{\mathrm{d}}{\mathrm{d}\,t}\boldsymbol{\rho}(t)\Big|_{t=0} = \lim_{\epsilon\to 0^+}\frac{\boldsymbol{\rho}(\epsilon)-\boldsymbol{\rho}(0)}{\epsilon} \\ &= -i[H,\boldsymbol{\rho}(t)] + \sum_{i,j=1}^{N-1} a_{ij} F_i\,\boldsymbol{\rho}(t)\,F_j^\dagger + \left\{\underbrace{\left(\frac{a_{N^2,N^2}}{2}\mathbb{1}_N + \frac{F^\dagger+F}{2}\right)}_{G}, \boldsymbol{\rho}\right\},\end{aligned} \tag{3.2}$$

with $a_{ij} := \frac{\mathrm{d}}{\mathrm{d}t} c_{ij}(t)\Big|_{t=0} = c'_{ij}(t=0)$. Since the trace of $\mathcal{L}(\boldsymbol{\rho}(t))$ must vanish, we have:

$$0 \overset{!}{=} \mathrm{Tr}[\mathcal{L}(\boldsymbol{\rho}(t))] = \mathrm{Tr}\left[\underbrace{\left(2\,G + \sum_{i,j=1}^{N-1} a_{ij} F_j^\dagger F_i\right)}_{=0}, \boldsymbol{\rho}(t)\right], \text{ which implies}$$
$$G = \frac{1}{2} \sum_{i,j=1}^{N^2-1} a_{ij} F_j^\dagger F_i. \tag{3.3}$$

After diagonalizing the coefficient matrix $(a_{ij})_{i,j\in\{1,\dots,N^2-1\}}$ with its eigenvalues $\gamma_k \geq 0$, we can write the generator as follows:

$$\mathcal{L}(\boldsymbol{\rho}(t)) = -i[H, \boldsymbol{\rho}(t)] + \sum_{k=1}^{N^2-1} \gamma_k \left(V_k\, \boldsymbol{\rho}(t)\, V_k^\dagger + \frac{1}{2}\Big\{V_k^\dagger V_k,\, \boldsymbol{\rho}(t)\Big\}\right). \tag{3.4}$$

For more details, see Section B.6.

3.2. The concept of an unravelling of the Lindblad equation

Unravellings are ensembles of stochastic quantum trajectories that are equivalent to the Lindblad equation in the sense that the state of the system $\boldsymbol{\rho}(t)$ at a time $t \geq 0$ is given by the average over all possible trajectories,

$$\boldsymbol{\rho}(t) = \langle \Theta(t,\omega) \rangle_{\omega\in\mathcal{U}} := \int_{\mathcal{U}} \Theta(t,\omega)\; \mathrm{d}\,\mathcal{P}(\omega). \tag{3.5}$$

The parameter $\omega \in \mathcal{U}$ labels the quantum trajectories $\mathbb{R}_{\geq 0} \ni\ t \mapsto \Theta(t,\omega) \in \{\text{density matrices}\}$ (see figure (3.1) for an illustration).
In particular, the expectation values of any quantum mechanical observable A, given the system is in the state $\boldsymbol{\rho}$, can be computed from these unravellings, according to

$$\mathcal{E}[A \,|\, \boldsymbol{\rho}] = \mathrm{Tr}[\boldsymbol{\rho}\, A] = \int_{\mathcal{U}} \mathrm{Tr}[\Theta(\omega)\, A]\; \mathrm{d}\,\mathcal{P}(\omega). \tag{3.6}$$

Unravellings provide different perspectives on the subject of time evolution of open quantum systems and can yield new insights since they are tackled by a different set of mathematical tools: Instead of linear differential equations one has the theory of Markov

chains, piecewise deterministic processes, and stochastic differential equations at hand [BP02].
Moreover, unravellings are often superior when it comes to numerical simulations of specific models, as a ket-state unravelling (defined in Section 3.3.4) requires less memory storage (linear with the dimension of the system, instead of quadratic [Lid19]).
While unravellings are usually done in terms of ket-states, we want to focus on the so called mixed-state unravellings, which assume values in the set of density matrices [BP02]. Among the different possible types of unravellings, we will use the quantum jump unravelling. Other types of unravellings are for example the quantum state diffusion model (see [GP92]), which relies on stochastic calculus. The main difference between these unravellings is that a trajectory from quantum state diffusion is continuous in time, in contrast to the quantum jumps coming from the quantum jump unravelling. One can say that quantum state diffusion relies on 'infinitesimal small' jumps happening at every time point, whereas the quantum jump model focuses on larger jumps, whose frequency is determined by some stochastic process.
For a given Hamiltonian H and a given finite set of Lindblad operators $\{V_k : k \in I\}$, the quantum jump unravelling is a piecewise deterministic process (see [BP02; Lid19]) where the continuous, deterministic time evolution is interrupted by discontinuous quantum jumps at times t_n, where a Lindblad operator V_{π_n} is applied.

3.3. The quantum jump unravelling: version for density matrices

The quantum jump unravelling is a special Piecewise-deterministic Markov process (PDMP), which takes values in the density matrices. PDMP have first been introduced by Mark H. A. Davis in [Dav84] and show both deterministic, as well as stochastic features. A typical path contains discontinuities (so-called 'jumps'), but between these jump times, the time evolution is deterministic and determined by a differential equation (see Figure 3.1 for an illustration). It is completely determined by the evolution equation, the rate of the jumps and the transition probabilities [Dav84].

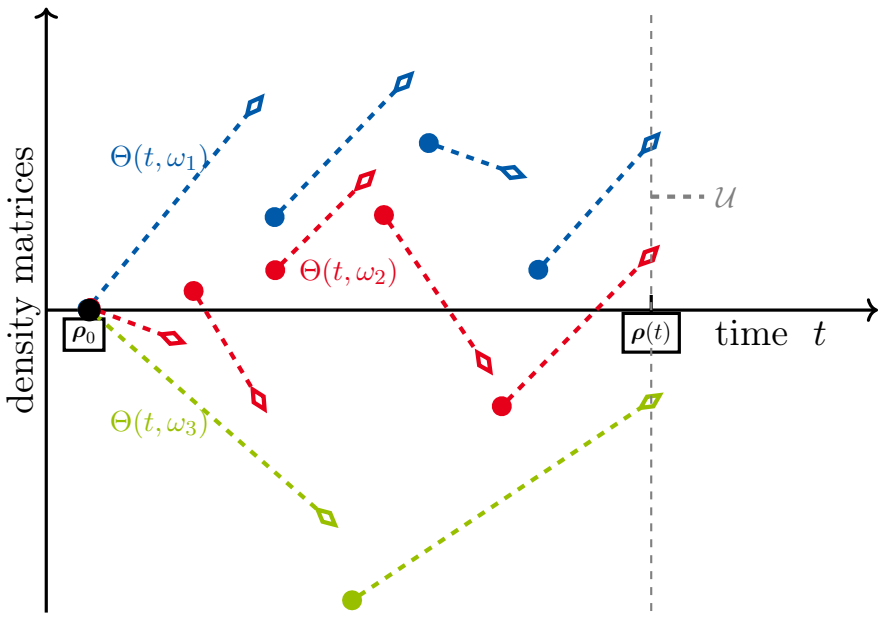


Figure 3.1.: Illustration of an ensemble of quantum trajectories, originating from the quantum jump unravelling, which is a special type of piecewise deterministic processes interrupted by quantum jumps. The average of the trajectories at a given time $t \geq 0$ yields the density matrix $\boldsymbol{\rho}(t)$.

So a trajectory of a quantum jump unravelling is completely determined by the sequence of times $(t_n)_{n\in\mathbb{N}} \subset \mathbb{R}^{\mathbb{N}}$ (which record when a quantum jump takes place) and the sequence $(\pi_n)_{n\in\mathbb{N}} \in \{V_k : k \in I\}^{\mathbb{N}}$ of indices of Lindblad operators (which records which operator has been applied at time t_n). We define $\mathcal{U}$ to be the set of all possible quantum trajectories, $\tau_n := t_{n+1} - t_n$ to be the length of the time interval between two jumps and $\Theta_n := \Theta(t_n)$ the state of the unravelling at time t_n.
We use with $\Theta(t,\omega)$ a different symbol for a state in the quantum trajectory $\omega \in \mathcal{U}$, to distinguish it from the solution $\boldsymbol{\rho}(t)$ of the Lindblad Equation (1.2) at time t_n, which is the average over all quantum trajectories, $\boldsymbol{\rho}(t) = \langle \Theta(\omega,t)\rangle_{\omega\in\mathcal{U}}$.
We point out that all the sequences of time events $(t_n)_{n\in\mathbb{N}}$, waiting times $(\tau_n)_{n\in\mathbb{N}}$ and quantum states $(\Theta_n)_{n\in\mathbb{N}}$ depend on a specific unravelling $\omega \in \mathcal{U}$ (which we will suppress in the following):

$$
\begin{aligned}
t_n &:= t_n(\omega) \\
\tau_n &:= \tau_n(\omega) \\
\Theta_n &:= \Theta_n(\omega).
\end{aligned}
\tag{3.7}
$$

Definition 23 (The operators for time evolution quantum jumps).

To avoid unnecessarily complicated notation, we denote by U_t the (non-unitary) time evolution according to the conditional Hamiltonian H_c, and by J_k the quantum jump operator, irrespective of the nature of the quantum state, be it a density matrix or a ket state. Their action on these two types of quantum states is thus given by

$$
\begin{aligned}
&U_t, J_k : \mathbb{C}^{N\times N} \to \mathbb{C}^{N\times N} && U_t, J_k : \mathbb{C}^{N} \to \mathbb{C}^{N} \\
&U_t(\Theta) = \mathsf{e}^{-i\,H_c\,t}\,\Theta\,\mathsf{e}^{i\,H_c^\dagger\,t} && U_t(\psi) = \mathsf{e}^{-i\,H_c\,t}\,\psi \\
&J_k(\Theta) = V_k\,\Theta\,V_k^\dagger && J_k(\psi) = V_k\,\psi
\end{aligned}
\tag{3.8}
$$

During the time interval $[t_n, t_{n+1})$ between two quantum jumps there is a deterministic time evolution $\Theta(t) = \frac{U_{t-t_n}(\Theta_n)}{\mathrm{Tr}[\dots]}$, and directly after a quantum jump induced by the Lindblad operator V_k the state of the quantum trajectory is given by $\Theta(t_n) = \frac{J_k(\Theta)}{\mathrm{Tr}[\dots]}$ $\Big($ we write $\frac{A}{\mathrm{Tr}[\dots]}$ instead of $\frac{A}{\mathrm{Tr}[A]}$ for any matrix $A\Big)$.

Definition 24 (The waiting time distribution function).

Given that the quantum trajectory is in state Θ, the waiting time t for the next jump is distributed according to

$$
\begin{aligned}
f(t\,|\,\Theta) &:= -\frac{d}{dt}\mathrm{Tr}\left[\underbrace{U_t(\Theta)}_{\mathrm{e}^{-i\,H_c\,t}\,\Theta\,\mathrm{e}^{i\,H_c^\dagger\,t}}\right] = -\mathrm{Tr}\left[\underbrace{(-i\,H_c)}_{-i\,H-\frac{1}{2}\,\Lambda}\,U_t(\Theta) + U_t(\Theta)\,\underbrace{(i\,H_c^\dagger)}_{i\,H-\frac{1}{2}\,\Lambda}\right] = \\
&= \underbrace{i\,\mathrm{Tr}\left[H\,U_t(\Theta)\right] - i\,\mathrm{Tr}\left[U_t(\Theta)\,H\right]}_{0} + \underbrace{\frac{1}{2}\mathrm{Tr}\left[\Lambda\,U_t(\Theta)\right] + \frac{1}{2}\mathrm{Tr}\left[U_t(\Theta)\,\Lambda\right]}_{\mathrm{Tr}[\Lambda\,U_t(\Theta)]} = \\
&= \mathrm{Tr}[\Lambda\,U_t(\Theta)] \overset{\Lambda=\sum_{k\in I}\gamma_k\,V_k^\dagger\,V_k}{=\!=\!=\!=\!=\!=\!=\!=\!=} \sum_{k\in I}\gamma_k\,\mathrm{Tr}\left[\underbrace{V_k\,\mathrm{e}^{-i\,H_c\,t}\,\Theta\,\mathrm{e}^{i\,H_c^\dagger\,t}\,V_k^\dagger}_{J_k\circ U_t(\Theta)}\right] = \\
&= \sum_{k\in I}\underbrace{\gamma_k\,\mathrm{Tr}\left[J_k\circ U_t(\Theta)\right]}_{f^{(k)}(t\,|\,\Theta)} =: \sum_{k\in I} f^{(k)}(t\,|\,\Theta).
\end{aligned}
\tag{3.9}
$$

with

$$
f^{(k)}(t\,|\,\Theta) := \gamma_k\,J_k\circ U_t(\Theta) = \gamma_k\,V_k\,\mathrm{e}^{-i\,H_c\,t}\,\Theta\,\mathrm{e}^{i\,H_c^\dagger\,t}\,V_k^\dagger \quad \text{(see [BP02])}. \tag{3.10}
$$

Note that $f(\cdot\,|\,\Theta)$ need not be a probability distribution since $\int_{\mathbb{R}_{\geq 0}} f(t\,|\,\Theta)\;\mathrm{d}\,t < 1$ is possible. In this case, we call Θ a possible trapping state, which will be discussed below.

3.3.1. The algorithm for an unravelling

The algorithm for obtaining an unravelling is the following (see [Lid19] or[BP02] for a detailed proof):

1) Choose an initial state $\boldsymbol{\rho}_0$ and set $\Theta_0 := \boldsymbol{\rho}_0 = \boldsymbol{\rho}(t_0)$.

2) Given that the system at time t_n is in state Θ_n, we calculate the waiting time τ_n to the next jump according to the following considerations: The probability for τ_n to lie in the interval $[0, T)$ is given by

$$
\mathcal{P}\left(\tau_n \in [0,T)\,|\,\Theta_n\right) = \int_0^T \underbrace{f(t\,|\,\Theta_n)}_{-\frac{d}{dt}\mathrm{Tr}[U_t(\Theta_n)]}\;\mathrm{d}\,t = 1 - \mathrm{Tr}\left[U_T(\Theta_n)\right]. \tag{3.11}
$$

One way of determining the waiting time is the so-called inversion method (see: [BP02]): We draw a random number η_n from the uniform distribution of the interval $[0, 1]$ and set

$$\tau_n := \begin{cases} f^{-1}(\eta_n \,|\, \Theta_n) & , \text{ if } \eta_n \in \text{image } \big(f(\cdot \,|\, \Theta_n)\big) \\ \infty & , \text{ otherwise} \end{cases} \tag{3.12}$$

An illustration of this procedure is given in Figure 3.2.

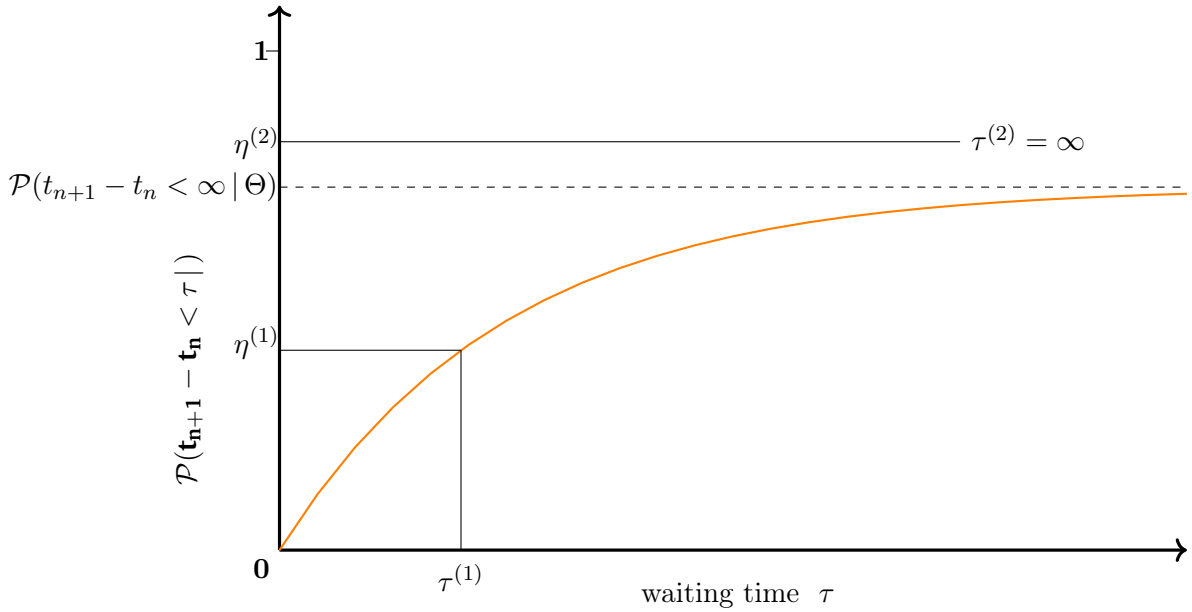


Figure 3.2.: Determining the waiting time $\tau \in \mathbb{R}_{\geq 0} \cup \{\infty\}$, given a state Θ: Choose a random variable η uniformly distributed over the interval $[0, 1]$ and determine the positive number τ that satisfies $f(\tau \,|\, \Theta) = \eta$. If no such number exists, set $\tau = \infty$. The two cases are indicated in the figure: When the random number is $\eta^{(1)}$, the corresponding waiting time is $\tau^{(1)}$; when it is $\eta^{(2)}$ the waiting time is infinite, $\tau^{(2)} = \infty$.

Now we have to distinguish two cases:

i) When $\lim_{t\to\infty} \mathrm{Tr}[U_t(\Theta_n)] = 0$, then the quantum jump will occur at a finite time, $\mathcal{P}(\tau_n < \infty \,|\, \Theta_n) = 1$.

ii) When $\lim_{t\to\infty} \mathrm{Tr}[U_t(\Theta_n)] \in (0,1]$, then there is a nonzero chance that the process no longer jumps, i.e. the waiting time is infinite ($\mathcal{P}(\tau_n = \infty \,|\, \Theta_n) > 0$) . In this case, the state of the quantum trajectory after the time t_n is given by $\frac{U_{t-t_n}(\Theta_n)}{\mathrm{Tr}[\dots]}$, and we call $\Theta_{\mathrm{trap}} := \Theta_n$ a possible trapping state, since it is possible that Θ_{trap} 'traps' the discrete quantum trajectory in the sense that the sequence of quantum states is finite, with Θ_{trap} being the last of these states.

We point out, that only the discrete quantum trajectory 'stops', in the sense that no more quantum jumps take place. The time evolution keeps evolving according to

$$\Theta_{\mathrm{trap}}(t) \xlongequal{t>t_{\mathrm{trap}}} \frac{U_{t-t_{\mathrm{trap}}}(\Theta_{\mathrm{trap}})}{\mathrm{Tr}[\dots]}$$

3) If the waiting time is finite ($\tau_n < \infty$) , choose the operator V_{π_n} that is applied at time t_n according to

$$\mathcal{P}\left(\pi_n = k \,|\, \Theta_n, \tau = \tau_n\right) = \frac{f^{(k)}(\tau_n \,|\, \Theta_n)}{\sum_{j\in J} f^{(j)}(\tau_n \,|\, \Theta_n)}, \tag{3.13}$$

with the $f^{(k)}(\cdot \,|\, \Theta)$ defined in Equation (3.10). The next state is then given by $\Theta_{n+1} = \frac{J_k\left(U_{\tau_n}(\Theta_n)\right)}{\mathrm{Tr}[\dots]}$.

Then for a fixed unravelling $\omega \in \mathcal{U}$, for positive times $t_n \in \mathbb{R}_{\geq 0}$, and (possibly infinite) times $t_{n+1} \in \mathbb{R}_{>0} \cup \{\infty\}$, we have

$$\begin{aligned} \text{for } t \in [t_n, t_{n+1}) : \Theta(t, \omega \,|\, \boldsymbol{\rho}_0) = \frac{U_{t-t_n}(\Theta_n)}{\mathrm{Tr}[\dots]} = \frac{U_{t-t_n} \circ J_{\pi_n} \circ U_{\tau_{n-1}} \circ \cdots \circ J_{\pi_1} \circ U_{\tau_0}(\boldsymbol{\rho}_0)}{\mathrm{Tr}[\dots]} \\ \Theta_n := \Theta(t_n, \omega) \end{aligned} \tag{3.14}$$

Depending on whether there are finitely many of infinitely many quantum jumps, the set of states in the quantum trajectory varies:

i) infinitely many jumps

$$\Omega_{\boldsymbol{\rho}_0}(\omega) := \{\boldsymbol{\rho}_0 = \Theta_0\} \cup \left\{ \frac{J_{\pi_n} \circ U_{\tau_{n-1}} \circ \cdots \circ J_{\pi_1} \circ U_{\tau_0}(\boldsymbol{\rho}_0)}{\mathrm{Tr}[\dots]} \;:\; n \in \mathbb{N},\, \omega = (t_i, \pi_i)_{i\in\mathbb{N}} \right\} = \\ = \{\Theta_n(\omega) \;:\; n \in \mathbb{N}_0,\, \omega = (t_i, \pi_i)_{i\in\mathbb{N}}\}\,, \tag{3.15}$$

and the sequence $(\Theta_n)_{n\in\mathbb{N}} = (\Theta(t_n, \omega \,|\, \boldsymbol{\rho}_0))_{n\in\mathbb{N}}$ is called the discrete quantum trajectory.

ii) finitely many jumps
Let $l \in \mathbb{N}$ be the number of jumps, then the last state must be a possible trapping state

$$\Theta_{\mathrm{trap}} := \frac{J_{\pi_l} \circ U_{t_l - t_{l-1}} \circ \ldots \circ J_{\pi_1} \circ U_{t_1}(\boldsymbol{\rho}_0)}{\mathrm{Tr}[\dots]}.$$

and the set of states in the quantum trajectory is given by

$$\Omega_{\boldsymbol{\rho}_0}(\omega) := \{\boldsymbol{\rho}_0 = \Theta_0\} \cup \left\{ \frac{J_{\pi_n} \circ U_{\tau_{n-1}} \circ \cdots \circ J_{\pi_1} \circ U_{\tau_0}(\boldsymbol{\rho}_0)}{\mathrm{Tr}[\dots]} \;:\; n \in \mathbb{N}_{\leq l},\, \omega = (t_i, \pi_i)_{i\leq l} \right\} \\ \cup \{\langle \Theta_{\mathrm{trap}}(t) \rangle_{t\geq 0}\}\,, \tag{3.16}$$

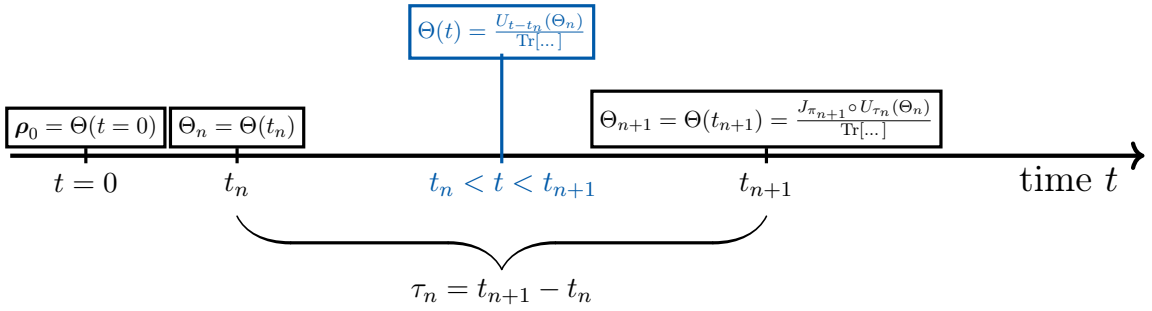


Figure 3.3.: Illustration of a quantum trajectory

3.3.2. The set of states Ω

In a classical Markov chain with finite state space, the states could be associated to the unit vectors: $\Omega = \{1, \ldots, |\Omega|\} \widehat{=} \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_{|\Omega|}\}$, with

$$\boldsymbol{e}_i = (\underbrace{0, \ldots, 0}_{i-1}, 1, \underbrace{0, \ldots, 0}_{|\Omega|-i}).$$

Now the system consists of both the states a quantum trajectory attains directly after the quantum jump as well as the time-averages of all possible trapping states that is

$$\Omega_{\boldsymbol{\rho}_0} = \bigcup_{\omega \in \mathcal{U}} \Omega_{\boldsymbol{\rho}_0}(\omega) \tag{3.17}$$

In order to apply the methods of the previous section, we again assume, that this state space is finite, $|\Omega_{\boldsymbol{\rho}_0}| < \infty$. Some important examples will be considered later.

3.3.3. Two examples of discrete quantum trajectories

In this Section we want to illustrate the abstract procedure given above by two concrete examples, which will come up again in Section 3.7.

Classical jump process between two states with two possible discrete quantum trajectories.

When we set the Hamiltonian to be trivial and the two Lindblad operators to be 'classical' jump operators, that is

$$H \propto \mathbb{1},\ V_1 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},\ V_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \tag{3.18}$$

we get the following two possible discrete quantum trajectories:

$$(\Theta_n)_{n\in\mathbb{N}} = \begin{cases} (\boldsymbol{\rho}_0, \Theta_{s_2}, \Theta_{s_3}, \Theta_{s_2}, \Theta_{s_3} \ldots,) \text{ with probability } \boldsymbol{\rho}_{11}(0) \text{ and} \\ (\boldsymbol{\rho}_0, \Theta_{s_3}, \Theta_{s_2}, \Theta_{s_3}, \Theta_{s_2} \ldots,) \text{ with probability } \boldsymbol{\rho}_{22}(0), \end{cases} \tag{3.19}$$

with the two states

$$\begin{aligned} \Theta_{s_2} &:= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \Theta_{s_3} := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \text{ and the initial state} \\ \boldsymbol{\rho}_0 &= \begin{pmatrix} \boldsymbol{\rho}_{11}(0) & \boldsymbol{\rho}_{12}(0) \\ \boldsymbol{\rho}_{21}(0) & \boldsymbol{\rho}_{22}(0) \end{pmatrix}. \end{aligned} \tag{3.20}$$

The set of all possible states Ω is given by

$$\Omega = \{\boldsymbol{\rho}_0, \Theta_{s_2}, \Theta_{s_3}\}.$$

Simple case of a possible trapping state

In case of a trivial Hamiltonian and a projection operator as a Lindblad operator, that is

$$H \propto \mathbb{1}, V_1 \quad = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \tag{3.21}$$

the initial state is a possible trapping state. This means (in this case) that the initial state could, but need not be the last state in the discrete quantum trajectory.
This leads to the following two possible discrete quantum trajectories:

$$(\Theta_n)_{n\in\mathbb{N}} = \begin{cases} (\boldsymbol{\rho}_0, \Theta_{s_2}, \Theta_{s_2}, \Theta_{s_2}, \dots,) & \text{with probability } \boldsymbol{\rho}_{11}(0) \text{ and} \\ (\boldsymbol{\rho}_0) & \text{with probability } \boldsymbol{\rho}_{22}(0), \end{cases}, \tag{3.22}$$

where Θ_{s_2} is defined as

$$\Theta_{s_2} := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \tag{3.23}$$

The set of all possible states Ω is given by

$$\Omega = \{\boldsymbol{\rho}_0, \Theta_{s_2}\}.$$

3.3.4. Connection between unravellings for *pure* quantum states and *ket*-states

If some state in the quantum trajectory is a pure state ($\Theta_n = P_{\psi_n} := |\psi_n\rangle \langle\psi_n| \mathrel{\widehat{=}} \psi_n$), the unravelling reduces to a ket-state unravelling from then on, and equations (3.9), (3.11), (3.13), and (3.14) become

$$f(t \,|\, \psi) = -\mathrm{Tr}[U_t(P_\psi)] = -\mathrm{Tr}[P_{U_t\,(\psi)}] = -\|U_t\,\psi\|^2 \tag{3.9*}$$

$$\mathcal{P}\left(\tau_n \in [0,T) \,|\, \psi_n\right) = \int_0^T \mathrm{d}\,t \underbrace{f(t \,|\, \psi)}_{-\frac{d}{dt}\|U_T(\psi_n)\|^2} = 1 - \|U_T(\psi_n)\|^2 \tag{3.11*}$$

$$\mathcal{P}\left(\pi_n = k \,|\, \psi_n,\, \tau = \tau_n\right) = \frac{f^{(k)}(\tau_n \,|\, \psi_n)}{\sum\limits_{j\in J} f^{(j)}(\tau_n \,|\, \psi_n)}. \tag{3.13*}$$

The states occurring during a ket-state unravelling $\omega \in \mathcal{U}$ are given by

$$\begin{aligned} \text{for } t \in [t_n, t_{n+1}) : \Psi(t, \omega \,|\, \psi_0) &= \frac{U_{t-t_n}\left(\Psi_n\right)}{\|\dots\|} \\ &= \frac{U_{t-t_n} \circ J_{\pi_{n+1}} \circ U_{\tau_n} \circ \dots \circ J_{\pi_1} \circ U_{\tau_0}\left(\Psi_0\right)}{\|\dots\|} \\ \Psi_n &:= \Psi(t_n, \omega) \end{aligned} \tag{3.14*}$$

The transition to a ket-state unravelling occurs for instance when a Lindblad operator of rank one ($V = |\psi\rangle\langle\varphi|$) is applied at some point in the quantum trajectory,

$$\Theta_{n+1} = \frac{|\psi\rangle\langle\varphi|\; U_{\tau_n}(\Theta_n)\; |\varphi\rangle\langle\psi|}{\operatorname{Tr}[]} = |\psi\rangle\langle\psi| \mathrel{\widehat{=}} \psi. \tag{3.24}$$

So, in some sense, Lindblad operators of rank one are ‘quasi-classical’.

3.4. The path to the stationary solution

3.4.1. How to ‘guess’ the stationary solution

In Chapter 2 we have computed the stationary solution of a classical Markov chain to

$$\begin{aligned} \boldsymbol{p}_\infty(\boldsymbol{p}_0) &\overset{(2.83)}{=\!=} \sum_{B\in\mathcal{B}} \mathcal{P}\left(\boldsymbol{p}_0 \rightsquigarrow B\right) \underbrace{\boldsymbol{p}_\infty(\boldsymbol{p}_0 \in B)}_{\sum_{i\in\Omega} p_\infty^{(i)}(\boldsymbol{p}_0\in B)\,\boldsymbol{e}_i} \\ &\overset{(2.73)}{=\!=} \sum_{B\in\mathcal{B}} \mathcal{P}\left(\boldsymbol{p}_0 \rightsquigarrow B\right) \sum_{i\in\Omega} \underbrace{p_\infty^{(i)}(\boldsymbol{p}_0 \in B)}_{\frac{q_\infty^{(i)}(\boldsymbol{p}_0\in B)\,\overline{\tau_i}}{\mathcal{N}}} \boldsymbol{e}_i \\ &= \sum_{B\in\mathcal{B}} \mathcal{P}\left(\boldsymbol{p}_0 \rightsquigarrow B\right) \sum_{i\in\Omega} \frac{q_\infty^{(i)}(\boldsymbol{p}_0 \in B)\,\overline{\tau_i}}{\mathcal{N}}\, \boldsymbol{e}_i, \end{aligned} \tag{3.25}$$

where the states correspond to a unit vector (for $s \in \Omega$ we have $s \mathrel{\widehat{=}} \boldsymbol{e}_s = (0, \dots, 1, \dots, 0)$) and the system is constant between two jumps (compare Figure 3.4a).

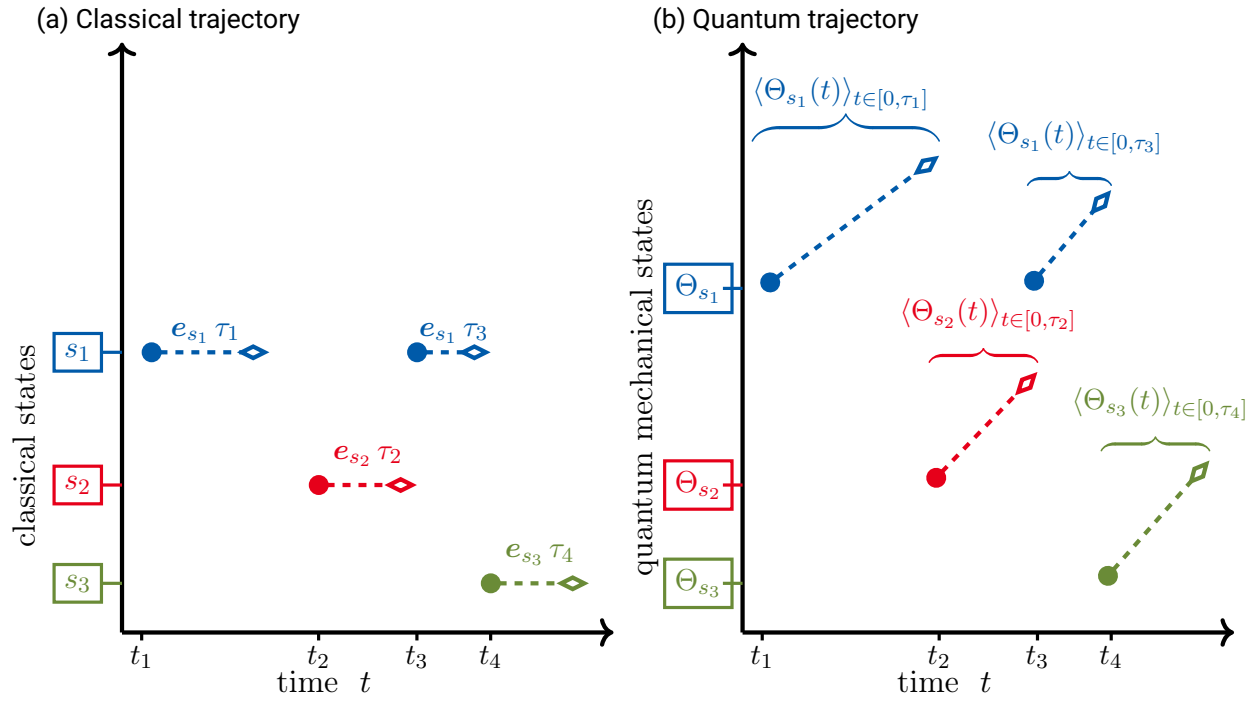


Figure 3.4.: Illustrating the difference between a classical trajectory and a trajectory coming from a quantum jump unravelling: While the system is constant between two time jumps and the states are associated to unit vectors, the states in a quantum trajectory are density matrices and the system evolves between two jumps according to the conditional time evolution $t \mapsto \frac{U_t(\Theta)}{\mathsf{Tr}[\dots]}$.

Things are more complicated when it comes to the quantum jump unravelling. Here, the states correspond to density matrices (for $s \in \Omega_{\boldsymbol{\rho}_0}$ we have $s \mathrel{\widehat{=}} \boldsymbol{\Theta}_s \in \mathbb{C}^{N\times N}$) and the system evolves between two jumps according to the conditional Hamiltonian H_c (compare Figure 3.4b).
Since the states are time dependent, we can no longer associate a state with $\Theta_s(t)$ for some $t \in [0,\tau]$, but we can do so with the time average: $s \mathrel{\widehat{=}} \langle\Theta_s(t)\rangle_{t\in[0,\tau_s]} = \int_0^{\tau_s} \Theta_s(t)\,\mathrm{d}\,t$, where the time average is taken for the duration of the waiting time. Since this waiting time τ_s is a random variable, with the probability density function $\tau_s \sim f(\bullet \,|\, \Theta_s)$, we get the following analogy:

$$\overline{\boldsymbol{\tau}_s} \cdot \boldsymbol{e}_s \mapsto \langle \boldsymbol{\tau} \cdot \langle \Theta_s(t) \rangle_{t\in[0,\boldsymbol{\tau}]} \rangle_{\boldsymbol{\tau} \sim f(\cdot \,|\, \Theta_s)}$$

In the special case of a time independent state we get the same result as for a classical Markov chain, namely

$$\underbrace{\langle \langle \Theta_s(t) \rangle_{t\in[0,\boldsymbol{\tau}]} \rangle_{\boldsymbol{\tau} \sim f(\cdot \,|\, \Theta_s)}}_{\Theta_s \, \boldsymbol{\tau}} = \Theta_s \underbrace{\langle \boldsymbol{\tau} \rangle_{\boldsymbol{\tau} \sim f(\cdot \,|\, \Theta_s)}}_{\overline{\tau_s}} = \Theta_s \cdot \overline{\tau_s}. \tag{3.26}$$

3.4.2. Outlining the procedure

Before going into the details, we are going to outline our approach:
The stationary solution $\boldsymbol{\rho}_\infty(\boldsymbol{\rho}_0)$ for an initial state $\boldsymbol{\rho}_0$ is obtained by taking the time average of expression (3.5) and interchanging the time average with the ensemble average,

$$\boldsymbol{\rho}_\infty(\boldsymbol{\rho}_0) = \langle \Theta_\infty(\omega \,|\, \boldsymbol{\rho}_0) \rangle_{\omega \in \mathcal{U}} = \int_{\mathcal{U}} \Theta_\infty(\omega \,|\, \boldsymbol{\rho}_0) \; \mathrm{d}\, \mathcal{P}(\omega \,|\, \boldsymbol{\rho}_0), \tag{3.27}$$

where the time average for a single trajectory is known to exist (see [KM04]) and is given by

$$\Theta_\infty(\omega \,|\, \boldsymbol{\rho}_0) := \langle \Theta(t, \omega \,|\, \boldsymbol{\rho}_0) \rangle_{t\geq 0} \overset{\text{[KM04]}}{=\!=} \lim_{T\to\infty} \frac{1}{T} \int_0^T \Theta(t, \omega \,|\, \boldsymbol{\rho}_0) \; \mathrm{d}\, t. \tag{3.28}$$

As we will see in Section 3.5, the set $\mathcal{U}$ of all possible unravellings, contains (under some assumptions) only a finite number of different long-term behaviors, which themselves do not depend on the initial state $\boldsymbol{\rho}_0$:

$$\begin{aligned} \mathcal{U} &= \dot{\bigcup_{i=1}^{n}} \mathcal{U}_i \\ \Theta_\infty(\omega \,|\, \boldsymbol{\rho}_0) &= \sum_{i=1}^{n} \Theta_\infty(\mathcal{U}_i) \, 1_{\mathcal{U}_i}(\omega). \end{aligned} \tag{3.29}$$

This will give us the following expression for the stationary state:

$$\begin{aligned}\boldsymbol{\rho}_\infty(\boldsymbol{\rho}_0) &= \int_{\mathcal{U}} \underbrace{\Theta_\infty(\omega \,|\, \boldsymbol{\rho}_0)}_{\sum_{i=1}^{n} \Theta_\infty(\mathcal{U}_i)\, 1_{\mathcal{U}_i}(\omega)} \; \mathrm{d}\,\mathcal{P}(\omega \,|\, \boldsymbol{\rho}_0) \\ &= \sum_{i=1}^{n} \Theta_\infty(\mathcal{U}_i) \underbrace{\int_{\mathcal{U}_i} \mathrm{d}\,\mathcal{P}(\omega \,|\, \boldsymbol{\rho}_0)}_{\mathcal{P}(\mathcal{U}_i \,|\, \boldsymbol{\rho}_0)} = \sum_{i=1}^{n} \mathcal{P}(\mathcal{U}_i \,|\, \boldsymbol{\rho}_0)\, \Theta_\infty(\mathcal{U}_i),\end{aligned} \tag{3.30}$$

where $\mathcal{P}(\mathcal{U}_i \,|\, \boldsymbol{\rho}_0)$ is the probability for a certain subset of unravellings $\mathcal{U}_i \subseteq \mathcal{U}$, whose analytical expression can be derived from the stationary solution of the Markov chain associated to the quantum state transition network introduced in Section 3.6 (not to be confused with the stationary solution of the Lindblad equation). For the classical analog expression, compare Section 2.5.2.
On the other hand, the density matrix $\Theta_\infty(\mathcal{U}_i)$ is indeed a stationary solution of the Lindblad Equation (1.2), independent of the initial condition, but dependent on the stationary solutions of Markov chains (compare equations (3.33)), as well as time averages, for which analytical expressions will be derived in Section 3.8.
Calculating the different building blocks of this expression is the goal of this chapter. We start in the next Section 3.5 by computing an explicit expression for the time average $\Theta_\infty(\omega \,|\, \boldsymbol{\rho}_0)$ of a single quantum trajectory.

3.5. The time average of a single quantum trajectory

In order to compute the time average for a single quantum trajectory, we perform similar calculations as we did in Section 2.4.6.
Similar to the classical case, the time average of a trajectory depends on, whether the number of jumps is finite or infinite:

i) finitely many jumps
When there are finitely many quantum jumps in the trajectory, then the last state in the 'discrete' quantum trajectory must be a possible trapping state Θ_{trap} (defined in the algorithm for the unravelling, Section 3.3.1), with a non-zero chance (namely $\lim_{t\to\infty} \mathrm{Tr}[U_t(\Theta_{\text{trap}})] > 0$) that no more jumps would occur. The time average of that trajectory is then given by $\langle \Theta_{\text{trap}}(t) \rangle_{t\geq 0} := \langle \Theta(t) \rangle_{t\geq t_{\text{trap}}} := \lim_{T\to\infty} \frac{1}{T} \int_0^T \frac{U_t(\Theta_{\text{trap}})}{\mathrm{Tr}[\dots]}\, \mathrm{d}\, t$.
In favor of a uniform description, we say that the trajectory makes a transition towards the time-averaged state (see Figure 3.5 for an illustration).

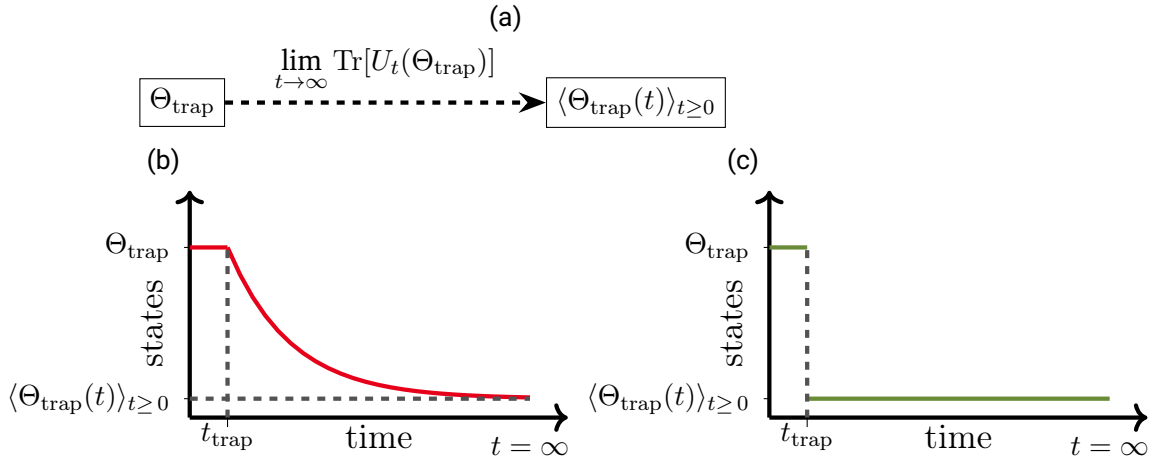


Figure 3.5.: Illustrating the transition from a *possible trapping state* Θ_{trap} to the *time-averaged trapping state* $\langle\Theta_{\mathsf{trap}}(t)\rangle_{t\geq 0}$: While the dynamics for $t \geq t_{\mathsf{trap}}$ is in fact continuously approaching the time-averaged state (compare Figure 3.5b), we treat it as an instantaneous transition as depicted in Figure 3.5c, in order to be compatible with the state transition network of the discrete-time Markov chain (see Figure 3.5).

ii) infinitely many jumps

In order to compute the time average of a quantum trajectory (which we know exists from [KM04]) for infinitely many jumps, we rely on the same notation as the classical case, defined in Equation (2.67).

In addition to the average waiting time (see Equation (2.69)) we come across expressions for the stationary probabilities within a minimal absorbing set, that is

$$\frac{J_s(T)}{J(T)} \xrightarrow{T\to\infty} q_\infty(\Theta_s \,|\, \boldsymbol{\rho}_0 \in B) \qquad \text{in probability, for all states } s \in B \in \mathcal{B}. \tag{3.31}$$

We use the notation $q_\infty(\Theta_s \,|\, \boldsymbol{\rho}_0 \in B)$ in contrast to $q_\infty^{(s)}(\boldsymbol{p}_0 \in B)$, in order to focus on the dependence of the density matrix Θ_s instead of the unit vector $\boldsymbol{e}_s$ in the classical case.

Another expression which appears is the "average of the same 'time averaged states'". What this expression means, is that a time averaged state $\langle \Theta_s(t) \rangle_{t \in [0,\tau_n]}$ depends on the waiting time, which is itself a random variable. By averaging over all outcomes of this random variable, we get:

$$\begin{aligned} &\frac{1}{K} \sum_{k=0}^{K-1} \tau\big(\alpha_s(k)\big) \cdot \langle \Theta_s(t) \rangle_{t \in [0,\tau(\alpha_s(k))]} \\ &\qquad \xrightarrow{K \to \infty} \big\langle \tau \cdot \langle \Theta_s(t) \rangle_{t \in [0,\tau]} \big\rangle_{\tau \sim f(\cdot \,|\, \Theta_s)} \\ &\qquad = \int_0^\infty \tau \, \langle \Theta_s(t) \rangle_{t \in [0,\tau]} \, f(\tau \,|\, \Theta_s) \;\, \mathrm{d}\,\tau \\ &\qquad = \int_0^\infty \mathrm{d}\,\tau \int_0^\tau \mathrm{d}\,t \, \Theta_s(t) \, f(\tau \,|\, \Theta_s). \end{aligned} \tag{3.32}$$

When we assume that the quantum trajectory exhibits infinitely many quantum jumps and is eventually captured by the minimal absorbing set $B \in \mathcal{B}$, we can compute its time average as follows:

$$
\begin{aligned}
\Theta_\infty(\omega \,|\, \boldsymbol{\rho}_0,\, \omega \in B \text{ eventually }) &\overset{\text{theorem (B.1)}}{=\!=\!=} \langle \Theta(\omega, t \,|\, \boldsymbol{\rho}_0) \cdot 1_{\{\omega \in B \text{ eventually }\}} \rangle_{t \geq 0} \\
&= \lim_{T\to\infty} \frac{1}{T} \int_0^T \Theta(t, \omega \,|\, \boldsymbol{\rho}_0)\; \mathrm{d}\, t \cdot 1_{\{\omega \in B \text{ eventually }\}} = \\
&= \lim_{T\to\infty} \frac{1}{T} \sum_{n=0}^{J(T)-1} \underbrace{\int_{t_n}^{t_{n+1}} \frac{U_{t-t_n}(\,\Theta_n\,)}{\operatorname{Tr}[\dots]}\; \mathrm{d}\, t}_{\int_0^{\tau_n} \frac{U_t(\,\Theta_n\,)}{\operatorname{Tr}[\dots]}\; \mathrm{d}\, t} \cdot 1_{\{\omega \in B \text{ eventually }\}} \\
&= \lim_{T\to\infty} \frac{1}{T} \sum_{n=0}^{J(T)-1} \underbrace{\int_0^{\tau_n} \frac{U_t(\,\Theta_n\,)}{\operatorname{Tr}[\dots]}\; \mathrm{d}\, t}_{\tau_n \cdot \langle \Theta_n(t) \rangle_{t\in[0,\tau_n]}} \cdot 1_{\{\omega \in B \text{ eventually }\}} \\
&= \lim_{T\to\infty} \frac{1}{T} \sum_{n=0}^{J(T)-1} \tau_n \cdot \langle \Theta_n(t) \rangle_{t\in[0,\tau_n]} \cdot \frac{1_{\{\omega \in B \text{ eventually }\}}}{1_{\{\omega \in \bigcup_{B'\in\mathcal{B}} B' \text{ eventually }\}}} \\
&= \lim_{T\to\infty} \frac{\sum_{s\in\Omega_{\boldsymbol{\rho}_0}} \frac{J_s(\omega,T)}{J(\omega,T)}\, 1_{\{\omega\in B \text{ event.}\}} \frac{1}{J_s(\omega,T)} \sum_{k=0}^{J_s(\omega,T)-1} \tau_{\alpha_s(k)} \cdot \langle \overbrace{\Theta_{\alpha_s(k)}}^{\Theta_s}(t) \rangle_{t\in[0,\tau_{\alpha_s(k)}]}}{\frac{T}{J(\omega,T)} \cdot 1_{\{\omega \in \bigcup_{B'\in\mathcal{B}} B' \text{ event.}\}}} \\
&= \sum_{s\in\Omega_{\boldsymbol{\rho}_0}} \frac{\overbrace{\left(\lim_{T\to\infty} \frac{J_s(\omega,T)}{J(\omega,T)} 1_{\{\omega\in B \text{ event.}\}}\right)}^{q_\infty(\Theta_s \,|\, \boldsymbol{\rho}_0 \in B)} \overbrace{\left(\lim_{T\to\infty} \frac{1}{J_s(\omega,T)} \sum_{k=0}^{J_s(\omega,T)-1} \tau_{\alpha_s(k)} \cdot \langle \Theta_s(t) \rangle_{t\in[0,\tau_{\alpha_s(k)}]}\right)}^{\langle \tau \cdot \langle \Theta_s(t) \rangle_{t\in[0,\tau]} \rangle_{\tau\sim f(\cdot\,|\,\Theta_s)}}}{\sum_{B'\in\mathcal{B}} \sum_{b'\in B'} \underbrace{\left(\lim_{T\to\infty} 1_{\{\omega\in B' \text{ event.}\}} \frac{J_b(\omega,T)}{J(\omega,T)}\right)}_{q_\infty(\Theta_{b'} \,|\, \boldsymbol{\rho}_0 \in B')} \underbrace{\left(\lim_{T\to\infty} \frac{1}{J_{b'}(\omega,T)} \sum_{k=0}^{J_{b'}(\omega,T)-1} \tau_{\alpha_s(k)}\right)}_{\overline{\tau_{b'}}}} \\
&= \frac{\sum_{b\in B} q_\infty(\Theta_b \,|\, \boldsymbol{\rho}_0 \in B)\, \langle \tau \cdot \langle \Theta_s(t) \rangle_{t\in[0,\tau]} \rangle_{\tau\sim f(\cdot\,|\,\Theta_s)}}{\operatorname{Tr}[\dots]} =: \Theta_B.
\end{aligned}
\tag{3.33}
$$

We note that the expression (3.33) does not depend on the initial state $\boldsymbol{\rho}_0$ and not on the details of the quantum trajectory $\omega \in \mathcal{U}$ but only on the minimal absorbing set $B \in \mathcal{B}$ the trajectory is eventually captured in.

When the states are constant between quantum jumps, the time averaged state is just the state times the waiting time ($\langle \Theta_s(t) \rangle_{t\in[0,\tau]} = \tau \cdot \Theta_s$) and Equation (3.33) reduces to

$$\Theta_B := \frac{\sum\limits_{s\in B} q_\infty(\Theta_s \,|\, \boldsymbol{\rho}_0 \in B)\ \overline{\tau_s}\ \Theta_s}{\mathrm{Tr}[\dots]}. \tag{3.34}$$

When additionally the conditional Hamiltonian is trivial ($H_c = \beta\, \mathbb{1}$ for some number $\beta \in \mathbb{R}_{\geq 0}$), then all waiting times have the same distribution ($\tau \sim \mathrm{Exp}(\beta)$) and formula (3.33) reduces again to

$$\begin{aligned} \Theta_B &:= \frac{\sum\limits_{s\in B} q_\infty(\Theta_s \,|\, \boldsymbol{\rho}_0 \in B)\ \overline{\tau_s} \cdot \Theta_s}{\mathrm{Tr}[\dots]} \\ &\overset{\overline{\tau_s}=\frac{1}{\beta}}{=\!=\!=} \sum_{s\in B} q_\infty(\Theta_s \,|\, \boldsymbol{\rho}_0 \in B)\ \Theta_s. \end{aligned} \tag{3.35}$$

For the general case we have:

$$\begin{aligned} \Theta_\infty(\omega \,|\, \boldsymbol{\rho}_0) &= \langle \Theta(\omega,t) \cdot 1_{\{\omega \in \bigcup_{B\in\mathcal{B}} B \text{ eventually }\}} \rangle_{t\geq 0} \\ &= \sum_{B\in\mathcal{B}} 1_{\{\omega\in B \text{ eventually }, \Theta_0 = \boldsymbol{\rho}_0\}} \cdot \underbrace{\langle \langle \Theta(\omega,t \,|\, \boldsymbol{\rho}_0) \cdot 1_{\{\omega\in B \text{ eventually }\}} \rangle_{t\geq 0}}_{\Theta_B}, \\ &= \sum_{B\in\mathcal{B}} 1_{\mathcal{U}(\boldsymbol{\rho}_0 \rightsquigarrow B)}(\omega) \cdot \underbrace{\left(\frac{\sum\limits_{b\in B} q_\infty(\Theta_b \,|\, \boldsymbol{\rho}_0 \in B)\, \langle \tau \cdot \langle \Theta_s(t) \rangle_{t\in[0,\tau]} \rangle_{\tau\sim f(\cdot\,|\,\Theta_s)}}{\mathrm{Tr}[\dots]} \right)}_{\Theta_B} \end{aligned} \tag{3.36}$$

where

$$\Theta_B := \begin{cases} \langle \Theta_{\mathrm{trap}}(t \,|\, \boldsymbol{\rho}_0) \rangle_{t\geq 0} & \text{, for finitely many quantum jumps and} \\ \frac{\sum\limits_{b\in B} q_\infty(\Theta_b \,|\, \boldsymbol{\rho}_0 \in B)\, \langle \tau \cdot \langle \Theta_s(t) \rangle_{t\in[0,\tau]} \rangle_{\tau\sim f(\cdot\,|\,\Theta_s)}}{\mathrm{Tr}[\dots]} & \text{, when the number of jumps is infinite.} \end{cases} \tag{3.37}$$

3.6. The transition probabilities of the Markov chain

The set of possible states in the Markov chain was defined in equation (3.17) as

$$\{\Theta_s \,:\, s \in \mathcal{S}\} := \Omega_{\boldsymbol{\rho}_0} = \bigcup_{\omega \in \mathcal{U}} \Omega_{\boldsymbol{\rho}_0}(\omega).$$

We choose the initial distribution vector $\boldsymbol{q}_{\boldsymbol{\rho}_0}$ to be the first unit vector $\boldsymbol{q}_{\boldsymbol{\rho}_0} = \boldsymbol{e}_1 = (1, 0, \ldots, 0) \in \mathbb{R}^{|\Omega_{\boldsymbol{\rho}_0}|}$, corresponding to the initial state $\boldsymbol{\rho}_0 = \Theta_0 \in \Omega_{\boldsymbol{\rho}_0}$, which is the first element of the set of possible states in the quantum trajectory.
We assume that in all quantum trajectories only finitely many states appear, $|\Omega_{\boldsymbol{\rho}_0}| < \infty$. The transition probability $\mathcal{P}(\Theta_{s_1} \to \Theta_{s_2})$ is the probability that the next state in the quantum trajectory equals Θ_{s_2}, provided that the last state was Θ_{s_1}. It can be computed to

$$\begin{aligned}
\mathcal{P}(\Theta_{s_2} \,|\, \Theta_{s_1}) &= \int_{\mathbb{R}_{\geq 0}} \underbrace{p(\tau, \Theta_{s_2} \,|\, \Theta_{s_1})}_{f(\tau \,|\, \Theta_{s_1}) \cdot \mathcal{P}(\Theta_{s_2} \,|\, \tau, \Theta_{s_1})} \quad \mathrm{d}\tau \\
&= \int_{\mathbb{R}_{\geq 0}} f(\tau \,|\, \Theta_{s_1}) \cdot \underbrace{\mathcal{P}(\Theta_{s_2} \,|\, \tau, \Theta_{s_1})}_{\sum_{k \in I(s_1 \to s_2)} \frac{f^{(k)}(\tau \,|\, \Theta_{s_1})}{f(\tau \,|\, \Theta_{s_1})}} \quad \mathrm{d}\tau \\
&= \sum_{k \in I(s_1 \to s_2)} \int_{\mathbb{R}_{\geq 0}} f^{(k)}(\tau \,|\, \Theta_{s_1}) \, \mathrm{d}\tau,
\end{aligned} \tag{3.38}$$

where the sum goes over all Lindblad operators V_k that yield a transition from state Θ_{s_1} to Θ_{s_2}, or more formally: $I(s_1 \to s_2) := \{k \in I \,:\, J_k \circ U_t(\Theta_{s_1}) = \Theta_{s_2} \text{ for some } t > 0\}$.
The following examples illustrate both the states of all possible discrete quantum trajectories as well as the transition probabilities.

3.7. Examples revisited

In this Section we present three examples that illustrate the state transition network $\Omega_{\boldsymbol{\rho}_0}$ for a specific choice of Hamiltonian and Lindblad operators. We note that $\Omega_{\boldsymbol{\rho}_0}$ depends on the initial state $\boldsymbol{\rho}_0$ and certain transition probabilities can vanish for special initial conditions (e.g. $\boldsymbol{\rho}_{11}(0) = 0$). The transition network consists of only those states, that can be reached from $\boldsymbol{\rho}_0$.

3.7.1. Example with trivial Hamiltonian and two transition operators

We consider a system of dimension $N = 2$ with a trivial Hamiltonian and two transition operators as Lindblad operators, that is:

$$H \propto \mathbb{1}, V_1 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, V_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \text{ and } \Lambda = \begin{pmatrix} \gamma_1 & 0 \\ 0 & \gamma_2 \end{pmatrix}. \tag{3.39}$$

The set of possible states appearing in the Markov chain can be computed to:

$$\Omega_{\boldsymbol{\rho}_0} = \left\{ \Theta_{s_1} = \begin{pmatrix} \boldsymbol{\rho}_{11}(0) & \boldsymbol{\rho}_{12}(0) \\ \boldsymbol{\rho}_{21}(0) & \boldsymbol{\rho}_{22}(0) \end{pmatrix}, \Theta_{s_2} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \Theta_{s_3} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\} \tag{3.40}$$

The transition matrix equals

$$Q = \begin{pmatrix} 0 & 0 & 0 \\ \boldsymbol{\rho}_{11}(0) & 0 & 1 \\ \boldsymbol{\rho}_{22}(0) & 1 & 0 \end{pmatrix}, \text{with the ordered states } (\Theta_{s_1} = \boldsymbol{\rho}_0, \Theta_{s_2}, \Theta_{s_3}). \tag{3.41}$$

In this example there are no possible trapping states.

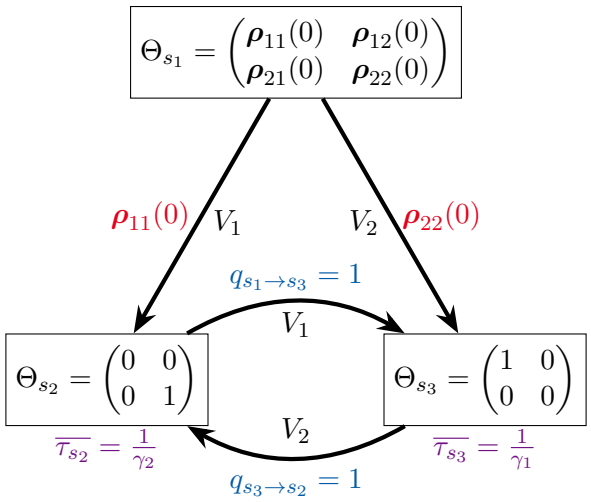


Figure 3.6.: State transition network for a system given by Equation (3.39). Every discrete quantum trajectory starting at the initial state $\Theta_{s_1} = \boldsymbol{\rho}_0$ is after one quantum jump in the minimal absorbing set $B = \{\Theta_{s_2}, \Theta_{s_3}\}$, where it oscillates between the two states Θ_{s_2} (with average waiting time $\overline{\tau_{s_2}} = \frac{1}{\gamma_2}$) and Θ_{s_3} (with average waiting time $\overline{\tau_{s_3}} = \frac{1}{\gamma_1}$).
For $\boldsymbol{\rho}_0 = \frac{1}{2}\,\mathbb{1}_2$, the three states are linearly dependent.

The stationary state can be computed as follows:

$$\begin{aligned}
\boldsymbol{\rho}_\infty(\boldsymbol{\rho}_0) &= \sum_{B\in\mathcal{B}} \mathcal{P}(B \,|\, \boldsymbol{\rho}_0) \sum_{s\in B} \frac{q_\infty(s \,|\, B)\ \langle \boldsymbol{\tau} \cdot \langle \Theta_s(t) \rangle_{t\in[0,\boldsymbol{\tau}]} \rangle_{\boldsymbol{\tau} \sim f(\cdot \,|\, \Theta_s)}}{\mathrm{Tr}[\dots]} \quad = \\
&\underset{B=\{\Theta_{s_2},\Theta_{s_2}\}}{\overset{\mathcal{B}=\{B\}}{=\!=\!=}} \sum_{B\in\{B_1\}} \underbrace{\mathcal{P}(B_1 \,|\, \boldsymbol{\rho}_0)}_{(\boldsymbol{\rho}_{11}(0)+\boldsymbol{\rho}_{22}(0))=1} \frac{1}{\mathrm{Tr}[\dots]} \sum_{i=1}^{2} \underbrace{q_\infty(s_i \,|\, B_1)}_{\frac{1}{2}}\ \overline{\boldsymbol{\tau}_{s_i}}\ \Theta_{s_i} = \\
&= \frac{1}{\mathrm{Tr}[\dots]} \left[\frac{1}{2}\frac{1}{\gamma_1} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \frac{1}{2}\frac{1}{\gamma_2} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right] \\
&= \frac{1}{\mathrm{Tr}[\dots]} \begin{pmatrix} \frac{1}{\gamma_1} & 0 \\ 0 & \frac{1}{\gamma_2} \end{pmatrix}.
\end{aligned}$$

3.7.2. Example with trivial Hamiltonian and one projection matrix as Lindblad operator

We consider a system of dimension $N = 2$ with a trivial Hamiltonian and a projection operator for a Lindblad operators, that is:

$$H \propto \mathbb{1},\ V_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},\ \Lambda = \begin{pmatrix} \gamma_1 & 0 \\ 0 & 0 \end{pmatrix}. \tag{3.42}$$

The set of possible states appearing in the Markov chain can be computed to:

$$\Omega_{\boldsymbol{\rho}_0} = \left\{ \Theta_{s_1} = \begin{pmatrix} \boldsymbol{\rho}_{11}(0) & \boldsymbol{\rho}_{12}(0) \\ \boldsymbol{\rho}_{21}(0) & \boldsymbol{\rho}_{22}(0) \end{pmatrix}, \Theta_{s_2} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \Theta_{s_3} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \right\}. \tag{3.43}$$

In this case, $\Theta_{s_1} = \boldsymbol{\rho}_0$ is a possible trapping state, since with non-vanishing probability, it will be the first and only state in the discrete quantum trajectory. The transition matrix is of the following form: $Q = \begin{pmatrix} 0 & 0 & 0 \\ \boldsymbol{\rho}_{11}(0) & 1 & 0 \\ \boldsymbol{\rho}_{22}(0) & 0 & 1 \end{pmatrix}$.

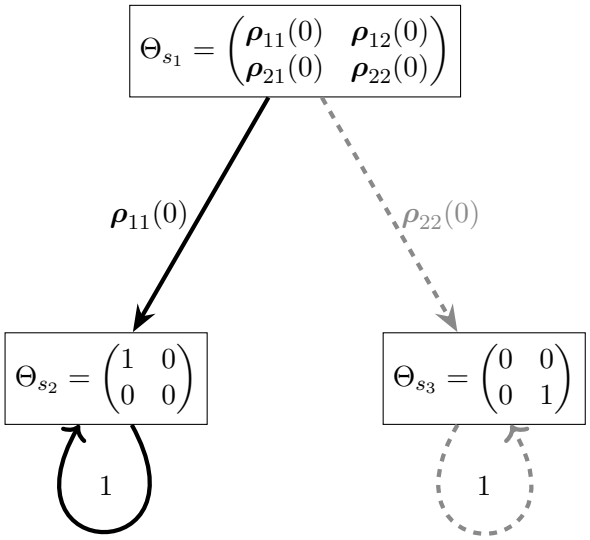


Figure 3.7.: State transition network for a system given by Equation (3.42). In the quantum trajectory, starting at $\Theta_{s_1} = \rho_0$, the Lindblad operator V_1 is either applied infinitely often (this happens with probability $\rho_{11}(0)$ and results in a constant trajectory, $\Theta(t) = \Theta_{s_2}$ for all $t \geq t_1$), or not at all. The latter happens with probability $\rho_{22}(0)$ and results in the quantum trajectory converging to the time-averaged state $\Theta(t) \xrightarrow{t\to\infty} \Theta_{s_3} = \langle \Theta_{s_1}(t) \rangle_{t\geq 0}$.

3.7.3. Example of a state transition network with $|\Omega| = 5$ states and a non-trivial Hamiltonian

We consider the following system with dimension $N = 4$, where the Hamilton operator and the two Lindblad operators are given by:

$$
\begin{aligned}
H &= \begin{pmatrix} E_1 & 0 & 0 & 0 \\ 0 & E_1 & 0 & 0 \\ 0 & 0 & E_2 & 0 \\ 0 & 0 & 0 & E_2 \end{pmatrix}, V_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, V_2 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \\
\Lambda &= \begin{pmatrix} E_1 - \frac{i\gamma_1}{2} & 0 & 0 & 0 \\ 0 & E_1 - \frac{i\gamma_1}{2} & 0 & 0 \\ 0 & 0 & E_2 - \frac{i\gamma_2}{2} & 0 \\ 0 & 0 & 0 & E_2 - \frac{i\gamma_2}{2} \end{pmatrix}.
\end{aligned} \tag{3.44}
$$

The set of possible states appearing in the Markov consists of the five states

$$
\Omega_{\boldsymbol{\rho}_0} = \{\Theta_{s_1} = \boldsymbol{\rho}_0, \Theta_{s_2}, \Theta_{s_3}, \Theta_{s_4}, \Theta_{s_5}\}. \tag{3.45}
$$

with

$$
\begin{aligned}
\Theta_{s_1} &= \begin{pmatrix} \rho_{11}(0) & \rho_{12}(0) & \rho_{13}(0) & \rho_{14}(0) \\ \rho_{21}(0) & \rho_{22}(0) & \rho_{23}(0) & \rho_{24}(0) \\ \rho_{31}(0) & \rho_{32}(0) & \rho_{33}(0) & \rho_{34}(0) \\ \rho_{41}(0) & \rho_{42}(0) & \rho_{43}(0) & \rho_{44}(0) \end{pmatrix} \\
\Theta_{s_2} &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{\rho_{22}(0)}{\rho_{11}(0)+\rho_{22}(0)} & \frac{\rho_{21}(0)}{\rho_{11}(0)+\rho_{22}(0)} \\ 0 & 0 & \frac{\rho_{12}(0)}{\rho_{11}(0)+\rho_{22}(0)} & \frac{\rho_{11}(0)}{\rho_{11}(0)+\rho_{22}(0)} \end{pmatrix}, \Theta_{s_3} = \begin{pmatrix} \frac{\rho_{44}(0)}{\rho_{33}(0)+\rho_{44}(0)} & \frac{\rho_{43}(0)}{\rho_{33}(0)+\rho_{44}(0)} & 0 & 0 \\ \frac{\rho_{34}(0)}{\rho_{33}(0)+\rho_{44}(0)} & \frac{\rho_{33}(0)}{\rho_{33}(0)+\rho_{44}(0)} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\
\Theta_{s_4} &= \begin{pmatrix} \frac{\rho_{11}(0)}{\rho_{11}(0)+\rho_{22}(0)} & \frac{\rho_{12}(0)}{\rho_{11}(0)+\rho_{22}(0)} & 0 & 0 \\ \frac{\rho_{21}(0)}{\rho_{11}(0)+\rho_{22}(0)} & \frac{\rho_{22}(0)}{\rho_{11}(0)+\rho_{22}(0)} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \Theta_{s_5} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{\rho_{33}(0)}{\rho_{33}(0)+\rho_{44}(0)} & \frac{\rho_{34}(0)}{\rho_{33}(0)+\rho_{44}(0)} \\ 0 & 0 & \frac{\rho_{43}(0)}{\rho_{33}(0)+\rho_{44}(0)} & \frac{\rho_{44}(0)}{\rho_{33}(0)+\rho_{44}(0)} \end{pmatrix},
\end{aligned}
$$

and the transition probabilities

$$
\begin{aligned}
q_{s_1 \to s_2} &= \rho_{11}(0) + \rho_{22}(0) \\
q_{s_1 \to s_3} &= \rho_{33}(0) + \rho_{44}(0)
\end{aligned} \tag{3.46}
$$

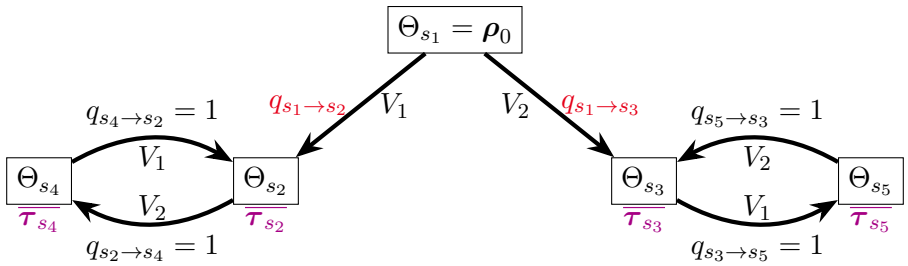


Figure 3.8.: State transition network for a system of dimension $N = 4$, where the Hamiltonian and Lindblad operators are given by Equation (3.44). Starting with the initial state $\boldsymbol{\rho}_0$, the second state in the discrete quantum trajectory is either Θ_{s_2} or Θ_{s_3}, which determines its future evolution: Either we have an oscillation between the states Θ_{s_2} and Θ_{s_4} in the minimal absorbing set $B_1 = \{\Theta_{s_2}, \Theta_{s_4}\}$, or an oscillation between the states Θ_{s_3} and Θ_{s_5} in the minimal absorbing set $B_2 = \{\Theta_{s_3}, \Theta_{s_5}\}$.

$$
\begin{aligned}
\boldsymbol{\rho}_\infty(\boldsymbol{\rho}_0) &= \sum_{B \in \mathcal{B}} \mathcal{P}(B \,|\, \boldsymbol{\rho}_0) \sum_{s \in B} \frac{q_\infty(\Theta_s \,|\, B) \;\; \langle \boldsymbol{\tau} \cdot \langle \Theta_s(t) \rangle_{t \in [0, \boldsymbol{\tau}]} \rangle_{\boldsymbol{\tau} \sim f(\cdot \,|\, \Theta_s)}}{\mathrm{Tr}[\dots]} \\
&\overset{(*)}{=\!=} \underbrace{\mathcal{P}(B_1 \,|\, \boldsymbol{\rho}_0)}_{q_{s_1 \to s_2}} \frac{\sum_{i \in \{2,4\}} \overbrace{q_\infty(\Theta_{s_i} \,|\, B_1)}^{\frac{1}{2}} \;\; \overline{\boldsymbol{\tau}_{s_i}} \;\; \Theta_{s_i}}{\mathrm{Tr}[\dots]} \\
&+ \quad \underbrace{\mathcal{P}(B_2 \,|\, \boldsymbol{\rho}_0)}_{q_{s_1 \to s_3}} \frac{\sum_{i \in \{3,5\}} \overbrace{q_\infty(\Theta_{s_i} \,|\, B_2)}^{\frac{1}{2}} \;\; \overline{\boldsymbol{\tau}_{s_i}} \;\; \Theta_{s_i}}{\mathrm{Tr}[\dots]} \\
&= \begin{pmatrix}
\frac{\gamma_2(\rho_{11}(0)+\rho_{44}(0))}{\gamma_1+\gamma_2} & \frac{\gamma_2(\rho_{12}(0)+\rho_{43}(0))}{\gamma_1+\gamma_2} & 0 & 0 \\
\frac{\gamma_2(\rho_{21}(0)+\rho_{34}(0))}{\gamma_1+\gamma_2} & \frac{\gamma_2(\rho_{22}(0)+\rho_{33}(0))}{\gamma_1+\gamma_2} & 0 & 0 \\
0 & 0 & \frac{\gamma_1(\rho_{22}(0)+\rho_{33}(0))}{\gamma_1+\gamma_2} & \frac{\gamma_1(\rho_{21}(0)+\rho_{34}(0))}{\gamma_1+\gamma_2} \\
0 & 0 & \frac{\gamma_1(\rho_{12}(0)+\rho_{43}(0))}{\gamma_1+\gamma_2} & \frac{\gamma_1(\rho_{11}(0)+\rho_{44}(0))}{\gamma_1+\gamma_2}
\end{pmatrix},
\end{aligned}
$$

where we used in step (*) that $\mathcal{B} = \{B_1, B_2\} = \{\{\Theta_{s_2}, \Theta_{s_4}\}, \{\Theta_{s_3}, \Theta_{s_5}\}$.

3.7.4. Recovering the classical case

It is possible to recover the 'classical' master equation as the diagonal of a special Lindblad equation, when we choose the Lindblad operators

$$\begin{aligned} V_{m\,n} &\overset{m\neq n}{=\!=} |n\rangle\langle m| \quad &&\text{, with rate } \gamma_{m\to n} \geq 0 \text{ and} \\ V_{m\,m} &= |m\rangle\langle m| &&\text{, with rate } \gamma_m > 0 \end{aligned} \tag{3.47}$$

as the set of all transition operators $|n\rangle\langle m|$, from state $|m\rangle$ to state $|n\rangle$ as well as all the projection operators $|m\rangle\langle m|$ for all $m, n \in \{1, \ldots, N\}$, where the dimension of the Lindblad equation must be chosen to coincide with that of the master equation, $N := |\Omega|$. There are two ways, to see this:

The right-hand side of the differential equation

When computing the right-hand side of the Lindblad equation componentwise for the Lindblad operators given above, we obtain the master equation for the diagonal terms, while the off-diagonal terms decay exponentially to zero:

$$\begin{aligned} \dot{\boldsymbol{\rho}}_{i\,i} &= \sum_{m=1}^{N} \boldsymbol{\rho}_{mm}\,\gamma_{m\to i} - \boldsymbol{\rho}_{ii}\,\gamma_{i\to m} \\ \dot{\boldsymbol{\rho}}_{i\,j} &\overset{i\neq j}{=\!=} -\frac{1}{2}\left(\sum_{n=1}^{N} \gamma_{i\to n} + \gamma_{j\to n}\right)\boldsymbol{\rho}_{i\,j} \\ &\leq -\frac{\gamma_{i\to i} + \gamma_{j\to j}}{2}\,\boldsymbol{\rho}_{i\,j}. \end{aligned} \tag{3.48}$$

Without introducing the projection operators $|m\rangle\langle m|$, the off-diagonal elements need not vanish.

The state transition network

An alternative way to see this is to compute the state transition network. The set of possible states is given by $\frac{|n\rangle\langle m|U_\tau(\Theta)|m\rangle\langle n|}{\mathrm{Tr}[\ldots]} = |n\rangle\langle n|$.
The transition probabilities can be computed to

$$
\begin{aligned}
&\mathcal{P}\left(\Theta(t_{k+1}) = |n\rangle\langle n| \,\middle|\, \Theta(t_k) = |m\rangle\langle m|\right) \\
&\qquad = \int_0^\infty \mathrm{d}\,\tau\, \gamma_{m\to n}\, \mathrm{Tr}\big[|n\rangle\langle m| \underbrace{U_\tau\left(|m\rangle\langle m|\right)}_{\mathrm{e}^{-t\,\gamma_{m\to}/2}\,|m\rangle\langle m|\,\mathrm{e}^{-t\,\gamma_{m\to}/2}} |m\rangle\langle n|\big] \\
&\qquad = \int_0^\infty \mathrm{d}\,\tau\, \mathrm{e}^{-t\,\gamma_{m\to}}\, \gamma_{m\to n} \\
&\qquad = \frac{\gamma_{m\to n}}{\gamma_{m\to}} \underbrace{\int_0^\infty \mathrm{d}\,\tau\, \gamma_{m\to}\, \mathrm{e}^{-t\,\gamma_{m\to}}}_{1},
\end{aligned}
\tag{3.49}
$$

which is precisely the transition probability of the associated, embedded discrete-time Markov chain (compare lemma 17) with the waiting time distribution

$$
\begin{aligned}
f\left(t \,|\, \Theta = |l\rangle\langle l|\right) &= \sum_{m,n=1}^{N} \gamma_{m\to n}\, \mathrm{Tr}\big[|n\rangle \underbrace{\langle m|\, \mathrm{e}^{-t\,\gamma_{l\to}}\, |l\rangle}_{\mathrm{e}^{-t\,\gamma_{l\to}}\,\delta_{l,m}} \langle l|\, |m\rangle\langle n|\big] = \gamma_{l\to}\, \mathrm{e}^{-t\,\gamma_{l\to}} \\
&\text{and the average waiting time} \\
\overline{\tau_l} &= \frac{1}{\gamma_{l\to}}.
\end{aligned}
\tag{3.50}
$$

3.8. Evaluating the stationary state Θ_B of the Lindbladian for the minimal absorbing set B

We now evaluate the stationary states Θ_B occurring in Equation (3.37).
We write the conditional Hamiltonian H_c (see Equation (1.3)) as a direct sum of Jordan blocks

$$
\begin{aligned}
H_c &= \bigoplus_{m=1}^{M} J_m = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_M \end{pmatrix}, \text{ with} \\
J_m &= \begin{pmatrix} \delta_m & 1 & & \\ & \ddots & \ddots & \\ & & \delta_m & 1 \\ & & & \delta_m \end{pmatrix} \in \mathbb{C}^{s_m \times s_m}
\end{aligned}
\tag{3.51}
$$

being a Jordan block of size $s_m \in \mathbb{N}$ and $\sum_{m=1}^{M} s_m = N$.
We discuss separately the two cases of a finite and an infinite number of quantum jumps.

3.8.1. Infinitely many quantum jumps

When the quantum trajectory contains infinitely many quantum jumps, we need to evaluate the expression $\langle \tau \cdot \langle \Theta_s(t) \rangle_{t\in[0,\tau]} \rangle_{\tau \sim f(\cdot\,|\,\Theta_s)}$ defined in Equation (3.32).
If Θ_s is not a possible trapping state, we can compute the (conditional) time evolution operator $\mathrm{e}^{-i\,H_c\,t}$ (for details see appendix B.2) and get

$$
\begin{aligned}
\langle \tau \cdot \langle \Theta_s(t) \rangle_{t\in[0,\tau]} \rangle_{\tau \sim f(\cdot\,|\,\Theta_s)} &= \int_0^\infty \underbrace{f(\tau\,|\,\Theta_s)}_{-\frac{\mathrm{d}}{\mathrm{d}\tau}\mathrm{Tr}[U_\tau(\Theta_s)]} \underbrace{\tau \cdot \langle \Theta(t) \rangle_{t\in[0,\tau]}}_{\int_0^\tau \frac{U_t(\Theta_s)}{\mathrm{Tr}[\dots]}\,\mathrm{d}\,t} \,\mathrm{d}\tau = \\
&\overset{\text{partial integration}}{=\!=\!=} -\mathrm{Tr}[U_\tau(\Theta_s)] \cdot \int_0^\tau \frac{U_t(\Theta_s)}{\mathrm{Tr}[\dots]}\,\mathrm{d}\,t \,\Big|_{\tau=0}^{\tau=\infty} + \int_0^\infty \mathrm{Tr}[U_\tau(\Theta_s)]\,\frac{U_\tau(\Theta_s)}{\mathrm{Tr}[U_\tau(\Theta_s)]}\,\mathrm{d}\tau = \\
&= \underbrace{-\lim_{\tau\to\infty} \mathrm{Tr}[U_\tau(\Theta_s)] \cdot \int_0^\tau \frac{U_t(\Theta_s)}{\mathrm{Tr}[\dots]}\,\mathrm{d}\,t}_{0} + \int_0^\infty U_\tau(\Theta_s)\,\mathrm{d}\tau = \\
&\overset{B.2}{=\!=} \int_0^\infty U_\tau(\Theta_s)\,\mathrm{d}\tau \overset{\text{Def } U_\tau(\cdot)}{=\!=\!=} \int_0^\infty \mathrm{e}^{-i\,H_c\,\tau}\,\Theta_s\,(\mathrm{e}^{-i\,H_c\,\tau})^\dagger\,\mathrm{d}\,\tau = \\
&= \int_0^\infty \underbrace{\begin{pmatrix} \mathrm{e}^{-i\,J_1\,\tau} & & 0 \\ & \ddots & \\ 0 & & \mathrm{e}^{-i\,J_M\,\tau} \end{pmatrix} \begin{pmatrix} \hat{\boldsymbol{\Theta}}^{(1,1)} & \dots & \hat{\boldsymbol{\Theta}}^{(1,M)} \\ \vdots & & \vdots \\ \hat{\boldsymbol{\Theta}}^{(M,1)} & \dots & \hat{\boldsymbol{\Theta}}^{(M,M)} \end{pmatrix} \begin{pmatrix} \mathrm{e}^{i\,J_1^T\,\tau} & & 0 \\ & \ddots & \\ 0 & & \mathrm{e}^{i\,J_M^T\,\tau} \end{pmatrix}}_{\begin{pmatrix} \mathrm{e}^{-i\,J_1\,\tau}\,\hat{\boldsymbol{\Theta}}^{(1,1)}\,\mathrm{e}^{i\,J_1^T\,\tau} & \dots & \mathrm{e}^{-i\,J_1\,\tau}\,\hat{\boldsymbol{\Theta}}^{(1,M)}\,\mathrm{e}^{i\,J_M^T\,\tau} \\ \vdots & & \vdots \\ \mathrm{e}^{-i\,J_M\,\tau}\,\hat{\boldsymbol{\Theta}}^{(M,1)}\,\mathrm{e}^{i\,J_1^T\,\tau} & \dots & \mathrm{e}^{-i\,J_M\,\tau}\,\hat{\boldsymbol{\Theta}}^{(M,M)}\,\mathrm{e}^{i\,J_M^T\,\tau} \end{pmatrix}} \mathrm{d}\,\tau = \\
&= \begin{pmatrix} \int_0^\infty \mathrm{e}^{-i\,J_1\,\tau}\,\hat{\boldsymbol{\Theta}}^{(1,1)}\,\mathrm{e}^{i\,J_1^T\,\tau}\,\mathrm{d}\tau & \dots & \int_0^\infty \mathrm{e}^{-i\,J_1\,\tau}\,\hat{\boldsymbol{\Theta}}^{(1,M)}\,\mathrm{e}^{i\,J_M^T\,\tau}\,\mathrm{d}\tau \\ \vdots & & \vdots \\ \int_0^\infty \mathrm{e}^{-i\,J_M\,\tau}\,\hat{\boldsymbol{\Theta}}^{(M,1)}\,\mathrm{e}^{i\,J_1^T\,\tau}\,\mathrm{d}\tau & \dots & \int_0^\infty \mathrm{e}^{-i\,J_M\,\tau}\,\hat{\boldsymbol{\Theta}}^{(M,M)}\,\mathrm{e}^{i\,J_Y^T\,\tau}\,\mathrm{d}\tau \end{pmatrix}.
\end{aligned}
\tag{3.52}
$$

For the component $\left(\sum_{\mu=1}^{m-1} s_\mu + j, \sum_{\nu=1}^{n-1} s_\nu + k\right)$, we get with $\delta_x = R_x + i\, I_x$:

$$\left(\langle \tau \cdot \langle \Theta_s(t)\rangle_{t\in[0,\tau]}\rangle_{\tau\sim f(\cdot\,|\,\Theta_s)}\right)_{\left(\sum_{\mu=1}^{m-1} s_\mu + j, \sum_{\nu=1}^{n-1} s_\nu + k\right)} = \left(\int_0^\infty \mathrm{e}^{-i\, J_m\, \tau}\, \hat{\mathbf{\Theta}}^{(m,n)}\, \mathrm{e}^{i\, J_n^T\, \tau}\ \mathrm{d}\tau\right)_{j\,k}$$

$$\overset{(*)}{=\!=} \begin{cases} \sum_{\alpha=j}^{s_m} \sum_{\beta=k}^{s_n} \begin{pmatrix} \alpha-j+\beta-k \\ \alpha-j \end{pmatrix} \frac{(-1)^{\alpha-j+\beta-k}\, (\hat{\mathbf{\Theta}}^{(m,n)})_{jk}}{[i\,(R_n - R_m) + I_n + I_m]^{\alpha-j+\beta-k}} & , \text{if } I_n \neq 0 \neq I_m \\ 0 & , \text{else.} \end{cases} \tag{3.53}$$

In step (*) in the last line we used the fact that since Θ_s is not a possible trapping state, a diagonal block of Θ_s must vanish whenever the corresponding eigenvalue of Λ vanishes $\left(\lambda_m = 0 \Longrightarrow \hat{\mathbf{\Theta}}_s^{(m,m)} = \mathbf{0}^{s_m \times s_m}\right)$, see appendix B.2, and an inequality that is valid for all density matrices: $|(\Theta_s)_{jk}| \leq (\Theta_s)_{jj} \cdot (\Theta_s)_{kk}$.
An important special case is when both Jordan blocks are of size one, that is $s_m = 1 = s_n$. In that case we have

$$\left(\int_0^\infty \underbrace{\mathrm{e}^{-i\, J_m\, t}}_{\mathrm{e}^{-i\,\delta_m\, t}}\, \hat{\mathbf{\Theta}}^{(m,n)}\, \underbrace{\mathrm{e}^{i\, J_n^T\, t}}_{\mathrm{e}^{i\,\delta_n^*\, t}}\ \mathrm{d}\, t\right)_{11} = \begin{cases} \frac{\left(\hat{\mathbf{\Theta}}^{(m,n)}\right)_{11}}{[i\,(R_n - R_m) + I_n + I_m]} & , \text{if } I_n \neq 0 \neq I_m \\ 0 & , \text{else.} \end{cases} \tag{3.54}$$

When the conditional Hamiltonian H_c is diagonalizable, then all Jordan blocks are of size one, the block matrices $\hat{\mathbf{\Theta}}^{(m,n)} \in \mathbb{C}$ are scalars, and expression (3.53) becomes:

$$\left(\langle \tau \cdot \langle \Theta_s(t)\rangle_{t\in[0,\tau]}\rangle_{\tau\sim f(\cdot\,|\,\Theta_s)}\right)_{m\,n} = \frac{\left(\hat{\mathbf{\Theta}}^{(m,n)}\right)_{11}}{i\,(\delta_n^* - \delta_m)} = \begin{cases} \frac{(\Theta_s)_{m\,n}}{[i\,(R_n - R_m) + I_n + I_m]} & , \text{if } I_n \neq 0 \neq I_m \\ 0 & , \text{else.} \end{cases} \tag{3.55}$$

3.8.2. Finitely many quantum jumps

When the number of jumps is finite, we have to evaluate $\Theta_B = \left\langle \frac{U_t(\Theta_{\text{trap}})}{\text{Tr}[\dots]} \right\rangle_{t\geq 0}$, with Θ_{trap} being a possible trapping state.
By definition of a possible trapping state, we know that $\text{Tr}[U_t(\Theta_{\text{trap}})]$ converges to a positive value: $\text{Tr}[U_t(\Theta_{\text{trap}})] \xrightarrow{t\to\infty} q \in (0, 1]$. Then we have

$$\frac{\left(U_t(\Theta_{\text{trap}})\right)_{\left(\sum_{\mu=1}^{m-1} s_\mu + j,\, \sum_{\nu=1}^{n-1} s_\nu + k\right)}}{\text{Tr}[U_t(\Theta_{\text{trap}})]} = \frac{\sum_{\alpha=j}^{s_m} \sum_{\beta=k}^{s_n} \mathrm{e}^{i\,(R_n - R_m)\,t}\, \mathrm{e}^{(I_n + I_m)\,t}\, \frac{t^{(\alpha-j)}}{(\alpha-j)!}\, \frac{t^{(\beta-k)}}{(\beta-k)!} \left(\hat{\mathbf{\Theta}}^{(m,n)}\right)_{\alpha,\beta}}{\underbrace{\text{Tr}[U_t(\Theta_{\text{trap}})]}_{\xrightarrow{t\to\infty} q}}$$

$$\xrightarrow[\text{if } I_n<0 \text{ or } I_m<0]{t\to\infty} 0 \tag{3.56}$$

and

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \mathrm{e}^{i\,\Delta E\, t}\; \mathrm{d}\, t = \left\{ \begin{array}{ll} 1 & , \text{ if } \Delta E = 0 \\ \lim\limits_{T\to\infty} \frac{\mathrm{e}^{i\,\Delta E\, T} - 1}{i\,\Delta E\, T} = 0 & , \text{ else} \end{array} \right\} = \delta_{\Delta\, E, 0}. \tag{3.57}$$

When we combine the equations (3.56) and (3.57), we get

$$\begin{aligned} \lim_{T\to\infty} \frac{1}{T} \int_0^T \left(\frac{U_t(\Theta_{\text{trap}})}{\text{Tr}[\dots]} \right)_{\left(\sum_{\mu=1}^{m-1} s_\mu + j,\, \sum_{\nu=1}^{n-1} s_\nu + k\right)} \mathrm{d}\, t \\ = \delta_{I_n,0}\, \delta_{I_m,0}\, \delta_{R_n,R_m} \cdot \frac{\left(\mathbf{\Theta}^{(m,n)}\right)_{j,\,k}}{\sum\limits_{\mu\in\{1,\dots,M\,:\, I_\mu = 0\}} \left(\mathbf{\Theta}^{m,n}\right)_{\mu,\mu}}. \end{aligned} \tag{3.58}$$

3.9. The stationary solution of the Lindblad equation and differences to the classical case

Now that we have obtained all building blocks for the stationary solution, let us put them together. Our starting point was the quantum jump unravelling (Section 3.2) and the fact that the time and ensemble average can be interchanged (Section B.1) due to the fact that the time average of a single trajectory exists. We assumed that the set of density matrices that occur immediately after quantum jumps is finite ($|\Omega_{\boldsymbol{\rho}_0}| < \infty$), so that a stationary solution of the corresponding Markov chain always exists. We evaluated the time average $\Theta_B := \langle \Theta(t, \omega \,|\, \boldsymbol{\rho}_0) \rangle_{t\geq 0})$ of a single trajectory $\omega \in \mathcal{U}$ (see Equation (3.37)), and we showed in Section 3.5 that it only depends on the minimal absorbing set $B \subseteq \Omega_{\boldsymbol{\rho}_0}$ the quantum trajectory is eventually captured in. We therefore separated the integral over the different unravellings into a sum over the minimal absorbing sets in the network, where the integrand Θ_B is constant over the set $\mathcal{U}(B)$ of quantum trajectories that all land

in the associated minimal absorbing set $B \subseteq \Omega_{\boldsymbol{\rho}_0}$. Formally, this leads to the following steps for obtaining the formula for the stationary density matrix $\boldsymbol{\rho}_\infty(\boldsymbol{\rho}_0)$, depending on the initial state $\boldsymbol{\rho}_0$:

$$\begin{aligned}
\boldsymbol{\rho}_\infty(\boldsymbol{\rho}_0) &= \langle \underbrace{\boldsymbol{\rho}(t \,|\, \boldsymbol{\rho}_0)}_{\langle \Theta(\omega, t \,|\, \boldsymbol{\rho}_0)\rangle_{\omega \in \mathcal{U}}} \rangle_{t \geq 0} \\
&= \langle \langle \Theta(\omega, t \,|\, \boldsymbol{\rho}_0)\rangle_{\omega \in \mathcal{U}} \rangle_{t \geq 0} = \langle \underbrace{\langle \Theta(\omega, t \,|\, \boldsymbol{\rho}_0)\rangle_{t \geq 0}}_{\Theta_\infty(\omega \,|\, \boldsymbol{\rho}_0)} \rangle_{\omega \in \mathcal{U}} \\
&= \int_{\mathcal{U}(\boldsymbol{\rho}_0 \rightsquigarrow)} \underbrace{\Theta_\infty(\omega \,|\, \boldsymbol{\rho}_0)}_{\sum_{B \in \mathcal{B}} 1_{\mathcal{U}(\boldsymbol{\rho}_0 \rightsquigarrow B)}(\omega) \cdot \Theta_B} \mathrm{d}\, \mathcal{P}(\omega \,|\, \boldsymbol{\rho}_0) \\
&= \sum_{B \in \mathcal{B}} \Theta_B \underbrace{\int_{\mathcal{U}(\boldsymbol{\rho}_0 \rightsquigarrow B)} \mathrm{d}\, \mathcal{P}(\omega \,|\, \boldsymbol{\rho}_0)}_{\mathcal{P}(\boldsymbol{\rho}_0 \rightsquigarrow B)} \\
&= \sum_{B \in \mathcal{B}} \mathcal{P}(\boldsymbol{\rho}_0 \rightsquigarrow B)\, \Theta_B.
\end{aligned} \tag{3.59}$$

This expression for stationary solutions of quantum master equations has structural similarities to the stationary solution for classical master equations, given in equation (3.25). Both expressions can be decomposed as sums over all minimal absorbing sets of the state transition network, where each summand is the probability $\mathcal{P}\left(\begin{smallmatrix} \boldsymbol{p}_0 \\ \boldsymbol{\rho}_0 \end{smallmatrix} \rightsquigarrow B\right)$ of a single trajectory to land in this minimal absorbing set (which again depends on the initial condition) times the stationary solution within such a minimal absorbing set.
There are also differences between the classical master equation and the quantum master equation. The states are represented by unit vectors $\boldsymbol{e}_s$ in the classical case an density matrices Θ_s in the quantum case. While a single trajectory of a jump unravelling is constant between two jumps for classical master equations, its time evolution is governed by the conditional Hamiltonian for the Lindblad equation (compare figure 3.4). Moreover the stationary solution of the classical master equation is always limiting, which is not necessarily the case for Lindblad equation. The most important difference however, is the fact, that different states of the classical Master equation are linearly independent (even orthogonal), which need not be the case of Lindblad equations (consider example 3.7.1, with $\boldsymbol{\rho}_0 = \frac{1}{2}\mathbb{1}$).

4. Conclusion and discussion

In this chapter, we summarize our results and give an outlook what future work could look like.

4.1. Summary

In Section 2 we computed an analytical expression for the stationary solutions of Markov chains. The most striking difference between continuous-time Markov chains and discrete-time Markov chains is that while discrete-time Markov chains admit stationary solutions (corresponding to normalized eigenvector of Q to the eigenvalue $\lambda = 1$), they need not have a limiting distribution (compare Figure 2.1), in contrast to continuous-time Markov chains. The algebraic reason for this is the fact that, every non-zero eigenvalue of continuous-time Markov chains has strictly negative real part ($\lambda \neq 0 \Rightarrow \mathrm{Re}\,[\lambda] < 0$) (compare Propositions 1), while discrete-time Markov chains can have eigenvalues with $\lambda \neq 1 = |\lambda|$, where the corresponding eigenvectors could lead to oscillating behavior.
Another reason based on the topology of the state transition network, is the periodicity of the network, where the period of a state $i \in \Omega$ is defined as

$$\text{period}\ (i) := \gcd\left(n \in \mathbb{N} : (Q^n)_{ii} > 0\right), \tag{4.1}$$

where gcd denotes the greatest common divisor.
Note that the period is a class property, meaning that all states in the same strongly connected component have the same period [Bré20; Pri13]. It is possible to define a discrete-time Markov chain for a given continuous-time Markov chain, such that the stationary solution is also attracting, by adding self-loops at every state, while choosing the transition probability between different states to be proportional the corresponding transition rates (see Section A.11 in the appendix for details).
Let us summarize our method as follows: While it is (at least in principle) possible to write down the solution operator $\mathrm{e}^{t\Gamma}$ and compute the limit, finding algebraic solutions for these eigenvalues is not possible for dimensions larger than four [Wüs13]. This is known as the Abel - Ruffini theorem. However, the minimal amount of information needed, is

not the precise value of the eigenvalue, but only the fact that all non-zero eigenvalues have negative real part and are thus declining to zero (compare proposition (1) and following discussion).
When we look at the structure of the transition network and ask where the probability will eventually flow, it leads to the definition of minimal absorbing sets, which can be shown to be strongly connected (see definition 6). This allows us to write the generator Γ in the form of Equation (2.32), where all transient states and minimal absorbing sets are clustered.
The search for a sufficient criterion for the uniqueness of the stationary solution (in which case the network is called relaxing) leads to the positivity of the solution operator (Section 2.3.3), which is guaranteed if the network is strongly connected (Theorem 13). Combining this with the structure of the generator matrix Γ, this leads to the observation that the number of (linearly independent) stationary solutions equals the number of minimal absorbing sets and that for each minimal absorbing set, we can construct a stationary solution, whose non-zero components correspond to the states within this minimal absorbing set (theorem 14).
This leads to the following formula for the stationary solution, which can be interpreted as a variation of Bayes's theorem:

$$\boldsymbol{p}_\infty(\boldsymbol{p}_0) = \sum_{B \in \mathcal{B}} \underbrace{\mathcal{P}(\boldsymbol{p}_0 \rightsquigarrow B)}_{\mathcal{P}(\boldsymbol{p}_0 \,|\, B)} \boldsymbol{p}_\infty(\boldsymbol{p}_0 \in B). \tag{4.2}$$

The i-th entry of the stationary solution $\boldsymbol{p}_\infty(\boldsymbol{p}_0 \in B)$ corresponding to the minimal absorbing set $B \in \mathcal{B}$ is proportional to the sum of all weights of in-trees rooted at state i if $i \in B$ (see Section 2.5.1), while the probability $\mathcal{P}(\boldsymbol{p}_0 \rightsquigarrow B)$ for reaching this minimal absorbing set is the sum over all possible paths in the state transition network of the embedded discrete-time Markov chain (see Section 2.5.2).
Section 2.6 provides a physical explanation for the dynamical behavior of the classical master equation for a strongly connected system: As long as the stationary solution $\boldsymbol{p}_\infty(\boldsymbol{p}_0)$ is not reached (which is never the case for a finite time if the initial condition was not already stationary), the entropy of the system increases, as shown in Equation (2.109). When a stationary solution additionally satisfies the detailed balance condition (see definition 19), it is invariant under time reversal, meaning that at the stationary solution there is no net circular flow (compare Kolmogorov's criterion in theorem 2.6).
In Chapter 3 we start with a derivation of the Lindblad equation for finite dimensions: Starting with the unitary time evolution of system and bath and tracing over the subspace spanning the bath, results in the Kraus operator representation, whose generator can be shown to be of the Lindblad form (see Section 3.1. This theorem has first been proved

by Gorini, Kossakowski, and Sudarshan [BP02] for finite dimensions, with Lindblad extending the proof for bounded operators. The question, of whether an analog generator for unbounded operators must also be of the Lindblad form is still open (date 2022).
Since the computational complexity of the solution scales quadratically with the dimension of the system, analytical solutions of larger systems are practically unfeasible (this has been referred to as the 'curse of dimensionality' [Lid19]). One way to tackle this problem is via unravellings, which are stochastic trajectories, whose ensemble average must equal the solution at a given time. This is usually done in terms of ket-state unravellings, in order to reduce complexity.
In order to avoid having to check that different decompositions

$$\sum_i p_i \, P_{\psi_i(0)} = \boldsymbol{\rho}_0 = \sum_i p'_i \, P_{\psi'_i(0)}$$

lead to the same density matrix in time

$$\sum_i p_i \, P_{\psi_i(t)} = \boldsymbol{\rho}(t \,|\, \boldsymbol{\rho}_0) = \sum_i p'_i \, P_{\psi'_i(t)},$$

we use density matrices as quantum states and the so-called quantum state unravelling , which is what Davies called a piecewise-deterministic process [Dav84].
The stationary solution of the Lindblad equation is given by its time average and we use the fact that Kümmerer and Maassen showed that the time average of a single trajectory exists and one can exchange the time- with the ensemble average (compare [KM04]).
When the number of quantum states directly after the quantum jump is finite, the so-called discrete quantum trajectory defines a classical, discrete-time Markov chain with transition probabilities and minimal absorbing sets, whose stationary probabilities are given in the previous chapter. Moreover, it is possible to find an expression for the time average between two jumps using the Jordan normal form of matrices.
In conclusion, we can say that analytical expressions for the stationary solutions of classical Markov chains on a finite state space exist and have an intuitive picture in the corresponding state transition network.
Stationary solutions of the Lindblad equation can be obtained from the quantum jump unravelling, by exchanging time- and ensemble average and interpreting the quantum states directly after a quantum jump as a Markov chain. In case that the number of these quantum states is finite (which is in particularly guaranteed, when the action of the conditional Hamiltonian on these states is trivial), an explicit expression of the steady state can be computed, using the stationary probabilities of the classical case.

4.2. Outlook

In this Section we want to look ahead and see what future research in this area could look like.
For almost all cases throughout this thesis, we restricted ourselves to the case of a finite system, $|\Omega| < \infty$. The reason why we made this additional assumption is that for finite systems we can work with matrices and use theorems like the one of Perron-Frobenius [Per07; Fro+12] or Gershgorin [Ger31], for which there is no analog for the infinite dimensional case. What is more, having an infinite index set, one has to clarify the space on which this operator is well defined, that is, satisfying the semi-group property. However, given the very specific structure of the set of differential equations, one can be confident that these problems could be overcome, at least in the most common cases [GL05].
There are some limit cases where the solution coming from mathematics would surely coincide with the physical intuition. An example would be an infinite dimensional system Ω_∞, for which there exists a finite sub-system, $\Omega_{<\infty} \subsetneqq \Omega_\infty$, where all the probability mass is contained in. A slightly more general case would be if for all $\epsilon > 0$ there existed a finite sub-system $\Omega_\epsilon \subsetneqq \Omega_\infty$, such that the probability mass in Ω_ϵ is at least $(1 - \epsilon)$ for all times .
On the other hand, a system like the one depicted in Figure 4.1, where the probability mass is just 'flowing towards infinity' will surely not have a limiting probability distribution.

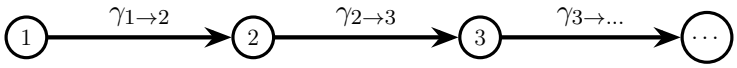


Figure 4.1.: Example of a transition network for a continuous-time Markov chain with an infinite state space: Since the probability $p_i(t)$ for every state converges to zero, there can be no limiting distribution.

There is slightly more material to be found in the literature concerning infinite, discrete-time Markov chains. In this case, an irreducible, recurrent and aperiodic Markov chain is not sufficient to ensure the existence of a limiting distribution - a counterexample would be the symmetric random walk on $\mathbb{Z}$ [Bré20; Pri13; Dou+18]. The reason for this is, that while a single trajectory starting at the origin will almost surely reach the origin again, it will take an infinite amount of time to do so. This phenomena is also known as the 'gambler's ruin' as this indicates that even a fair game (throwing a fair dice and winning or loosing a coin, depending on the outcome) could lead to ruin of the gambler, when the game is played for only a finite time. A Markov chain with this characteristic is called

null recurrent [Bré20; Dou+18; Pri13]. However, when having a finite expectation value of the return time (which is called positive recurrent) rules this out and guarantees the existence of a limiting distribution.
Analytical expressions for classical Markov chains on an infinite state space are also crucial for a possible generalization of our result for quantum master equations, where the assumptions of a finite state space is needed (compare Section B.3).
One has to admit that this assumption is a highly non-generic case, in particular, when the conditional Hamiltonian acts non-trivially on the quantum states, which will then depend on the continuous parameter of the waiting time, resulting (in general) in a state space that is uncountably infinite.
However, it is possible to weaken this assumption a little bit. If trajectories were (even for an infinite state space) eventually captured by some minimal absorbing set, it would most likely suffice if the conditional Hamiltonian were to act trivially only for all the quantum states that lie within some minimal absorbing sets (compare the example in Section B.4).
Another way to generalize this, would be to allow an infinite state space, but requiring that the number of cluster point for every possible unravelling must be finite. Then one were in a position to divide every member of the discrete quantum trajectory into a finite partition of subsequences converging to different limits, where every member of the discrete quantum trajectory belongs to exactly one subsequence. However, this too would not cover the general case, since even a sequence consisting of only a countably infinite number of terms, may have uncountably many cluster points, as can be seen from a rotation by an angle that is no rational multiple of $2\,\pi$.

A. Appendix A: Master equation

A.1. Properties of the transition matrix of a discrete-time - and the solution of a continuous-time Markov chain

Lemma 25 (Eigenvalues of the transition matrix).

The number 1 is an eigenvalue of the transition matrix Q for a discrete-time Markov chain with a finite state space, while all eigenvalues have a modulus less or equal to one

Proof. With $\mathbf{1} := (1, \ldots, 1)^T$ we have

$$\mathbf{1}^T Q = \left(\underbrace{\sum_{i=1}^{|\Omega|} Q_{i,1}}_{1}, \ldots, \underbrace{\sum_{i=1}^{|\Omega|} Q_{i,|\Omega|}}_{1} \right) = \mathbf{1}^T, \tag{A.1}$$

so $Q^T \mathbf{1} = 1 \cdot \mathbf{1}$. This means that 1 is an eigenvalue of both Q^T and (since $\mathsf{det}(Q^T - 1\,\mathbb{1}) = \mathsf{det}\left((Q - 1\,\mathbb{1})^T\right) = \mathsf{det}(Q - 1\,\mathbb{1})$) as well as Q.
Let λ_M be one of the (possible multiple) eigenvalues of Q with the largest absolute value, corresponding the the eigenvector $\boldsymbol{v}$, which we can assume has the one-norm of one. Then we have:

$$|\lambda_M| = |\lambda_M| \underbrace{\|\boldsymbol{v}\|_1}_{1} = \| \underbrace{\lambda_M \boldsymbol{v}}_{Q\,\boldsymbol{v}} \| \leq \underbrace{\|Q\|_1}_{1} \underbrace{\|\boldsymbol{v}\|_1}_{1} = 1. \tag{A.2}$$

□

Lemma 26 (Properties of the solution function).

When $\boldsymbol{p}_0$ is a probability vector, then the solution of the master equation $\boldsymbol{p}(t \,|\, \boldsymbol{p}_0) := \mathrm{e}^{\Gamma\, t}\, \boldsymbol{p}_0$ is also a probability vector for all times $t \geq 0$.

Proof.
When for some state $i \in \{1, \dots, |\Omega|\}$ and some time $t \geq 0$, the probability $p^{(i)}(t)$ of being in that state at that time equals one (zero), then the time derivative $\frac{d}{dt}p^{(i)}(t)$ of that probability is negative (positive), that is

$$\begin{aligned} p^{(i)}(t) = 1 \implies \frac{d}{dt}p^{(i)}(t) &\overset{(2.21)}{=\!=} \sum_{\substack{j=1 \\ j\neq i}}^{|\Omega|} \left(\underbrace{p^{(j)}(t)}_{0}\, \gamma_{j\to i} - \underbrace{p^{(i)}(t)}_{1}\, \gamma_{i\to j} \right) \\ &= -\sum_{j=1}^{|\Omega|} \gamma_{i\to j} \leq 0 \\ p^{(i)}(t) = 0 \implies \frac{d}{dt}p^{(i)}(t) &\overset{(2.21)}{=\!=} \sum_{\substack{j=1 \\ j\neq i}}^{|\Omega|} \left(p^{(j)}(t)\, \gamma_{j\to i} - \underbrace{p^{(i)}(t)}_{0}\, \gamma_{i\to j} \right) \\ &= \sum_{j=1}^{|\Omega|} p^{(j)}(t)\, \gamma_{j\to i} \geq 0. \end{aligned} \tag{A.3}$$

This means, that the solution $\boldsymbol{p}(t \,|\, \boldsymbol{p}_0)$ will always have non-negative entries. In order to show that it is also a probability vector, we notice that the time derivative of the sum of all its components vanishes:

$$\frac{d}{dt}\left(\sum_{i=1}^{|\Omega|} p^{(i)}(t) \right) = \sum_{i=1}^{|\Omega|} \left(\frac{d}{dt}p^{(i)}(t) \right) \overset{(2.21)}{=\!=} \sum_{\substack{i,j=1 \\ i\neq j}}^{|\Omega|} \left(p^{(j)}(t)\, \gamma_{j\to i} - p^{(i)}(t)\, \gamma_{i\to j} \right) = 0. \tag{A.4}$$

This means, that the sum of the components of the solution $\boldsymbol{p}(t \,|\, \boldsymbol{p}_0)$ is constant with time, so $\sum_{i=1}^{|\Omega|} p^{(i)}(t) = \sum_{i=1}^{|\Omega|} p^{(i)}(t=0) = 1$. □

A.2. Gershgorin circle theorem

Lemma 27 (Gershgorin circle theorem and consequences for the eigenvalues of Γ).

Let $\mathcal{B}(z, r) := \{x \in \mathbb{C} : |z - x| \leq r\}$ denote the closed ball around the complex number z with radius $r > 0$. For a $N \times N$ complex matrix ($N \in \mathbb{N}$), the spectrum $\sigma(A)$ of A (that

is the set of all eigenvalues of A) lies within the union of all Gershgorin circles:

$$\sigma(A) \subseteq \bigcup_{i=1}^{N} \mathcal{B}\left(A_{ii}, \sum_{\substack{j=1,\\ j\neq i}}^{N} |A_{ji}| \right)$$

Proof. Let $\boldsymbol{v}$ be an eigenvector of the matrix $A \in \mathbb{C}^{N\times N}$ to the eigenvalue λ, that is $A\,\boldsymbol{v} = \lambda\,\boldsymbol{v}$. We can choose that for some component $i \in \{1,\dots,N\}$ the corresponding entry in the eigenvector equals one, $v_i = 1$, and for all other components $j \in \{1,\dots,N\}$ we have $|v_j| \leq 1$. Now we make the following estimation:

$$\lambda = \lambda \underbrace{v_i}_{1} = (\lambda\,\boldsymbol{v})_i = (A\,\boldsymbol{v})_i = \sum_{\substack{j=1,\\ j\neq i}}^{N} A_{ij}\,v_j + A_{ii} \underbrace{v_i}_{1}$$

$$\Longrightarrow |\lambda - A_{ii}| = \left| \sum_{\substack{j=1,\\ j\neq i}}^{N} A_{ij}\,v_j \right| \leq \sum_{\substack{j=1,\\ j\neq i}}^{N} |A_{ij}| \cdot \underbrace{|v_j|}_{\leq 1} \leq \sum_{\substack{j=1,\\ j\neq i}}^{N} |A_{ij}|$$

□

When we apply the Gershgorin circle theorem to the generator matrix Γ, we get the following result:
Γ has only eigenvalues with a non-positive real part $\Big(\mathrm{Re}\,(\sigma(\Gamma)) \subseteq \mathbb{R}_{\leq 0}\Big)$ and $\lambda = 0$ is the only eigenvalue of Γ where the real part equals zero $\Big(\sigma(\Gamma) \cap (i\,\mathbb{R}) = \{0\}\Big)$.
This follows directly from figure A.1 and the fact that $-\Gamma_{ii} = \sum_{\substack{j=1,\\ j\neq i}}^{N} |\Gamma_{ji}|$ for all $i \in \{1,\dots,N\}$: The origin is the only point in the intersection of the Gershgorin circle and the half-space $\{z \in \mathbb{C} \,|\, \mathrm{Re}(z) \geq 0\}$.

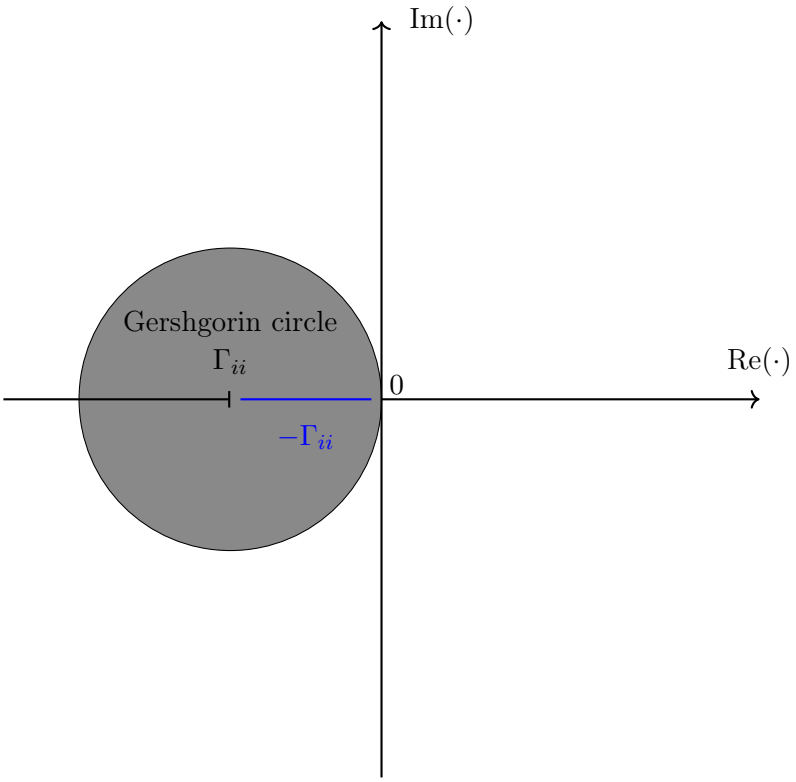


Figure A.1.: The union of all Gershgorin circles containing all eigenvalues of Γ. Hence $\mathrm{Re}(\sigma(\Gamma)) \subseteq \mathbb{R}_{\leq 0}$ and $\sigma(\Gamma) \cap (i\,\mathbb{R}) = \{0\}$.

A.3. Properties of the generator matrix of the master equation

Lemma 28 (Proof of various facts about the generator matrix Γ, listed in Proposition 1).

i) The number zero is an eigenvalue of the generator Γ.

ii) All eigenvalues of the generator matrix Γ apart from zero have strictly negative real part, that is $\mathrm{Re}\,[\sigma(\Gamma)\backslash\{0\}] \subset \mathbb{R}_{<0}$.

iii) The geometric multiplicity for the eigenvalue $\lambda = 0$ agrees with the algebraic multiplicity, that is $g_{\lambda=0} = a_{\lambda=0}$.

iv) Every stationary solution of the master equation lies in the kernel of the generator matrix and vice versa.

v) The column sum of every matrix power of the generator equals zero $\left(\sum_{i=1}^{|\Omega|}(\Gamma^n)_{ij} = 0\right)$, while the column sum of the solution operator equals one $\left(\sum_{i=1}^{|\Omega|}(\mathrm{e}^{\Gamma t})_{ij} = 1\right)$ for all $j \in \{1, \dots, |\Omega|\}$.

Proof. □

i) Since sum of all columns of Γ is zero $\left(\sum_{i=1}^{|\Omega|}\Gamma_{ij} = 0 \text{ for all } j \in \{1, \dots, |\Omega|\}\right)$, the column vectors must be linearly independent and hence Γ can not be invertible, so $\lambda = 0$ must be an eigenvalue. Another way to see this, is to compute

$$\underbrace{(1, \dots, 1)}_{|\Omega| \text{ times}} \cdot \Gamma = \underbrace{(0, \dots, 0)}_{|\Omega| \text{ times}} = 0 \cdot \underbrace{(1, \dots, 1)}_{|\Omega| \text{ times}},$$

which means that $\underbrace{(1, \dots, 1)}_{|\Omega| \text{ times}}^T$ is an eigenvector of Γ^T (and hence also of Γ) to the eigenvalue $\lambda = 0$.

ii) This follows directly from Gershgorin's circle theorem (see Section 27 in the appendix) .

iii) Let $\boldsymbol{h}_{\lambda,m} \in \text{kern } (\Gamma - \lambda \mathbb{1})^m \backslash \text{kern } (\Gamma - \lambda \mathbb{1})^{m-1}$ be a generalized eigenvector of the generator of the transition matrix Γ of rank $m \in \mathbb{N}$ to the eigenvalue λ. The action of the time evolution operator $\mathrm{e}^{\Gamma t}$ on this vector is given by

$$\mathrm{e}^{\Gamma t}\, \boldsymbol{h}_{\lambda,m} = \mathrm{e}^{\lambda t} \sum_{k=0}^{m-1} \frac{t^k}{k!}\, \boldsymbol{h}_{\lambda,m-k}\,, \tag{A.5}$$

where we used the Jordan normal form of the generator Γ (see [HJ12] for details). For non-zero eigenvalues $\lambda \neq 0$, this tends to zero since $\mathrm{Re}[\lambda] < 0$.

For $\lambda = 0$, we must have $m = 1$ since the solution of the master Equation (2.21) is bounded. If we had $m \geq 2$, then for any initial state $\boldsymbol{p}_0 \in (\mathbb{R}_{>0})^{|\Omega|}$ with strictly

positive entries we could choose a small number $\epsilon > 0$ and a normalization constant $\mathcal{N} > 0$ such that $\frac{\boldsymbol{p}_0 + \epsilon\, \boldsymbol{h}_{\lambda=0,m=2}}{\mathcal{N}}$ is a probability vector. The time evolution of this modified initial state would be

$$\mathrm{e}^{\Gamma\, t} \left(\frac{\boldsymbol{p}_0 + \epsilon\, \boldsymbol{h}_{\lambda=0,m=2}}{\mathcal{N}} \right) = \underbrace{\frac{\mathrm{e}^{\Gamma\, t}\, \boldsymbol{p}_0}{\mathcal{N}}}_{\in \left(\mathbb{R}_{\geq 0}\right)^{|\Omega|}} + \frac{\epsilon}{\mathcal{N}} \left(\boldsymbol{h}_{\lambda=0,m=2} + t\, \boldsymbol{h}_{\lambda=0,m=1} \right), \tag{A.6}$$

which is unbounded, in contradiction to the fact that the solution of the master equation remains a probability vector for all times (see lemma 26 in the appendix).

A proof that uses no other properties other than the structure of the generator matrix is given in lemma 44.

This means, that for the eigenvalue $\lambda = 0$ we only have 'normal eigenvectors' of rank $m = 1$: $\boldsymbol{h}^{(i)}_{\lambda=0,m=1} = \boldsymbol{v}^{(i)}_{\lambda=0}$ for $i \in \{1, \ldots, g_{\lambda=0}\}$.

iv) When $\boldsymbol{p} \in \ker\,(\Gamma)$, then $\mathrm{e}^{\Gamma\, t}\, \boldsymbol{p} = \left(\mathbb{1} + \sum\limits_{n \geq 1} \frac{(\Gamma\, t)^n}{n!} \right) \boldsymbol{p} = \boldsymbol{p}$, so $\boldsymbol{p}$ is stationary. When on the other hand, $\boldsymbol{p}$ is stationary, that is $\mathrm{e}^{\Gamma\, t}\, \boldsymbol{p} = \boldsymbol{p}$ for all times $t \geq 0$, we have $\Gamma\, \boldsymbol{p} = \lim\limits_{t \to 0^+} \frac{\mathrm{e}^{\Gamma\, t} - \mathbb{1}}{t}\, \boldsymbol{p} = 0$, so $\boldsymbol{p}$ lies in the kernel of Γ.

v) This is proven by induction over the matrix power $n \in \mathbb{N}_{\geq 1}$. For $n = 1$ this follows from the definition of the generator matrix Γ. For the induction step, we have:

$$\sum_{i=1}^{|\Omega|} \underbrace{(\Gamma^{n+1})_{ij}}_{\sum\limits_{k=1}^{|\Omega|} (\Gamma^n)_{ik}\, \Gamma_{kj}} = \sum_{k=1}^{|\Omega|} \Gamma_{kj} \underbrace{\sum_{i=1}^{|\Omega|} (\Gamma^n)_{ik}}_{0} = 0, \tag{A.7}$$

while for the solution operator we use its representation as a power series

$$\sum_{i=1}^{|\Omega|} \underbrace{(\mathrm{e}^{\Gamma\, t})_{ij}}_{\sum\limits_{n \in \mathbb{N}_0} \frac{(\Gamma\, t)^n}{n!}} = \sum_{n \in \mathbb{N}_0} \frac{t^n}{n!} \underbrace{\sum_{i=1}^{|\Omega|} (\Gamma^n)_{ij}}_{\delta_{n,0}} = 1. \tag{A.8}$$

A.4. Weakly chained diagonal dominant (WCDD) matrices

Lemma 29 (WCDD matrices are non-singular).

Proof. The fact that SDD matrices are non-singular follows from Gershgorin's circle theorem [HW06] (for a proof, see Section 27 in the Appendix).
Now let $A \in \mathbb{C}^{N\times N}$ be WCDD and assume that A is singular. Then A as well as A^T has an eigenvalue $\lambda = 0$ and a corresponding eigenvector $\boldsymbol{x} \in \ker(A^T - 0\cdot\mathbb{1})$, w.l.o.g. assume that there is an $i \in \{1,\dots,N\}$ such that $1 = |x_i| \geq |x_k|\, \forall k \in \{1,\dots,N\},\, k \neq i$. Then we have

$$0 = (A^T\boldsymbol{x})_i = \sum_{\substack{k=1,\\ k\neq i}}^{N} \underbrace{(A^T)_{ik}}_{A_{ki}} x_k + A_{ii}\, x_i \implies -A_{ii}\, x_i = \sum_{\substack{k=1,\\ k\neq i}}^{N} A_{ki}\, x_k \tag{A.9}$$

$$\implies |A_{ii}| = |-A_{ii}\, x_i| \overset{(A.9)}{=} \left|\sum_{\substack{k=1,\\ k\neq i}}^{N} A_{ki}\, x_k\right| \leq \sum_{\substack{k=1,\\ k\neq i}}^{N} |A_{ki}| \underbrace{|x_k|}_{\leq 1} \overset{(*)}{\leq} \sum_{\substack{k=1,\\ k\neq i}}^{N} |A_{ki}| \overset{A\dots\text{WDD}}{\leq} |A_{ii}| \tag{A.10}$$

Hence, in line (A.10) we have equality everywhere. In particular:

i) The last equality tells us that the i-the row is not SDD.

ii) Equality in (*) means that whenever $A_{ki} \neq 0 \implies |x_k| = 1$.

Since A is WCDD we know there exists a path $i = i_0 \to i_1 \to \cdots \to i_k = j$ to the SDD row number j. In particular, we have $A_{i_1,i_0} \neq 0 \overset{ii)}{\implies} |x_{i_1}| = 1$. Repeating the argument from the beginning, we get from i) that the i_1-th row is not SDD. When we keep iterating, we finally get, that the j-th row is not SDD which is a contradiction. □

A.5. The solution operator - positivity and long-term behavior

Lemma 30 (Positivity of entries of the solution operator and the reachability of the corresponding states).

Given a network $\mathcal{S} = (\Omega, \mathcal{E}, \gamma)$ with two states $i, j \in \Omega$ such that state i is reachable from state j, $j \rightsquigarrow i$, then the i-j-th entry of the solution operator is strictly positive, $\left(\mathrm{e}^{t\,\Gamma}\right)_{ij} > 0$ for all $t > 0$.
In particular, for a strongly connected network $\mathcal{S}$, the solution operator has only strictly positive entries for all times, $\mathrm{e}^{t\,\Gamma} \in (\mathbb{R}_{>0})^{|\Omega|\times|\Omega|}$ for all $t > 0$.

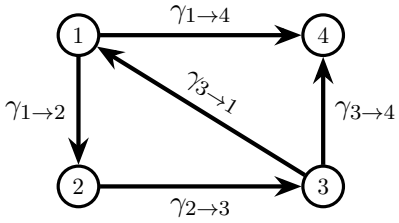
Proof. □

Let $i, j \in \Omega$ be states in the network, such that state i is reachable from state j, $j \rightsquigarrow i$. Since we know that the limit $\mathrm{e}^{t\,\Gamma} = \lim_{n\to\infty} (\mathbb{1} + \frac{t\,\Gamma}{n})^n$ exists, it suffices to show that that the sequence

$$\left(\left(\left(\mathbb{1} + \frac{t\,\Gamma}{n}\right)^n\right)_{ij}\right)_{n\in\mathbb{N}} = \sum_{k_1=1}^{|\Omega|} \cdots \sum_{k_{n-1}=1}^{|\Omega|} \left(\mathbb{1} + \frac{\Gamma\, t}{n}\right)_{i,\,k_1} \cdots \left(\mathbb{1} + \frac{\Gamma\, t}{n}\right)_{k_{n-1},\,j} \tag{A.11}$$

has a strictly positive lower bound.

(a) Original network $\mathcal{S}$



(b) Network $\tilde{\mathcal{S}}_n$ associated to $\mathbb{1} + \frac{t\,\Gamma}{n}$

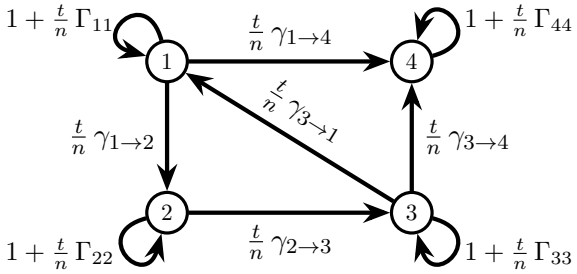


Figure A.2.: Illustrating the difference between the original network $\mathcal{S}$ and network $\tilde{S}_n$ associated to the matrix $\mathbb{1} + \frac{t\,\Gamma}{n}$, with the modified link strength and additional self-loops.
The walk $(1, 1, 2, 3, 3, 4) \in \mathcal{WP}(1 \xrightarrow{5} 4,\, \tilde{S}_n) \subseteq \mathcal{W}(1 \xrightarrow{5} 4, \tilde{S}_n)$, after removing the self-loops, becomes a *path*, where as the walk $(1, 2, 3, 1, 2, 3, 4) \in \mathcal{W}(1 \xrightarrow{6} 4, \tilde{S}_n) \backslash \mathcal{WP}(1 \xrightarrow{6} 4, \tilde{S}_n)$ does not.

We call $\tilde{\mathcal{S}}_n := (\Omega, \tilde{\mathcal{E}}, \tilde{\gamma})$ the network associated to the matrix $\mathbb{1} + \frac{t\,\Gamma}{n}$, provided that $n \in \mathbb{N}$ is large enough, such that all links are positive (in particular, $1 + \frac{t}{n}\,\Gamma_{ii}$ is positive for all $i \in \Omega$). It can be recovered from the original network $\mathcal{S}$, by modifying the links between different states with a factor of $\frac{t}{n}$ and adding a self-loop at every state with a weight of $1 + \frac{t\,\Gamma_{ss}}{n}$ for all $s \in \Omega$.

We need the following definitions:

$$
\begin{aligned}
\Gamma_{\min} :=& \min\{\Gamma_{ll} \,:\, l \in \{1,\dots,|\Omega|\}\} < 0 \\
\mathcal{W}(j \xrightarrow{n} i,\, \mathcal{S}) :=& \{ \text{ all walks } \omega \in \Omega^{n+1} \text{ of length } n \in \mathbb{N} \text{ from state } j \in \Omega \\
& \quad \text{to state } i \in \Omega \text{ in the network } \mathcal{S}\} \\
\mathcal{P}(j \xrightarrow{n} i,\, \mathcal{S}) :=& \{ \text{ all paths } \omega \in \Omega^{n+1} \text{ of length } n \in \mathbb{N} \text{ from state } j \in \Omega \\
& \quad \text{to state } i \in \Omega \text{ in the network } \mathcal{S}\} \\
\mathcal{WP}(j \xrightarrow{n} i,\, \mathcal{S}) :=& \{ \text{ all walks } \omega \in \Omega^{n+1} \text{ of length } n \in \mathbb{N} \text{ from state } j \in \Omega \\
& \quad \text{to state } i \in \Omega \text{ in the network } \mathcal{S}, \\
& \quad \text{which become paths after removing the self-loops}\} \\
\mathcal{P}(j \rightsquigarrow i,\, \mathcal{S}) :=& \bigcup_{n\in\mathbb{N}} \mathcal{P}(j \xrightarrow{n} i,\, \mathcal{S}) \\
=& \{ \text{ all paths of arbitrary length from state } j \in \Omega \\
& \quad \text{to state } i \in \Omega \text{ in the network } \mathcal{S}
\end{aligned}
\tag{A.12}
$$

Clearly, we have

$$\mathcal{P}(j \xrightarrow{n} i,\, \mathcal{S}) \subseteq \mathcal{WP}(j \xrightarrow{n} i,\, \mathcal{S}) \subseteq \mathcal{W}(j \xrightarrow{n} i,\, \mathcal{S}).$$

A non-zero summand in the right hand side of Equation (A.11) can be interpreted as the weight $\gamma_{\tilde{\omega}}$ of a walk $\tilde{\omega} \in \mathcal{W}(j \xrightarrow{n} i,\, \tilde{\mathcal{S}}_n)$ of length $|\tilde{\omega}| = n$ from j to i.
When we look only at the weight of a special walk

$$\tilde{\omega} \in \mathcal{WP}(j \xrightarrow{|\tilde{\omega}|} i, \tilde{\mathcal{S}}_n) \subseteq \mathcal{W}(j \xrightarrow{|\tilde{\omega}|} i, \tilde{\mathcal{S}}_n),$$

then this weight can be separated into the weight of the corresponding path $\omega \in \mathcal{P}(j \xrightarrow{|\omega|} i,\, \mathcal{S})$ times a scaling factor $\left(\frac{t}{n}\right)^{|\omega|}$ times the weight of the self-loops:

$$
\begin{aligned}
\gamma_{\tilde{\omega}} = \prod_{\alpha=1}^{n} \gamma_{\tilde{\omega}_\alpha \to \tilde{\omega}_{\alpha+1}} &= \underbrace{\prod_{k=1}^{|\omega|} \frac{t}{n}\, \gamma_{\omega_k \to \omega_{k+1}}}_{= \left(\frac{t}{n}\right)^{|\omega|} \gamma_\omega} \cdot \underbrace{\prod_{s \in \{\text{self-loops}(\tilde{\omega})\}} \left(1 + \frac{t\,\Gamma_{ss}}{n}\right)}_{\geq \left(1+\frac{t\,\Gamma_{\min}}{n}\right)^{|\tilde{\omega}|-|\omega|}} \geq \\
&\geq \gamma_\omega \left(\frac{t}{n}\right)^{|\omega|} \left(1 + \frac{t\,\Gamma_{\min}}{n}\right)^{|\tilde{\omega}|-|\omega|},
\end{aligned}
\tag{A.13}
$$

where the product is taken over all self-loops of $\tilde{\omega}$.

(a) The *walk*
$\tilde{\omega} = (1,1,2,3,3,4) \in \mathcal{WP}(1 \xrightarrow{|\tilde{\omega}|=5} 4, \tilde{\mathcal{S}}_n)$

(b) The *path*
$\omega = (1,2,3,4) \in \mathcal{P}(1 \xrightarrow{|\omega|=3} 4, \tilde{\mathcal{S}}_n)$
and self-loops.

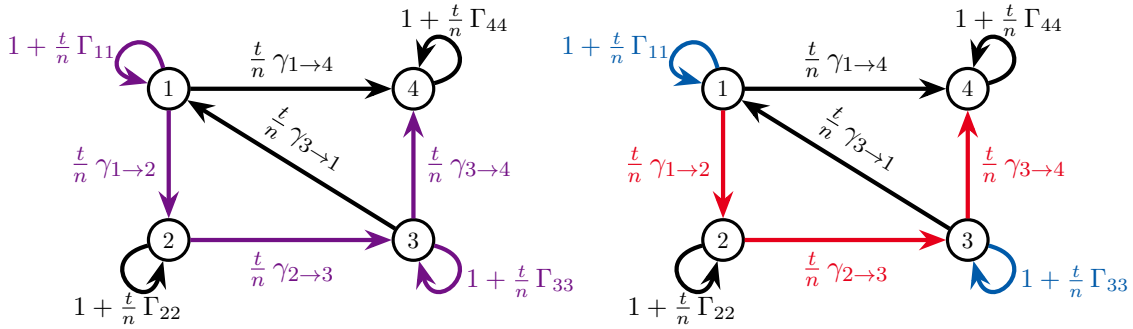


Figure A.3.: Illustrating the separation of the weight of a *walk* $\tilde{\omega} \in \mathcal{WP}(1 \xrightarrow{|\tilde{\omega}|=5} 4, \tilde{\mathcal{S}}_n)$ into the weight of a *path* $\omega \in \mathcal{P}(1 \xrightarrow{|\omega|=3} 4, \tilde{\mathcal{S}}_n)$ and a product of the weights of self-loops.
For a fixed path $\omega = (1,2,3,4) \in \mathcal{P}(1 \xrightarrow{|\omega|=3} 4, \tilde{\mathcal{S}}_n)$, there are $\binom{|\tilde{\omega}|=5}{|\omega|=3} = 10$ many $\tilde{\omega} \in \mathcal{WP}(1 \xrightarrow{|\tilde{\omega}|=5} 4, \tilde{\mathcal{S}}_n)$ that include the original path, with only additional self-loops, namely
$\tilde{\omega}_1 = (1,1,1,2,3,4)$, $\tilde{\omega}_2 = (1,1,2,2,3,4)$, $\tilde{\omega}_3 = (1,1,2,3,3,4)$,
$\tilde{\omega}_4 = (1,1,2,3,4,4)$, $\tilde{\omega}_5 = (1,2,2,2,3,4)$, $\tilde{\omega}_6 = (1,2,2,3,3,4)$,
$\tilde{\omega}_7 = (1,2,2,3,4,4)$, $\tilde{\omega}_8 = (1,2,3,3,3,4)$, $\tilde{\omega}_9 = (1,2,3,3,4,4)$,
and $\tilde{\omega}_{10} = (1,2,3,4,4,4)$.

For a given path $\omega \in \mathcal{P}(j \xrightarrow{|\omega|} i)$ and some fixed natural number $|\tilde{\omega}|$ greater or equal to $|\omega|$, $|\tilde{\omega}| \in \mathbb{N}_{\geq |\omega|}$, there are $\binom{|\tilde{\omega}|}{|\omega|}$ many walks $\tilde{\omega} \in \mathcal{WP}(j \xrightarrow{|\tilde{\omega}|} i)$, which include the original path, but have additional self-loops (out of $|\tilde{\omega}|$ many transitions, choose the positions of $|\omega|$ non-trivial ones).
For a last estimation, we notice that the following two sequences are both monotonously increasing (denoted by $\nearrow$) and converge, which allows us to make the following estimation for a fixed, but sufficiently small number $\epsilon_0 > 0$ and for sufficiently large number $n \in \mathbb{N}_{\geq |\Omega|}$

:

$$\left.\begin{array}{l} \frac{n!}{(n-|\omega|)!\cdot n^{|\omega|}} \xrightarrow{n\to\infty} 1 \\ \left(\frac{n!}{(n-|\omega|)!\cdot n^{|\omega|}}\right)_{n\in\mathbb{N}} \nearrow \end{array}\right\} \Longrightarrow \frac{n!}{(n-|\omega|)!\cdot n^{|\omega|}} \quad \geq (1-\epsilon_0) \tag{A.14}$$

$$\left.\begin{array}{l} \left(1+\frac{t\,\Gamma_{\min}}{n}\right)^{n-|\omega|} \xrightarrow{n\to\infty} \mathrm{e}^{t\,\Gamma_{\min}} \\ \left(\left(1+\frac{t\,\Gamma_{\min}}{n}\right)^{n-|\omega|}\right)_{n\in\mathbb{N}} \nearrow \end{array}\right\} \Longrightarrow \left(1+\frac{t\,\Gamma_{\min}}{n}\right)^{n-|\omega|} \quad \geq (1-\epsilon_0)\,\mathrm{e}^{t\,\Gamma_{\min}}$$

$$\begin{aligned}
\left(\left(\mathbb{1}+\frac{t\,\Gamma}{n}\right)^{n}\right)_{ij} &= \sum_{k_1=1}^{|\Omega|}\cdots\sum_{k_{n-1}=1}^{|\Omega|}\left(\mathbb{1}+\frac{\Gamma\,t}{n}\right)_{i,\,k_1}\cdot\ldots\cdot\left(\mathbb{1}+\frac{\Gamma\,t}{n}\right)_{k_{n-1},\,j} \\
&= \sum_{\tilde{\omega}\in\mathcal{W}\left(j\xrightarrow{n} i,\,\tilde{\mathcal{S}}_n\right)} \gamma_{\tilde{\omega}} \\
&\overset{\mathcal{W}\supseteq\mathcal{WP}}{\geq} \sum_{\tilde{\omega}\in\mathcal{WP}\left(j\xrightarrow{n} i,\,\tilde{\mathcal{S}}_n\right)} \underbrace{\gamma_{\tilde{\omega}}}_{\geq\gamma_\omega\left(\frac{t}{n}\right)^{|\omega|}\left(1+\frac{t\,\Gamma_{\min}}{n}\right)^{n-|\omega|}} \\
&\overset{(A.13)}{\geq} \sum_{\omega\in\mathcal{P}\left(j\rightsquigarrow i,\,\tilde{\mathcal{S}}_n\right)} \binom{n}{|\omega|}\gamma_\omega\left(\frac{t}{n}\right)^{|\omega|}\left(1+\frac{t\,\Gamma_{\min}}{n}\right)^{n-|\omega|} \\
&= \sum_{\omega\in\mathcal{P}\left(j\rightsquigarrow i,\,\tilde{\mathcal{S}}_n\right)} \underbrace{\frac{n!}{(n-|\omega|)!\,n^{|\omega|}}}_{\geq(1-\epsilon_0)} \frac{\gamma_\omega\,t^{|\omega|}}{|\omega|!} \underbrace{\left(1+\frac{t\,\Gamma_{\min}}{n}\right)^{n-|\omega|}}_{\geq(1-\epsilon_0)\,\mathrm{e}^{t\,\Gamma_{\min}}} \\
&\overset{(A.14)}{\geq} \frac{(1-\epsilon_0)^2}{2}\,\mathrm{e}^{t\,\Gamma_{\min}} \sum_{\omega\in\mathcal{P}\left(j\rightsquigarrow i,\,\tilde{\mathcal{S}}_n\right)} \frac{\gamma_\omega\,t^{|\omega|}}{|\omega|!} \\
&> 0.
\end{aligned} \tag{A.15}$$

In the fourth step, we used both the estimate of Equation (A.13) and the fact that every 'walk-path' can be separated into a path and self-loops and in the last step the fact that $|\omega| \leq |\Omega| \leq n$.
This concludes the proof.

Lemma 31 (The structure of the 'stationary' solution operator $\lim_{t\to\infty} \mathrm{e}^{t\,\Gamma}$).

Let $B \in \mathcal{B}$ be a minimal absorbing set and $\boldsymbol{p}_B$ the corresponding stationary state, given by Equation (2.77). Then $\boldsymbol{p}_B \cdot (\underbrace{1,\ldots,1}_{|B|})$ is the matrix, whose columns are the vector $\boldsymbol{p}_B$:

$$\boldsymbol{p}_B \cdot (\underbrace{1,\ldots,1}_{|B|}) = \begin{pmatrix} | & \ldots & | \\ \boldsymbol{p}_B & \ldots & \boldsymbol{p}_B \\ | & \ldots & | \end{pmatrix} = \begin{pmatrix} \boldsymbol{p}_B^{(1)} & \cdots & \boldsymbol{p}_B^{(1)} \\ \vdots & \ldots & \vdots \\ \boldsymbol{p}_B^{(|B|)} & \ldots & \boldsymbol{p}_B^{(|B|)} \end{pmatrix} \in (\mathbb{R}_{>0})^{|B|\times|B|} \text{ with}$$

$$\boldsymbol{p}_B = \begin{pmatrix} \sum_{T\in\mathcal{T}(\to 1,B)} \gamma_T \\ \vdots \\ \sum_{T\in\mathcal{T}(\to |B|,B)} \gamma_T \end{pmatrix} \tag{A.16}$$

$$\lim_{t\to\infty} \mathrm{e}^{\Gamma\, t} = \begin{pmatrix} \mathbf{0}^{|B_0|\times|\Omega|} & & & \\ \mathcal{P}(B_0 \to B_1)\, \boldsymbol{p}_{B_1} \cdot (\underbrace{1,\ldots,1}_{|B_1|}) & \boldsymbol{p}_{B_1} \cdot (\underbrace{1,\ldots,1}_{|B_1|}) & & \\ & & \ddots & \\ \mathcal{P}(B_0 \to B_{|\mathcal{B}|})\, \boldsymbol{p}_{B_{|\mathcal{B}|}} \cdot (\underbrace{1,\ldots,1}_{|B_1|}) & & & \boldsymbol{p}_{B_{|\mathcal{B}|}} \cdot (\underbrace{1,\ldots,1}_{B_{|\mathcal{B}|}|}) \end{pmatrix} =$$

$$= \begin{pmatrix} 0 & & \ldots & \ldots & 0 \\ \vdots & & \ldots & \ldots & \vdots \\ 0 & & \ldots & \ldots & 0 \\ \mathcal{P}\,(B_0 \to B_1) \begin{pmatrix} | & \ldots & | \\ \boldsymbol{p}_{B_1} & \ldots & \boldsymbol{p}_{B_1} \\ | & \ldots & | \end{pmatrix} & \begin{pmatrix} | & \ldots & | \\ \boldsymbol{p}_{B_1} & \ldots & \boldsymbol{p}_{B_1} \\ | & \ldots & | \end{pmatrix} & & \\ \vdots & & \ddots & \\ \mathcal{P}\,(B_0 \to B_{|\mathcal{B}|}) \begin{pmatrix} | & \ldots & | \\ \boldsymbol{p}_{B_{|\mathcal{B}|}} & \ldots & \boldsymbol{p}_{B_{|\mathcal{B}|}} \\ | & \ldots & | \end{pmatrix} & & & \begin{pmatrix} | & \ldots & | \\ \boldsymbol{p}_{B_{|\mathcal{B}|}} & \ldots & \boldsymbol{p}_{B_{|\mathcal{B}|}} \\ | & \ldots & | \end{pmatrix} \end{pmatrix}$$

Remark 32 (Non-negative entries of $\mathrm{e}^{t\,\Gamma}$ are not sufficient).

Let $v \in \mathbb{R}^N$ be a vector and $A \in (\mathbb{R}_{\geq 0})^{N \times N}$ is a matrix with non-negative entries such that for all $i \in \{1, \ldots, N\}$ we have $(A_{i,1}\, v_1, \ldots, A_{i,N}\, v_N) \in (\mathbb{R}_{\geq 0})^N \cup (\mathbb{R}_{\leq 0})^N$. Then we can not conclude, that $\boldsymbol{v} \in (\mathbb{R}_{\geq 0})^N \cup (\mathbb{R}_{\leq 0})^N$, as the following example show:

$$
\begin{aligned}
A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \boldsymbol{v} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\
\left(\underbrace{A_{1,1}}_{0}\, v_1, \underbrace{A_{1,2}}_{1}\, \underbrace{v_2}_{-1} \right) = (0, -1) \in (\mathbb{R}_{\leq 0})^{N=2} \text{ and} \\
\left(\underbrace{A_{2,1}}_{1}\, \underbrace{v_1}_{1}, \underbrace{A_{2,2}}_{1}\, v_2 \right) = (1, 0) \in (\mathbb{R}_{\geq 0})^{N=2}
\end{aligned}
\tag{A.17}
$$

Lemma 33 (Any linear subspace of $\mathbb{R}^N$ contained in $(\mathbb{R}_{\geq 0})^N \cup (\mathbb{R}_{\leq 0})^N$ must be on-dimensional).

Proof. □

An abstract argument is, that the vector space, when removing the zero vector, is no longer path connected, and this is only the case for one-dimensional vector spaces.
A technical proof using only tools from linear algebra is given below. The idea is to use assume the vector space were at least two dimensional and then construct a vector with a strictly positive and a strictly negative entry.
So let $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ be two linear independent vectors and without loss of generality we assume that both the two vectors as well as their difference have only non-negative entries, $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_1 - \boldsymbol{v}_2 \in (\mathbb{R}_{\geq 0})^N$.
Then for all components $i \in \{1, \ldots, N\}$ we can choose an $\epsilon_i \in [0, 1]$ such that $\epsilon_i\, v_1^{(1)} = v_2^{(1)}$. We define $\epsilon_m\, (\epsilon_M)$ be the smallest (largest) of these epsilons and $i_m\, (i_M)$ to be the corresponding index, that is

$$
\begin{aligned}
\epsilon_m &:= \mathsf{min}\{\, \epsilon_i \,:\, i \in \{1, \ldots, N\}\,\} & i_m &:= arg\, \mathsf{min}\{\, \epsilon_i \,:\, i \in \{1, \ldots, N\}\,\} \\
\epsilon_M &:= \mathsf{max}\{\, \epsilon_i \,:\, i \in \{1, \ldots, N\}\,\} & i_M &:= arg\, \mathsf{max}\{\, \epsilon_i \,:\, i \in \{1, \ldots, N\}\,\}.
\end{aligned}
\tag{A.18}
$$

Since $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are linearly independent, we know that $\epsilon_m < \epsilon_M$. Then for any $\epsilon \in (\epsilon_m, \epsilon_M)$, that is $\epsilon_m < \epsilon < \epsilon_M$, the vector $\epsilon\, \boldsymbol{v}_1 - \boldsymbol{v}_2$ has both strictly positive as well as strictly negative components:

$$\begin{aligned}
\text{component } i_M: \underbrace{\epsilon}_{<\epsilon_M} v_1^{(i_M)} - v_2^{(i_M)} &< \epsilon_M\, v_1^{(i_M)} - v_2^{(i_M)} \overset{\text{Def } \epsilon_M, i_M}{=\!=\!=} 0 \\
\text{component } i_m: \underbrace{\epsilon}_{>\epsilon_m} v_1^{(i_m)} - v_2^{(i_m)} &> \epsilon_m\, v_1^{(i_m)} - v_2^{(i_m)} \overset{\text{Def } \epsilon_m, i_m}{=\!=\!=} 0.
\end{aligned} \tag{A.19}$$

This is a contradiction to the assumption that two linearly independent vectors exists, which concludes the proof.

A.6. In-trees and in-forests of a network

Definition 34 (In-trees).

We call a network $T = (\Omega, \mathcal{E}_T)$ with no self-loops an in-tree (also called anti-arborescence [GM78]) rooted at state $i_0 \in \Omega$ if for all states $i \in \Omega$ there is a unique directed walk leading from state i towards the root i_0. An example is given in Figure A.4, where the root is state number $i_0 = 2$. Note, that since the walk is unique, it has to be a path, meanig that all states must be different. The weight of an in-tree is the product of the link-strength of all edges of the in-tree: $\gamma_T := \prod_{(i,j)\in\mathcal{E}_T} \gamma_{i\to j}$.

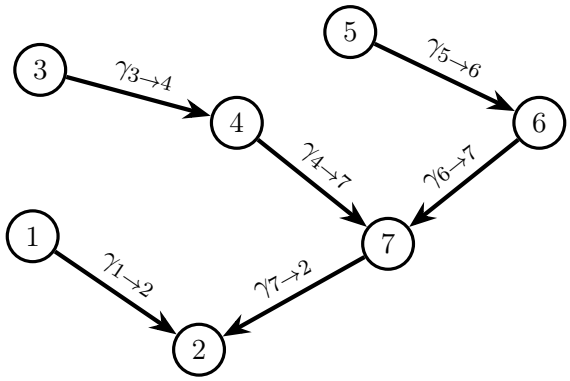


Figure A.4.: Example of an in-tree rooted at state number 2. There is a unique directed walk from each state of the tree leading to state number 2. Its weight equals $\gamma_T = \prod_{(i,j)\in\mathcal{E}(T)} \gamma_{i\to j} = \gamma_{1\to2}\,\gamma_{7\to2}\,\gamma_{7\to2}\,\gamma_{4\to7}\,\gamma_{6\to7}\,\gamma_{3\to4}\,\gamma_{5\to6}$

Definition 35 (In-forests).

We call a network $F_{\to B} = (\Omega, \mathcal{E}_F)$ an in-forest rooted at the states $B \subseteq \Omega$ if its weakly connected components are in-trees, that is $F_{\to B} = \bigoplus_{b \in B} T_{\to b}$. In-trees are special in-forests, with exactly one weakly connected component.
This means that for all states $i \in \Omega$ there is a unique state $b \in B$ and a unique directed path leading from state i towards this state $b \in B$ (for $i \in B$, this would be the trivial walk (i)).
In particular, this means that the set $\{\mathcal{R}(\to b) : b \in B\}$ is a partition of Ω, that is

$$\begin{aligned} &\bigcup_{b \in B} \mathcal{R}(\to b) = \Omega \text{ and} \\ &\mathcal{R}(\to b_1) \cap \mathcal{R}(\to b_2) = \emptyset, \text{ for } b_1 \neq b_2. \end{aligned} \tag{A.20}$$

Analogously to in-trees, the weight of an in-forest is defined as the product over the links of its edges: $\gamma_F := \prod_{(i,j) \in \mathcal{E}_F} \gamma_{i \to j} = \prod_{b \in B} \gamma_{T \to b}$.
An example is given in Figure A.5, where the roots are given by $B = \{2, 3, 6\}$.

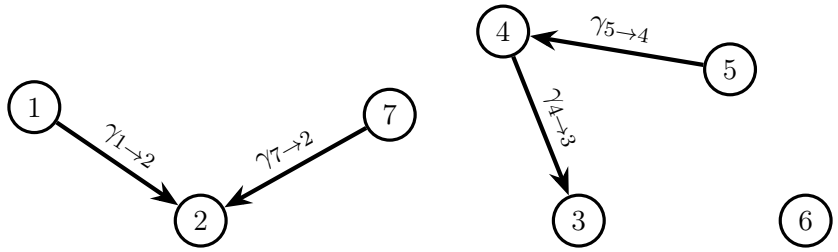


Figure A.5.: Example of an in-forest rooted at the set $B = \{2, 3, 6\}$. There is a unique directed path from each state of the tree leading to either one of the states 2, 3 or 6. Its weight equals
$\gamma_F = \prod_{(i,j) \in \mathcal{E}_F} \gamma_{i \to j} = \gamma_{T_{\to 2}} \cdot \gamma_{T_{\to 3}} \cdot \gamma_{T_{\to 6}} = \gamma_{1\to 2}\, \gamma_{7\to 2} \cdot \gamma_{5\to 4}\, \gamma_{4\to 3}$.

Further, we define the set of possible in-trees and in-forests of a given network as follows:

Definition 36 (The set of partitions of a set, the set of possible in-trees and of possible in-forests of a network).

Le Ω be any set and $\emptyset \neq J \subseteq \Omega$ be a non-empty subset. We denote with $\mathrm{Part}(\Omega, J \subseteq \Omega)$ the set of all partitions of Ω into $|J|$ many elements Ω_j (that is $\dot{\bigcup}_{j \in J} \Omega_j = \Omega$) such that $j \in \Omega_j$ for all $j \in J$.

Let $\mathcal{S} = (\Omega, \mathcal{E})$ be a network with $i_0 \in \Omega$ and $J \subseteq \Omega$, $J \neq \emptyset$. Then

$$\begin{aligned}
\mathcal{T}(\to i_0, \mathcal{S}) &:= \{(\Omega, \mathcal{E}_T) \subseteq \mathcal{S} \,:\, (\Omega, \mathcal{E}_T) \text{ is an in-tree rooted at the state } i_0 \in \Omega \,\} \\
\mathcal{F}(\to J, \mathcal{S}) &:= \{(\Omega, \mathcal{E}_F) \subseteq \mathcal{S} \,:\, (\Omega, \mathcal{E}_F) \text{ is an in-forest rooted at the states } B \subseteq \Omega \,\} \\
&= \bigcup_{\substack{\{\Omega_j : j \in J\} \in \\ \mathrm{Part}(\Omega, J \subseteq \Omega)}} \quad \bigcup_{\substack{(\mathcal{E}_{T_1}, \ldots, \mathcal{E}_{T_{|J|}}) \in \bigoplus_{j \in J} \mathcal{E}_{\Omega_j} \\ (\Omega_j, \mathcal{E}_{T_j}) \in \mathcal{T}(\to j, (\Omega_j, \mathcal{E}_{\Omega_j}))}} \left\{ \bigoplus_{j \in J} (\Omega_j, \mathcal{E}_{T_j}) \right\}, \\
&\text{with } \mathcal{E}_{\Omega_j} := \{(a, b) \in \Omega_j^2 \,:\, (a, b) \in \mathcal{E}\}.
\end{aligned} \tag{A.21}$$

It should be noted, that an in-tree T of a network $\mathcal{S}$ need not be a subnetwork defined in 4. In fact, T can only a subnetwork of original network $\mathcal{S}$, if $\mathcal{S}$ itself is an in-tree.
An illustration is given in the figures A.6 and A.7.

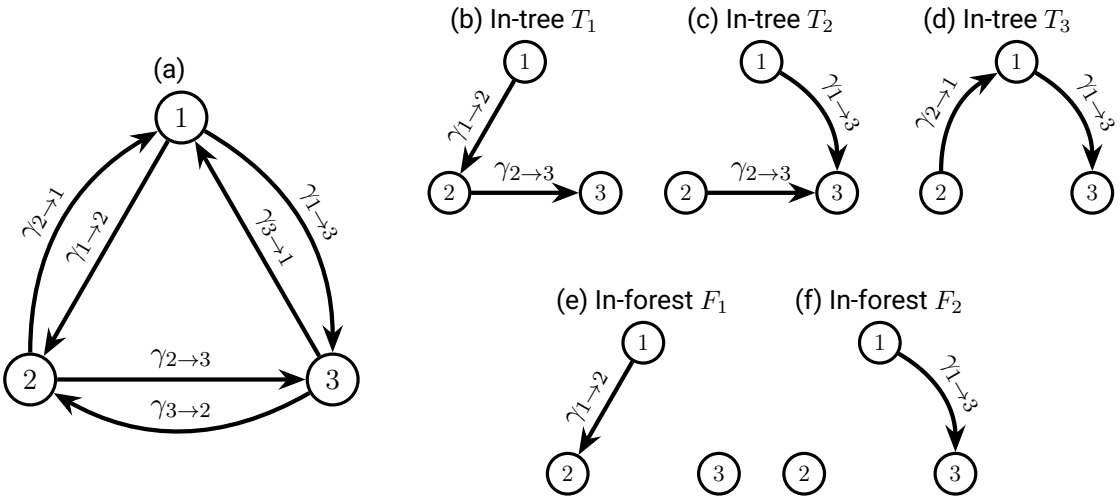
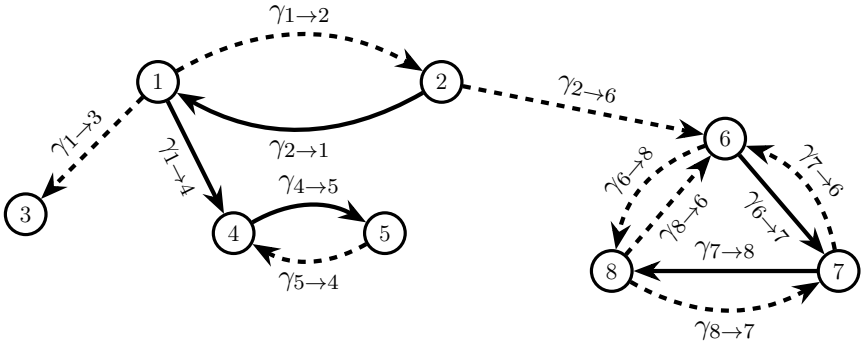
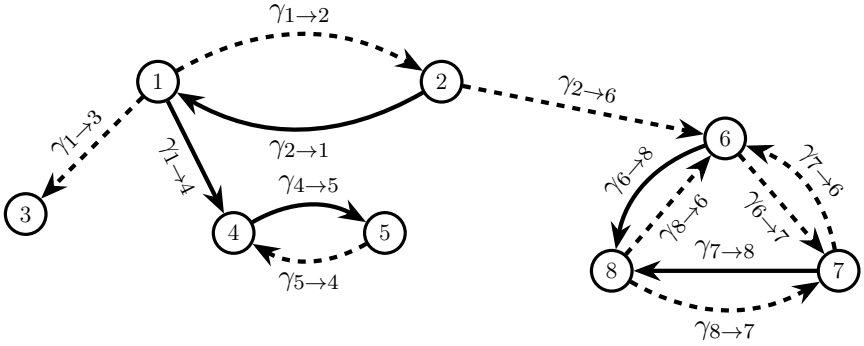


Figure A.6.: Example of a network, whose in-trees rooted at state 3 are given by $\mathcal{T}(\to 3, \mathcal{S}) = \{T_1, T_2, T_3\}$ and whose in-forests rooted at the states $\{2, 3\}$ are given by $\mathcal{F}(\to \{2, 3\}, \mathcal{S}) = \{F_1, F_2\}$.

(a) $(\mathcal{E}_{T_3}, \mathcal{E}_{T_5}, \mathcal{E}_{T_8}) = (\{\}, \{(2,1),\ (1,3),\ (4,5)\}, \{(6,7),\ (7,8)\})$



(b) $(\mathcal{E}_{T_3}, \mathcal{E}_{T_5}, \mathcal{E}_{T_8}) = (\{\}, \{(2,1),\ (1,3),\ (4,5)\}, \{(6,8),\ (7,8)\})$



(c) $(\mathcal{E}_{T_3}, \mathcal{E}_{T_5}, \mathcal{E}_{T_8}) = (\{\}, \{(2,1),\ (1,3),\ (4,5)\}, \{(6,8),\ (7,6)\})$

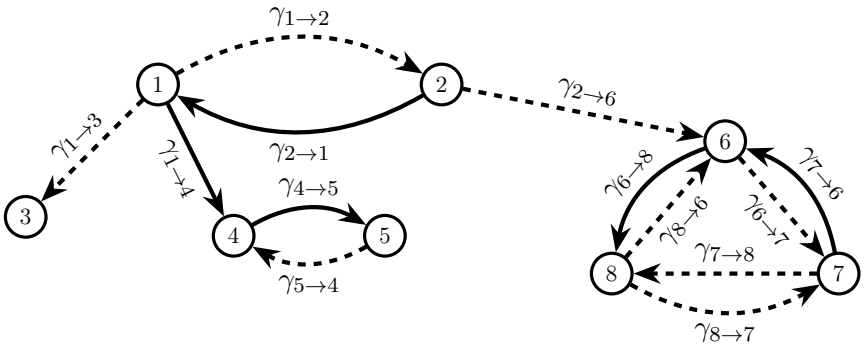


Figure A.7.: Example of three possible in-forests rooted in the set $J = \{3, 5, 8\}$, while fixing the partition $(\Omega_3, \Omega_5, \Omega_8) = (\{3\}, \{1, 2, 4, 5\}, \{6, 7, 8\})$. Other elements of Part$(\Omega, J)$ are $(\{1, 2, 3\}, \{4, 5\}, \{6, 7, 8\})$,$(\{3\}, \{4, 5\}, \{1, 2, 6, 7, 8\})$, $(\{1, 3\}, \{4, 5\}, \{2, 6, 7, 8\})$ and $(\{3\}, \{1, 4, 5\}, \{2, 6, 7, 8\})$, making a total of 15 in-forests, rooting in the set $J = \{3, 5, 8\}$.

For a given link $(a, b) \in \mathcal{E}$, we call the state $a \in \Omega$ a starting point. Then we can characterize in-trees as follows:

Lemma 37 (Characterization of in-trees and in-forests).

A network $(\Omega, \mathcal{E}_T)$ is an in-tree rooted at the state $i_0 \in \Omega$, if and only if the following three conditions are satisfied:

i) All states $j \in \Omega \backslash \{i_0\}$ are starting points of some link

ii) $|\mathcal{E}_T| = |\Omega| - 1$

iii) The network contains non cycles

Proof. Let $(\Omega, \mathcal{E}_T)$ be an in-tree. Since for every state $j \in \Omega$ there is a unique walk to the root ($j \rightsquigarrow i_0$), every state $j \in \Omega \backslash \{i_0\}$ must be the starting point of some link. This also implies $|\mathcal{E}_T| \geq |\Omega| - 1$. At the same time, no state can be the starting point of more than one link, since then the walk to the root would no longer be unique (if $a \to b \rightsquigarrow i_0$ and $a \to c \rightsquigarrow i_0$ for $b \neq c$, there are are at least two walk from state a to the root i_0). This is also the reason why there can be no cycles: With a cycle one is able to construct walks of arbitrary length (if $a \rightleftarrows b \to i_0$, we have $a \to b \to i_0$, $a \to b \to a \to b \to i_0$, etc.).
Vice versa, assume that a given network satisfies the above conditions. We can construct a walk from every state to the root as follows: Take an arbitrary state $j \in \Omega$. If $j = i_0$, we have the trivial walk (i_0) and are done. Otherwise there will be a link $j \to j_2$ to some other state $j_2 \in \Omega$. If $j_2 = i_0$, we are done, otherwise we repeat the process. Since there are only finitely many state and no cycles, this process must come to an end, meaning, the constructed walk must reach a state, which is no starting point. This state can only be the root i_0, as it is the only state, which is no starting point. Moreover, condition $i)$ together with condition $iii)$ implies that every state other than the root is the starting point of exactly one link, making the walk constructed above unique. This means, that the network is indeed an in-tree rooted at state i_0. □

Since every connected component of an in-forest in an in-tree, we also have a characterization of in-forests:

A network $(\Omega, \mathcal{E}_F)$ is an in-forest rooted at the states $J \subseteq \Omega$, if and only if the following three conditions are satisfied:

i) All states $j \in \Omega \backslash J$ are starting points of some link

ii) $|\mathcal{E}_F| = |\Omega| - |J|$

iii) The network contains non cycles.

Lemma 38 (In-trees and minimal absorbing sets).

For a given network $\mathcal{S} = (\Omega, \mathcal{E})$, the set of possible in-trees rooted at the state $a \in \Omega$ is non-empty, if and only if the state a lies in the only minimal absorbing set, that is

$$\mathcal{T}(\to a, \mathcal{S}) \neq \emptyset \iff a \in \bigcap_{B \in \mathcal{B}} B \tag{A.22}$$

Proof. '$\Longleftarrow$' We know from definition 6 that all states lead to some minimal absorbing set. If this minimal absorbing set $B \in \mathcal{B}$ is unique, then all states must lead to B and since B is strongly connected, then there must be a path from all other states to state $a \in B$, so there is at least one in-tree of $\mathcal{S}$ rooted in state a.

'$\Longrightarrow$' We prove this by contraposition: Suppose, there were a minimal absorbing set $B_0 \in \mathcal{B}$ with $a \notin B_0$. Since B_0 is absorbing, the state a cannot be reached by any state in B_0 ($B_0 \nrightsquigarrow a$), so the set of in-trees of $\mathcal{S}$ rooted at state a is empty.

□

Lemma 39 (In-forests and minimal absorbing sets).

The set of possible in-forests rooted at the states $J \subseteq \Omega$ is non-empty, if and only if the intersection of every minimal absorbing set with J is non-empty:

$$\mathcal{F}(\to J, \mathcal{S}) \neq \emptyset \iff \begin{array}{l}\text{for all } B \in \mathcal{B} \text{ the intersection of } B \\ \text{with } J \text{ is non-empty: } B \cap J \neq \emptyset\end{array} \tag{A.23}$$

Proof. '$\Longrightarrow$' We prove this by contraposition: Let $B_0 \in \mathcal{B}$ be a minimal absorbing set such that $B_0 \cap J = \emptyset$. Then no state in J can be reached by any state in B_0 (that is $j \nrightsquigarrow b_0$ for all $b_0 \in B_0$ and $j \in J$) and thus there can be no in-forest rooted at the state $J \subseteq \Omega$.

'$\Longleftarrow$' We prove this again by contraposition, so we assume that there is an in-forest rooted at the states $J \subseteq \Omega$ and we choose an arbitrary minimal absorbing set $B \in \mathcal{B}$. From the definition of an in-forest, we know that for all states $b \in B$ there exists a state $j_b \in J$ such that $b \rightsquigarrow j_b$. Since B is absorbing, this can only be true, if $j_b \in B$. Since $B \in \mathcal{B}$ was arbitrary, the claim follows.

□

Corollary 40 (Indicator term of multiple minimal absorbing sets).

For some fixed natural number $j \in \mathbb{N}$, the term $\sum_{\substack{J \in \mathrm{Pow}\,(\Omega) \\ |J|=j}} \sum_{F \in \mathcal{F}(\to J, \mathcal{S})} \gamma_F$ vanishes, if and only if there are strictly more than j many minimal absorbing sets, that is $|\mathcal{B}| > j$.

Proof. We have

$$0 = \sum_{\substack{J \in \mathrm{Pow}\,(\Omega) \\ |J|=j}} \sum_{F \in \mathcal{F}(\to J, \mathcal{S})} \gamma_F, \tag{A.24}$$

if and only if for all subsets $J \subseteq \Omega$ with $|J| = j$ elements, there are no in-forests rooted at the set J. By lemma 39, this is the case if and only if for all $J \subseteq \Omega$ with $|J| = j$ there exists a minimal absorbing set $B_J \in \mathcal{B}$ such that $B_J \cap J = \emptyset$. This can only be the case, when there are more than $|J| = j$ minimal absorbing set (since otherwise one could choose $J = \{b_1, \ldots, b_{|\mathcal{B}|}, b_{|\mathcal{B}|+1}, \ldots, b_{|J|}\}$) and vice versa, when $|\mathcal{B}| > |J|$ there can be no in-forests rooted at the states in J. □

A.7. Analytical expression for the principal minors of the generator matrix

In the following, we need the notion of minors of matrices, which are determinants of smaller submatrices.

Definition 41 (minors of matrices).

Let $M, N \in \mathbb{N}_{\geq 2}$ be natural numbers, $A \in \mathbb{C}^{M \times N}$ be a matrix and $I \subsetneqq \{1, \ldots, M\}$, $J \subsetneqq \{1, \ldots, N\}$ be subsets of the first $M\,(N)$ natural numbers. We denote by $A_{I,J} \in \mathbb{C}^{M-|I|, N-|J|}$ the submatrix of A that is obtained by deleting all rows $i \in I$ and all columns $j \in J$ and set $[A]_{I,J}$ to be the determinant of $A_{I,J}$ $\Big([A]_{I,J} := \det(A_{I,J})\Big)$, provided that $A_{I,J}$ is a square matrix. $[A]_{I,J}$ is called a principal minor if $I = J$ (compare the two example in Figure A.8).

(a)

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},$$

$$A_{\{2\},\{2\}} = \begin{pmatrix} a_{11} & \square & a_{13} \\ \square & \square & \square \\ a_{31} & \square & a_{33} \end{pmatrix}$$

$$[A]_{2,2} := [A]_{\{2\},\{2\}} = \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} = a_{11}\, a_{33} - a_{31}\, a_{13}$$

(b)

$$M = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} & m_{15} \\ m_{21} & m_{22} & m_{23} & m_{24} & m_{25} \\ m_{31} & m_{32} & m_{33} & m_{34} & m_{35} \\ m_{41} & m_{42} & m_{43} & m_{44} & m_{45} \end{pmatrix},$$

$$M_{\{1,3\},\{1,3,4\}} = \begin{pmatrix} \square & \square & \square & \square & \square \\ \square & m_{22} & \square & \square & m_{25} \\ \square & \square & \square & \square & \square \\ \square & m_{42} & \square & \square & m_{45} \end{pmatrix}$$

$$[M]_{\{1,3\},\{1,3,4\}} = \begin{vmatrix} m_{22} & m_{25} \\ m_{42} & m_{45} \end{vmatrix} = m_{22}\, m_{45} - m_{42}\, m_{45}$$

Figure A.8.: Example of the *minors* of two matrices: In example A.8a, we delete the row number 2 and the column number 2, whereas in example A.8b we delete the rows 1 and 3 and the column with the numbers 1, 3 and 4. $[A]_{\{2\},\{2\}}$ is a principal minor whereas $[M]_{\{1,3\},\{1,3,4\}}$ is not.

In the following we want to show that the principal minors of the (negative) generator matrix can be computed via sums of weights of in-forests.

Theorem 42 (Principal minors of the generator matrix).

Let Γ be the generator matrix of a network $\mathcal{S} = (\Omega, \mathcal{E})$, given by Equation (2.21), and $J \subseteq \{1, \ldots, |\Omega|\}$ be a subset of the first $|\Omega|$ numbers. Then the principal minors of the negative generator matrix are given by:

$$\begin{aligned} [-\Gamma]_{J,J} &= \sum_{F \in \mathcal{F}(\to J, \mathcal{S})} \gamma_F = \sum_{\{\Omega_k \,:\, k \in J\} \in \mathrm{Part}(\Omega, J \subseteq \Omega)} \prod_{j \in J} \sum_{T_j \in \mathcal{T}(\to j, \mathcal{S}_k)} \gamma_{T_j} = \\ &= \sum_{\{\Omega_k \,:\, k \in I\} \in \mathrm{Part}(\Omega, J \subseteq \Omega)} \sum_{(T_k)_{k \in J} \in \left(\bigoplus_{k \in J} \mathcal{T}(\to k, \mathcal{S}_k) \right)} \prod_{j \in J} \gamma_{T_j}, \end{aligned} \tag{A.25}$$

with the subnetwork $\mathcal{S}_k := (\Omega_k, \{(a, b) \in \mathcal{E} \,:\, a, b, \in \Omega_k\})$.
For $J = \{j\}$, this reduces to

$$[-\Gamma]_{j,j} = (-1)^{|\Omega|-1} [\Gamma]_{j,j} = \sum_{T_j \in \mathcal{T}(\to j, \mathcal{S})} \gamma_{T_j} \tag{A.26}$$

Proof. We follow the proof given in [MG13]. First, a few definitions:

- We denote by $\mathrm{Perm}(\Omega \backslash J)$ the set of all permutations of the set $\Omega \backslash J$
- For a permutation $\pi \in \mathrm{Perm}(\Omega \backslash J)$, we set $\mathrm{Fix}(\pi) := \{i \in \{1, \dots, |\Omega \backslash J|\} : \pi(i) = i\}$ to be the set of fixed points of π and $C_\pi := \{\text{cycles of } \pi\}$ the set of all cycles of the permutation π.
- The weight of all the cycles of the permutation π is defined as $\gamma_{C_\pi} := \prod_{i \in \Omega \backslash (J \cup \mathrm{Fix}(\pi))} \gamma_{i \to \pi(i)}$

We can write any permutation $\pi \in \mathrm{Perm}(\Omega \backslash J)$ as a product of cycles, namely $\pi = c_1 \circ \dots \circ c_{|C_\pi|}$, with the cycles $c_i := (\omega_1^{(i)}, \dots, \omega_{k_i}^{(i)}) := \begin{pmatrix} \omega_1^{(i)} & \dots & \omega_{k_i}^{(i)} \\ \omega_2^{(i)} & \dots & \omega_1^{(i)} \end{pmatrix}$ for $i \in \{1, \dots, |C_\pi|\}$.
The principal minors of the negative generator matrix are given by

$$[-\Gamma]_{J,J} = \sum_{\pi \in \mathrm{Perm}(\Omega \backslash J)} \mathrm{sgn}(\pi) \prod_{\alpha \in \Omega \backslash J} (-\Gamma)_{\pi(\alpha), \alpha}. \tag{A.27}$$

The sign of a permutation $\pi \in \mathrm{Perm}(\Omega \backslash J)$ can be computed to

$$\mathrm{sgn}(\pi) = \prod_{i=1}^{|C_\pi|} \underbrace{\mathrm{sgn}(c_i)}_{(-1)^{k_i - 1}} = (-1)^{\sum_{i=1}^{|C_\pi|} k_i - |C_\pi|} = (-1)^{K_\pi + |C_\pi|}, \tag{A.28}$$

with $K_\pi := \sum_{i=1}^{|C_\pi|} k_i = |\Omega \backslash (J \cup \mathrm{Fix}(\pi))|$.
For the product, we get:

$$\prod_{\alpha\in\Omega\setminus J}(-\Gamma)_{\pi(\alpha),\alpha} = \prod_{i\in\mathrm{Fix}(\pi)} \underbrace{(-\Gamma)_{i,i}}_{\sum_{j_i=1}^{|\Omega|}\gamma_{i\to j_i}} \cdot \prod_{\beta\in\Omega\setminus(J\cup\mathrm{Fix}(\pi))} \underbrace{(-\Gamma)_{\pi(\beta),\beta}}_{(-1)\,\gamma_{\beta\to\pi(\beta)}}$$

$$= \underbrace{\prod_{i\in\mathrm{Fix}(\pi)}\sum_{j_i=1}^{|\Omega|}\gamma_{i\to j_i}}_{\sum_{\boldsymbol{j}\in\{1,\dots,|\Omega|\}^{|\mathrm{Fix}(\pi)|}}\prod_{i\in\mathrm{Fix}(\pi)}\gamma_{i\to j_i}} \underbrace{\prod_{\beta\in\Omega\setminus(J\cup\mathrm{Fix}(\pi))}(-1)\,\gamma_{\beta\to\pi(\beta)}}_{(-1)^{K_\pi}\,\gamma_{C_\pi}} \tag{A.29}$$

$$= (-1)^{K_\pi}\,\gamma_{C_\pi} \sum_{\boldsymbol{j}\in\{1,\dots,|\Omega|\}^{|\mathrm{Fix}(\pi)|}} \prod_{i\in\mathrm{Fix}(\pi)}\gamma_{i\to j_i},$$

where the first sum s over the 'vector index' $\boldsymbol{j}\in\{1,\dots,|\Omega|\}^{|\mathrm{Fix}(\pi)|}$.
When we put these together, we end up with

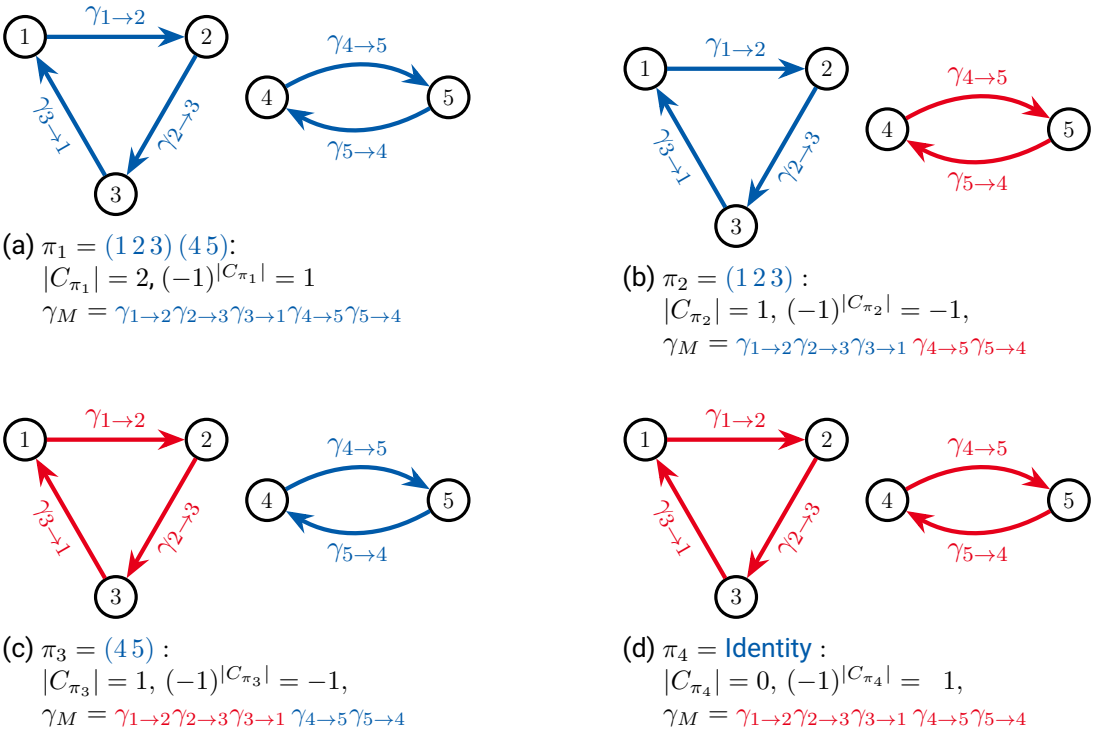
$$[-\Gamma]_{J,J} = \sum_{\pi\in\mathrm{Perm}(\Omega\setminus J)} \mathrm{sgn}(\pi) \prod_{\alpha\in\Omega\setminus J}(-\Gamma)_{\pi(\alpha),\alpha} =$$

$$= \sum_{\pi\in\mathrm{Perm}(\Omega\setminus J)} (-1)^{|C_\pi|} \sum_{\boldsymbol{j}\in\{1,\dots,|\Omega|\}^{|\mathrm{Fix}(\pi)|}} \underbrace{\gamma_{C_\pi}\prod_{i\in\mathrm{Fix}(\pi)}\gamma_{i\to j_i}}_{\text{monomial}} \tag{A.30}$$

For a given network $\mathcal{S}$, we call a product of weights of links a monomial of size $k\in\mathbb{N}$, if this product constitutes of k different links with k different starting points.
Every summand in Equation (A.30) is a product over $K_\pi + |\mathrm{Fix}(\pi)| = |\Omega| - |J|$ many different links, whose starting points are the set $\Omega\setminus J$.
Now we consider a fixed monomial γ_M and ask, with how many permutations $\pi\in\mathrm{Perm}(\Omega\setminus J)$, we can construct this monomial.
Since the term $\prod_{i\in\mathrm{Fix}(\pi)}\gamma_{i\to j_i}$ can yield additional cycles, we set $C_M := \{\text{ cycles in M}\}\supseteq C_\pi$ to be the set of cycles in M.
For every cycle in C_M, we can either construct it with permutations via the term γ_{C_M} or with the help of fixed points via the term $\prod_{i\in\mathrm{Fix}(\pi)}\gamma_{i\to j_i}$. So for every $k\in\{0,\dots,|C_M|\}$ there are $\binom{|C_M|}{k}$ many different ways to choose a permutation $\pi\in\mathrm{Perm}(\Omega\setminus J)$, such that $|C_\pi| = k$.
Although the monomial γ_M will be the same for all permutations, the sign can vary

due to the term $(-1)^{|C_\pi|}$. So for a fixed monomial γ_M with a positive number of cycles ($|C_M| > 0$), we have:

$$\gamma_M \sum_{k=0}^{|C_M|} \binom{|C_M|}{k} (+1)^{|C_M|-k} (-1)^k = \gamma_M (1-1)^{|C_M|} = 0. \tag{A.31}$$



(a) $\pi_1 = (1\,2\,3)\,(4\,5)$:
$|C_{\pi_1}| = 2, (-1)^{|C_{\pi_1}|} = 1$
$\gamma_M = \gamma_{1\to2}\gamma_{2\to3}\gamma_{3\to1}\gamma_{4\to5}\gamma_{5\to4}$

(b) $\pi_2 = (1\,2\,3)$:
$|C_{\pi_2}| = 1, (-1)^{|C_{\pi_2}|} = -1,$
$\gamma_M = \gamma_{1\to2}\gamma_{2\to3}\gamma_{3\to1}\,\gamma_{4\to5}\gamma_{5\to4}$

(c) $\pi_3 = (4\,5)$:
$|C_{\pi_3}| = 1, (-1)^{|C_{\pi_3}|} = -1,$
$\gamma_M = \gamma_{1\to2}\gamma_{2\to3}\gamma_{3\to1}\,\gamma_{4\to5}\gamma_{5\to4}$

(d) $\pi_4 =$ Identity :
$|C_{\pi_4}| = 0, (-1)^{|C_{\pi_4}|} = \ 1,$
$\gamma_M = \gamma_{1\to2}\gamma_{2\to3}\gamma_{3\to1}\,\gamma_{4\to5}\gamma_{5\to4}$

Figure A.9.: Example of the associated network of two cycles of the monomial

$$\gamma_{\mathcal{M}} \ = \ \gamma_{1\to2}\,\gamma_{2\to3}\,\gamma_{3\to1} \ \ \gamma_{4\to5}\,\gamma_{5\to4} \ = \sum_{\pi\in\mathsf{Perm}(\{1,\dots,5\})} \gamma_{C_\pi} \prod_{i\in\mathsf{Fix}(\pi)} \sum_{j_i=1}^{|\Omega|=5} \gamma_{i\to j_i}.$$

There are four different permutations that lead to these cycles with $|C_M| = 2$.

This means, that all monomials with a positive number of cycles cancel out and only the monomials with no cycles ($|C_M| = 0$) remain.
These monomials correspond to a subnetwork of $\mathcal{S}$, with no cycles and $||-|\mathbf{J}|$ many links, whose starting points are all states in $\setminus\mathbf{J}$.

By lemma 37, these networks are exactly the in-forests of the network $\mathcal{S}$, rooted at the set J. This finishes the proof.

□

Lemma 43 (First minors of the generator matrix). In order to evaluate $[\Gamma]_{i,j}$, we first have a look at a more general case:
Let $A \in \mathbb{C}^{N\times(N-1)}$ be an arbitrary matrix, whose columns sum up to zero, that is $(\underbrace{1,\ldots,1}_{N \text{ times}})\, A = (\underbrace{0,\ldots,0}_{N-1 \text{ times}})$. Then the first minors $[A]_{\{i\},\emptyset}$ of A, which are obtained by deleting row number i and taking the determinant of the resulting sub-matrix, satisfy the relation

$$[A]_{\{i\},\emptyset} = (-1)^{i-1}\, [A]_{\{1\},\emptyset}, \tag{A.32}$$

with the notation introduced in Definition 41.

Proof. The statement is trivial for $i = 1$, so let us assume that $i \geq 2$. We define the square matrix

$$B := \begin{pmatrix} \boldsymbol{e}_1^T A + \boldsymbol{e}_i^T A \\ \boldsymbol{e}_2^T A \\ \vdots \\ \boxed{\boldsymbol{e}_i^T A} \\ \vdots \\ \boldsymbol{e}_N^T A \end{pmatrix} = \begin{pmatrix} A_{1,1} + A_{i,1} & \ldots & A_{1,N-1} + A_{i,N-1} \\ A_{2,1} & \ldots & A_{2,N-1} \\ \vdots & & \vdots \\ \boxed{A_{i,1}} & \ldots & \boxed{A_{i,N-1}} \\ \vdots & & \vdots \\ A_{N,1} & \ldots & A_{N,N-1} \end{pmatrix} \in \mathbb{K}^{N\times(N-1)}, \tag{A.33}$$

where a box around an entry $\boxed{A_{ij}}$ means that this entry is missing. We notice that all its column sums vanishes, so B must be singular and its determinant must be equal to zero:

$$\begin{aligned} 0 = \det(B) &= \underbrace{\begin{vmatrix} A_{1,1} & \ldots & A_{1,N-1} \\ \vdots & & \vdots \\ \boxed{A_{i,1}} & \ldots & \boxed{A_{i,N-1}} \\ \vdots & & \vdots \\ A_{N,1} & \ldots & A_{N,N-1} \end{vmatrix}}_{[A]_{\{i\},\emptyset}} + \underbrace{\begin{vmatrix} A_{i,1} & \ldots & A_{i,N-1} \\ \vdots & & \vdots \\ \boxed{A_{i,1}} & \ldots & \boxed{A_{i,N-1}} \\ \vdots & & \vdots \\ A_{N,1} & \ldots & A_{N,N-1} \end{vmatrix}}_{(-1)^{i-2}\,[A]_{\{1\},\emptyset}} = \\ &\overset{(*)}{=\!=} [A]_{\{i\},\emptyset} + (-1)^{i-2}\, [A]_{\{1\},\emptyset} \end{aligned} \tag{A.34}$$

In the last step $(*)$, we have used the fact that the determinant changes its sign when interchanging two rows and we have move the first row $\boldsymbol{e}_i\, A = (A_{i,1}, \ldots, A_{i,N-1}) = (A\, \boldsymbol{e}_i)$ successively $i-2$-times to the bottom. □

In particular, the column sums of the matrix $\Gamma_{\emptyset,\{j\}}$ (that is, deleting the j-th column from the generator matrix) equals zero, so we can compute:

$$\begin{aligned}[\Gamma]_{ij} &\overset{\text{Def}}{=\!=} \underbrace{\left[\Gamma_{\emptyset,\{j\}}\right]_{\{i\},\emptyset}}_{(-1)^{i-1}\,[\Gamma_{\emptyset,\{j\}}]_{\{1\},\emptyset}} = (-1)^{i-1} \underbrace{\left[\Gamma_{\emptyset,\{j\}}\right]_{\{1\},\emptyset}}_{(-1)^{j-1}\,[\Gamma_{\emptyset,\{j\}}]_{\{j\},\emptyset}} \\ &= \underbrace{(-1)^{i-1}\,(-1)^{j-1}}_{(-1)^{i+j}} \underbrace{\left[\Gamma_{\emptyset,\{j\}}\right]_{\{j\},\emptyset}}_{[\Gamma]_{jj}} = (-1)^{i+j}\,[\Gamma]_{jj}\,.\end{aligned} \tag{A.35}$$

This means that we can compute all first minors $[\Gamma]_{ij}$ of the generator matrix Γ from its first principal minors $[\Gamma]_{jj}$. This can be used to compute the stationary state of a strongly connected network, as done in Section 2.5.1.

Lemma 44 (Connection between $a_{\lambda=0}$ and $|\mathcal{B}|$).

The algebraic multiplicity $a_{\lambda=0}$ for the eigenvalue $\lambda = 0$ of the generator matrix Γ equals the number of minimal absorbing sets.

Proof. Let the characteristic polynomial of the matrix Γ be given by

$$\det\ (t\,\mathbb{1} - \Gamma) = t^{|\Omega|} + \chi_1\, t^{|\Omega|-1} + \ldots + \chi_{|\Omega|-j}\, t^j + \ldots \chi_{|\Omega|-1}\, t^1 + \underbrace{\chi_{|\Omega|}}_{0}\cdot \tag{A.36}$$

When $a := a_{\lambda=0}$ is the algebraic multiplicity of the eigenvalue $\lambda = 0$, we have:

$$\begin{aligned}\chi_{|\Omega|} = 0 = \ldots = \chi_{|\Omega|-(a-1)} \text{ that is} \\ \chi_{|\Omega|-i} = 0 \text{ for all } i \in \{0\ldots, a-1\}\end{aligned} \tag{A.37}$$

We know from [HJ12] that the absolute value of $\chi_{|\Omega|-j}$ for $j \in \{1, \ldots, |\Omega|-1\}$ is given by the sum over all principal minors of Γ of order j, namely

$$|\chi_{|\Omega|-j}| \overset{\text{[HJ12]}}{=\!=} \sum_{\substack{J\in \text{Pow}\ (\Omega)\\ |J|=j}} [\Gamma]_{J,J} \overset{\text{Thm 42}}{=\!=} \sum_{\substack{J\in \text{Pow}\ (\Omega)\\ |J|=j}} \sum_{F\in\mathcal{F}(\to J,\mathcal{S})} \gamma_F. \tag{A.38}$$

By Corollary 40, this vanishes, if and only if there are strictly more than j minimal absorbing sets in $\mathcal{S}$

$$\left.\begin{array}{c} 0 = \chi_{|\Omega|-(a-1)} \\ 0 = \chi_{|\Omega|-(a-1)} \iff |\mathcal{B}| \geq a \end{array}\right\} \Longrightarrow |\mathcal{B}| \geq a \tag{A.39}$$

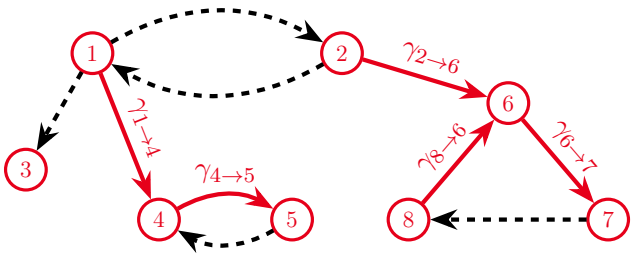
This means that there are at least a many minimal absorbing set, which results in

$$a_{\lambda=0} \overset{(A.39)}{\leq} |\mathcal{B}| \overset{14}{=\!=} g_{\lambda=0} \leq a_{\lambda=0}. \tag{A.40}$$

This also gives an elementary proof that the geometrical multiplicity $g_{\lambda=0}$ must coincide with the algebraic multiplicity $a_{\lambda=0}$ of the generator matrix Γ for the eigenvalue $\lambda = 0$. Since for any column stochastic matrix Q we can interpret the matrix $\Gamma_Q := Q - \mathbb{1}_{|\Omega|}$ as a generator matrix, this is also an elementary proof that the algebraic multiplicity $a_{\lambda=1}$ of the stochastic matrix Q agrees with its geometric multiplicity $g_{\lambda=1}$:

$$a_{\lambda=1}(Q) = a_{\lambda=0}(\underbrace{Q - \mathbb{1}_{|\Omega|}}_{\Gamma_\Omega}) \overset{A.40}{=\!=} g_{\lambda=0}(\underbrace{Q - \mathbb{1}_{|\Omega|}}_{\Gamma_\Omega}) = g_{\lambda=1}(Q). \tag{A.41}$$

(a) Three in-trees rooted at the states 3, 5 and 7



(b) Three in-trees rooted at the states 3, 4 and 8

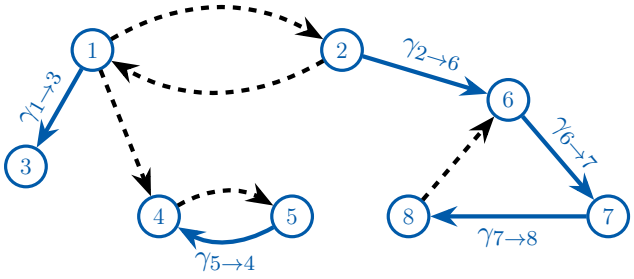


Figure A.10.: Example of two decomposition of the network Ω into three in-trees each. Since Ω has three minimal absorbing sets, it is not possible to decompose it into less than three in-trees.

$$
\begin{aligned}
T_3 &= (\Omega_3, \mathcal{E}_3) = \{\{3\}, \emptyset\} \\
T_5 &= (\Omega_5, \mathcal{E}_5) = \{\{1,4,5\}, \{(1,4),\ (4,5)\}\} \\
T_7 &= (\Omega_7, \mathcal{E}_7) \\
&= \{\{2,6,7,8\}, \{(2,6),\ (6,7),\ (8,6)\}\} \\
F_{\{3,5,7\}} &= \{\Omega, \mathcal{E}_{F_1}\} \\
\mathcal{E}_{F_1} &= \{(1,4),\ (4,5),\ (2,6),\ (6,7),\ (8,6)\}
\end{aligned}
\qquad
\begin{aligned}
T_3 &= (\Omega_3, \mathcal{E}_3) = \{\{1,3\}, \{(1,3)\}\} \\
T_4 &= (\Omega_4, \mathcal{E}_4) = \{\{4,5\}, \{(5,4)\}\} \\
T_8 &= (\Omega_8, \mathcal{E}_8) \\
&= \{\{2,6,7,8\}, \{(2,6),\ (6,7),\ (7,8)\}\} \\
F_{\{3,4,8\}} &= \{\Omega, \mathcal{E}_{F_2}\} \\
\mathcal{E}_{F_2} &= \{(1,3),\ (5,4),\ (2,6),\ (6,7),\ (7,8)\}
\end{aligned}
$$

□

A.8. Facts about trajectories in discrete-time Markov chains

Lemma 45 (Trajectories in discrete-time Markov chains and minimal absorbing sets).

In a classical, discrete-time Markov chain with a finite state space a trajectory $(X_n)_{n\in\mathbb{N}}$ will eventually be captured by a minimal absorbing set, that is $\sum_{B\in\mathcal{B}} \mathcal{P}(X_n \in B\{ \text{ eventually } \}) = 1$.

Proof. We know that for every state $s \in \Omega$ in the network there is a path to a minimal absorbing set $B \subseteq \Omega$ (see definition 6). This means that there are a finite number of states $s = s_0, \ldots, s_L \in B$ with $q_{s_j \to s_{j+1}} > 0$ [FD22]. We now let l_b be the length of the shortest (non-zero) path to the minimal absorbing set $b \in \mathcal{B}$ and q_B be the largest probability for reaching the set $B \in \mathcal{B}$ in l_B time steps, namely

$$
\begin{aligned}
l_B &:= \min_{l\in\mathbb{N}_0} \left\{(\omega_0 = s, \ldots, \omega_l \in B) \,:\, q_{\omega_j \to \omega_{j+1}} > 0 \text{ for all } j \in \{0, \ldots, l\}\right\} \\
L &:= \max_{B\in\mathcal{B}} \{l_B\}
\end{aligned}
$$

$$
\begin{aligned}
q_B &:= \max_{\boldsymbol{\omega}\in\Omega^{l_B+1}} \left\{ \prod_{j=0}^{l_B-1} q_{\omega_j \to \omega_{j+1}},\ \text{with } \omega_0 = s \text{ and } \omega_{l_B} \in B \right\} \\
q_0 &:= \min_{B\in\mathcal{B}} \{q_B\}.
\end{aligned}
$$

Further, be let $L\,(q_0)$ be the largest (smallest) of the $l_B\,(q_B)$. For a fixed natural number $m \in \mathbb{N}$ we define the event that a trajectory $\omega \in \Omega^{\mathbb{N}}$ after $L \cdot m$ time steps is not contained in a minimal absorbing set

$$A_m := \{\omega \in \Omega^{\mathbb{N}} : \omega_{L\cdot m} \notin \bigcup_{B\in\mathcal{B}} B\}. \tag{A.42}$$

The corresponding probability for A_m to occur can be estimated to $\mathcal{P}(A_m) \leq (1-q_0)^m$. Since the sum over all probabilities of A_m is finite ($\sum_{m\in\mathbb{N}} \mathcal{P}(A_m) < \infty$), we know by the lemma of Borel-Cantelli that the probability for A_m to occur infinitely often vanishes, $\mathcal{P}(A_m \text{ infinitely often }) = 0$.

□

A.9. From Markov chains to exponential waiting times

In this Section we prove the fact that a continuous-time Markov chain leads to exponential waiting time. To see this, we first compute the probability that the waiting time τ in some state $j \in \Omega$ is larger than $t+\epsilon$, provided we already know that the waiting time is larger than t:

$$\begin{aligned}
\mathcal{P}(\tau > t+\epsilon \,|\, \tau > t,\, X_0 = j) &= \frac{\mathcal{P}\left(\tau > t+\epsilon \,|\, X_0 = j\right)}{\mathcal{P}\left(\tau > t \,|\, X_0 = j\right)} \\
&= \frac{\mathcal{P}\left(X_{t+\epsilon} = j,\, X_\delta = j\,\forall\,\delta \in (0,t] \,|\, X_0 = j\right)}{\mathcal{P}\left(\, X_\delta = j\,\forall\,\delta \in (0,t] \,|\, X_0 = j\right)} \\
&= \mathcal{P}\left(X_{t+\epsilon} = j \,|\; X_\delta = j\,\forall\,\delta \in [0,t]\right) \\
&\overset{\text{Markov}}{=\!=\!=} \mathcal{P}\left(X_{t+\epsilon} = j \mid X_t = j\right) \\
&\overset{\text{time homogeneity}}{=\!=\!=\!=} \mathcal{P}\left(X_\epsilon = j \mid X_0 = j\right) \\
&= 1 - \gamma_{j\to}\,\epsilon + o(\epsilon) \text{ for } \epsilon \to 0^+,
\end{aligned} \tag{A.43}$$

where $\gamma_{j\to} = \sum_{k\in\Omega} \gamma_{j\to k}$ is the sum of all outgoing transition rates.
When taking the (right-sided) time derivative of the logarithm of the above equation, we

get:

$$\begin{aligned}
&\frac{\mathrm{d}}{dt}\ln\Big(\mathcal{P}\big(\tau > t \,|\, X_0 = j\big)\Big) \\
&\qquad = \lim_{\epsilon\to 0+}\frac{1}{\epsilon}\left[\ln\left(\underbrace{\mathcal{P}\left(\tau > t+\epsilon \,|\, X_0 = j\right)}_{\mathcal{P}(\tau>t\,|\,X_0=j)\,\cdot\,\mathcal{P}(\tau>t+\epsilon\,|\,\tau>t,\,X_0=j)}\right) - \ln(\mathcal{P}\left(\tau > t \,|\, X_0 = j\right))\right] \\
&\qquad = \lim_{\epsilon\to 0+}\frac{1}{\epsilon}\left[\ln\left(\underbrace{\mathcal{P}\left(\tau > t \,|\, X_0 = j\right)}_{1-\gamma_{j\to}\,\cdot\,\epsilon+o(\epsilon)}\right)\right] \\
&\qquad \overset{(*)}{=\!=} -\gamma_{j\to},
\end{aligned} \tag{A.44}$$

where we used in step $(*)$ the fact that $f(\epsilon) = 1 - \gamma_{j\to}\,\cdot\,\epsilon + o(\epsilon)$ for $\epsilon \to 0^+$ implies $\ln\big(f(\epsilon)\big) = -\gamma_{j\to}\,\cdot\,\epsilon + o(\epsilon)$ for $\epsilon \to 0^+$.
Equation (A.44) implies $\mathcal{P}\left(\tau > t \,|\, X_0 = j\right) = \mathrm{e}^{-t\,\gamma_{j\to}}$, when we keep in mind that the waiting time is almost surely positive.

A.10. The effect of including self-loops for discrete-time Markov chains

Suppose, we are given an irreducible, continuous-time Markov chain on a finite state space with generator matrix Γ and (unique) stationary solution $\boldsymbol{p}_*$. How can we construct a discrete-time Markov chain $Q(\Gamma)$, with the same stationary solution ?
This question has not been answered, with Section 2.4, since the associated, embedded discrete-time Markov chain Q_Γ has (usually) a different solution (compare Figure 2.9). But, choosing the transition probabilities proportional to the corresponding transition rates, while at the same time adding a self-loop at every state not only results in the same stationary solution, but guaranteeing at the same time that this stationary solution is attracting.

Theorem 46. For a given generator matrix Γ of a strongly connected network, a transition matrix $(Q_{ij})_{i,j\in\Omega} := (q_{j\to i})_{i,j\in\Omega}$ defined as

$$\begin{aligned}
q_{j\to i} &:= \quad \gamma_{j\to i}\,\Delta t \in [0,1) \\
q_{j\to j} &:= 1 - \gamma_{j\to}\,\Delta t \in (0,1]
\end{aligned}$$

with $\frac{1}{\Delta t} > \max\limits_{j\in\Omega}\gamma_{j\to}$.

has the same stationary solution $\boldsymbol{p}_*$ as the original network and this stationary solution is attracting.

Proof. The fact that the stationary solutions coincide, follows from the fact that $Q(\Gamma) = \mathbb{1} + \Delta t\, \Gamma$ and that the stationary solution for a strongly connected network is unique:

$$\dim \underbrace{\left(\operatorname{kern}\left(Q(\Gamma) - 1 \cdot \mathbb{1}\right)\right)}_{\Delta t\, \Gamma} = \dim \left(\Delta t\, \Gamma - 0 \cdot \mathbb{1}\right) \overset{\text{Thm 13}}{=\!=} 1. \tag{A.45}$$

What remain is to show that this stationary solution is indeed attracting, that is $\lim_{n\to\infty} Q(\Gamma)^n\, \boldsymbol{q}_0 = \boldsymbol{p}_*$.
The choice of Δt guarantees that every state has a self-loop ($q_{j\to j} > 0$), which ensures that the whole network is aperiodic and every discrete-time Markov chain on a finite state space, which is both irreducible and aperiodic (and therefore also positive recurrent) has a unique limiting distribution [Bré20; Dou+18; Pri13].
Another way to see this, is to apply Gershgorin's circle theorem, which guarantees that every eigenvalue is either equal to one or has a modulus strictly less than one: $|\lambda| \leq 1$ and $|\lambda| = 1 \iff \lambda = 1$. By the reasoning of subSection 2.1.2, this is sufficient for an attracting solution.

□

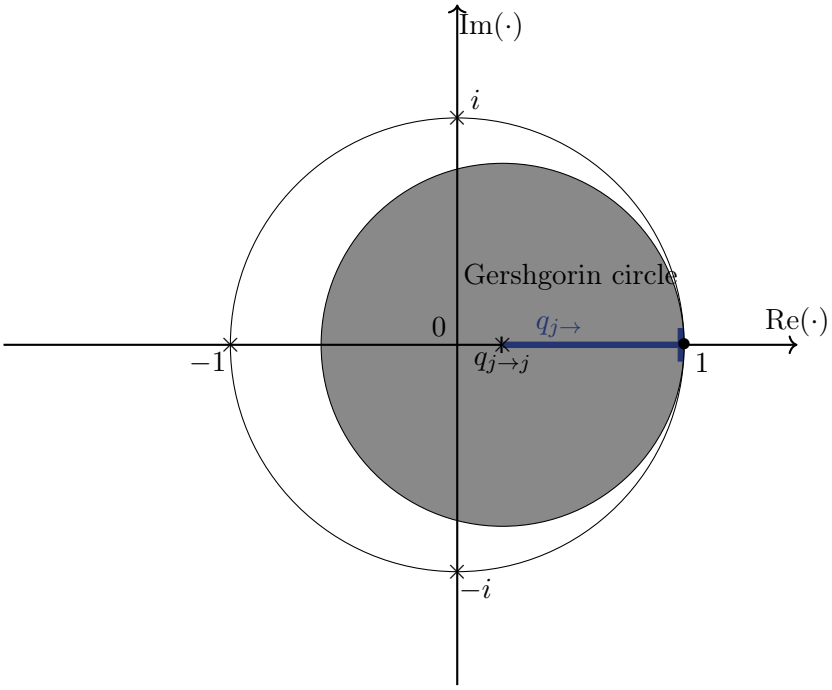


Figure A.11.: The union of all Gershgorin circles of the matrix $Q(\Gamma)$. Apart from the (non-degenerate) eigenvalue $\lambda = 1$, all eigenvalues have an absolute value strictly less than one, resulting in the fact that all (discrete) trajectories $\boldsymbol{q}_n(\boldsymbol{q}_0) = (Q(\Gamma))^n \ \boldsymbol{q}_0$ converge to the stationary distribution.

B. Appendix B: Lindblad equation

B.1. The time average and the ensemble average

From [KM04] we know that for almost all $\omega \in \mathcal{U}$ the time average

$$\Theta_\infty(\omega \,|\, \boldsymbol{\rho}_0) := \langle \Theta(t, \omega \,|\, \boldsymbol{\rho}_0) \rangle_{t \geq 0} := \lim_{T \to \infty} \frac{1}{T} \int_0^T dt \; \Theta(t, \omega \,|\, \boldsymbol{\rho}_0) \tag{B.1}$$

exists and it commutes with the ensemble average

$$\langle \langle \Theta(t, \omega \,|\, \boldsymbol{\rho}_0) \rangle_{t \geq 0} \rangle_{\omega \in \mathcal{U}} = \langle \langle \Theta(t, \omega \,|\, \boldsymbol{\rho}_0) \rangle_{\omega \in \mathcal{U}} \rangle_{t \geq 0}. \tag{B.2}$$

Here $\Theta_\infty(\omega \,|\, \boldsymbol{\rho}_0)$ is a random variable whose expectation value is the time average of $\boldsymbol{\rho}(t \,|\, \boldsymbol{\rho}_0)$.

B.2. Auxiliary calculations for Section 3.8.1

The form of the (conditional) time evolution operator

With respect to a suitable basis, we can write the conditional Hamiltonian in the Jordan normal form, that is

$$\begin{aligned} H_c &= \bigoplus_{m=1}^{M} J_m = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_M \end{pmatrix}, \text{ with} \\ J_m &= \begin{pmatrix} \delta_m & 1 & & \\ & \ddots & \ddots & \\ & & \delta_m & 1 \\ & & & \delta_m \end{pmatrix} \in \mathbb{C}^{s_m \times s_m}, \text{ with } \sum_{m=1}^{M} s_m = N. \end{aligned} \tag{B.3}$$

The eigenvalues of H_c are then given by $\{\delta_1, \ldots, \delta_M\}$

This makes it easier to compute matrix exponentials, in particular the conditional time evolution operator

$$\mathrm{e}^{-i\,H_c\,t} = \begin{pmatrix} \mathrm{e}^{-i\,J_1\,t} & & 0 \\ & \ddots & \\ 0 & & \mathrm{e}^{-i\,J_M\,t} \end{pmatrix} = \bigoplus_{m=1}^{M} \mathrm{e}^{-i\,J_m\,t} = \bigoplus_{m=1}^{M} \mathrm{e}^{-i\,\delta_m\,t} \begin{pmatrix} 1 & \frac{t^1}{1!} & \dots & & \frac{t^{(s_m-1)}}{(s_m-1)!} \\ & 1 & \frac{t^1}{1!} & & \frac{t^{(s_m-2)}}{(s_m-2)!} \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & \frac{t^1}{1!} \\ & & & & 1 \end{pmatrix} \tag{B.4}$$

When applying the conditional time evolution to a density matrix Θ_s, we get

$$\begin{aligned} U_\tau(\Theta_s) &= \begin{pmatrix} \mathrm{e}^{-i\,J_1\,\tau} & & 0 \\ & \ddots & \\ 0 & & \mathrm{e}^{-i\,J_M\,\tau} \end{pmatrix} \begin{pmatrix} \hat{\boldsymbol{\Theta}}^{(1,1)} & \dots & \hat{\boldsymbol{\Theta}}^{(1,M)} \\ \vdots & & \vdots \\ \hat{\boldsymbol{\Theta}}^{(M,1)} & \dots & \hat{\boldsymbol{\Theta}}^{(M,M)} \end{pmatrix} \begin{pmatrix} \mathrm{e}^{i\,J_1^T\,\tau} & & 0 \\ & \ddots & \\ 0 & & \mathrm{e}^{i\,J_M^T\,\tau} \end{pmatrix} = \\ &= \begin{pmatrix} \mathrm{e}^{-i\,J_1\,\tau}\,\hat{\boldsymbol{\Theta}}^{(1,1)}\,\mathrm{e}^{i\,J_1^T\,\tau} & \dots & \mathrm{e}^{-i\,J_1\,\tau}\,\hat{\boldsymbol{\Theta}}^{(1,M)}\,\mathrm{e}^{i\,J_M^T\,\tau} \\ \vdots & & \vdots \\ \mathrm{e}^{-i\,J_M\,\tau}\,\hat{\boldsymbol{\Theta}}^{(M,1)}\,\mathrm{e}^{i\,J_1^T\,\tau} & \dots & \mathrm{e}^{-i\,J_M\,\tau}\,\hat{\boldsymbol{\Theta}}^{(M,M)}\,\mathrm{e}^{i\,J_M^T\,\tau}. \end{pmatrix} \end{aligned} \tag{B.5}$$

When we look at the components $\left(\sum_{\mu=1}^{m-1} s_\mu + j, \sum_{\nu=1}^{n-1} s_\nu + k\right)$ of Equation (B.5), we get:

$$\begin{aligned} \left(U_\tau(\Theta_s)\right)_{\left(\sum_{\mu=1}^{m-1} s_\mu + j, \sum_{\nu=1}^{n-1} s_\nu + k\right)} &= \left(\mathrm{e}^{-i\,J_m\,\tau}\,\hat{\boldsymbol{\Theta}}^{(m,n)}\,\mathrm{e}^{i\,J_n^T\,\tau}\right)_{j\,k} = \\ &= \sum_{\alpha=1}^{s_m} \sum_{\beta=1}^{s_n} \underbrace{\left(\mathrm{e}^{-i\,J_m\,t}\right)_{j\,\alpha}}_{\mathrm{e}^{-i\,\delta_m\,t}\,\frac{t^{(\alpha-j)}}{(\alpha-j)!}\,1_{\{j\le\alpha\}}} \left(\hat{\boldsymbol{\Theta}}^{m,n}\right)_{\alpha\,\beta} \underbrace{\left(\mathrm{e}^{i\,J_m^T\,t}\right)_{\beta\,k}}_{\mathrm{e}^{i\,\delta_n\,t}\,\frac{t^{(\beta-k)}}{(\beta-k)!}\,1_{\{k\le\beta\}}} = \\ &= \sum_{\alpha=j}^{s_m} \sum_{\beta=k}^{s_n} \underbrace{\mathrm{e}^{i\,(\delta_n^*-\delta_m)\,t}}_{\mathrm{e}^{i\,(R_n-R_m)\,t}\,\mathrm{e}^{-(I_n+I_m)\,t}} \frac{t^{(\alpha-j)}}{(\alpha-j)!}\,\frac{t^{(\beta-k)}}{(\beta-k)!} \left(\hat{\boldsymbol{\Theta}}^{(m,n)}\right)_{\alpha,\beta}. \end{aligned} \tag{B.6}$$

Auxiliary Calculations for Equation (3.52)

When $\Theta_s \in \Omega_{\boldsymbol{\rho}_0}$ is a state in the quantum trajectory that is not a possible trapping state, then

$$\lim_{\tau\to\infty} \mathrm{Tr}[U_\tau(\Theta_s)] \cdot \int_0^\tau \frac{U_t(\Theta_s)}{\mathrm{Tr}[U_t(\Theta_s)]}\,\mathrm{d}\,t = 0. \tag{B.7}$$

Proof. It suffices to show that the left hand side of Equation (B.7) is a positive semi-definite matrix with vanishing trace. To see this, we take an arbitrary vector $\psi \in \mathbb{C}^N$ and compute

$$\langle \psi \,|\, U_t(\Theta_s)\,\psi\rangle = \left\langle \mathrm{e}^{i\,H_c^\dagger\,t}\,\psi \,|\, \Theta_s\,\mathrm{e}^{i\,H_c^\dagger\,t}\,\psi \right\rangle \geq 0, \tag{B.8}$$

where we used the fact that the density matrix Θ_s is positive semi-definite. Since multiplying by a positive scalar and integrating over the interval $[0,\tau]$ does not change positivity, we have shown the first part of our claim.
For the second part we use the Jordan normal form of H_c (see Equation (B.3))
We recall that for non-possible trapping states Θ_s the trace $U_t(\Theta_s)$ vanishes as t tends to infinity (see the algorithm for the unravelling in Section 3.3.1, 2 ii)) and compute

$$\begin{aligned} 0 \overset{\substack{\Theta_s \text{ is no possible}\\ \text{trapping state}}}{=\!=\!=} \lim_{t\to\infty} \mathrm{Tr}[U_t(\Theta_s)] &= \lim_{t\to\infty} \sum_{m=1}^{M}\sum_{j=1}^{s_m} \frac{(U_\tau(\Theta_s))_{\left(\sum_{\mu=1}^{m-1} s_\mu + j,\, \sum_{\mu=1}^{m-1} s_\mu + j\right)}}{\left(\mathrm{e}^{-i\,J_m\,\tau}\,\hat{\boldsymbol{\Theta}}^{(m,m)}\,\mathrm{e}^{i\,J_m^T\,\tau}\right)_{j\,j}} = \\ &= \lim_{t\to\infty} \sum_{m=1}^{M}\sum_{j=1}^{s_m}\sum_{\alpha,\beta=j}^{s_m} \mathrm{e}^{-2\,I_m\,t}\,\frac{t^{(\alpha-j)}}{(\alpha-j)!}\,\frac{t^{(\beta-j)}}{(\beta-j)!}\left(\hat{\boldsymbol{\Theta}}^{(m,m)}\right)_{\alpha,\beta} \end{aligned} \tag{B.9}$$

This means that whenever the imaginary part of an eigenvalue of the conditional Hamiltonian equals zero, then the corresponding block of Θ_s must vanish

$$\mathrm{Im}\,[\delta_m] = 0 \Longrightarrow \hat{\boldsymbol{\Theta}}^{(m,m)} = \mathbf{0}^{s_m\times s_m}. \tag{B.10}$$

When we now take the trace over the left side of Equation (B.7), we get

$$\lim_{\tau\to\infty} \mathrm{Tr}[U_\tau(\Theta_s)] \cdot \underbrace{\mathrm{Tr}\left[\int_0^\tau \frac{U_t(\Theta_s)}{\mathrm{Tr}[U_t(\Theta_s)]}\,\mathrm{d}\,t\right]}_{\tau} \overset{\text{(B.10)}}{=\!=} 0, \tag{B.11}$$

since multiplying Equation (B.9) by a factor of τ does not influence the limiting behavior. □

B.3. The need for a finite state space

Now that we have derived a formula for the stationary solution of the Lindblad equation, which is valid when certain assumptions are satisfied, we will look at these assumptions and check whether they can be dropped.
One crucial assumptions was the need for a finite state space $|\mathcal{S}| < \infty$. When dropping this assumption, the formula does no longer hold true in general, as the following counter-example shows:

$$V_1 = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\sqrt{2}} \end{pmatrix}, \quad V_2 = \begin{pmatrix} 0 & 1/2 \\ 0 & 1/2 \end{pmatrix}, \quad \Lambda = \sum_{k=1}^{2} \gamma_k V_k^\dagger V_k = \begin{pmatrix} \gamma_1 & 0 \\ 0 & \frac{\gamma_1+\gamma_2}{2} \end{pmatrix} \tag{B.12}$$

For $\gamma_1 = \gamma_2$, the set of possible states appearing in the Markov chain can be computed to

$$\Omega_{\boldsymbol{\rho}_0} = \left\{ \boldsymbol{\rho}_0, \Theta_{s_k} \ : \ k \in \mathbb{N}_0 \right\} \tag{B.13}$$

with

$$\begin{aligned} \Theta_{s_{2\,k}} &= \frac{1}{\frac{\boldsymbol{\rho}_{22}(0)}{2^k} + \boldsymbol{\rho}_{11}(0)} \begin{pmatrix} \boldsymbol{\rho}_{11}(0) & \frac{\boldsymbol{\rho}_{11}(0)}{2^{k/2}} \\ \frac{\boldsymbol{\rho}_{21}(0)}{2^{k/2}} & \frac{\boldsymbol{\rho}_{22}(0)}{2^k} \end{pmatrix}, \\ \Theta_{s_{2\,k+1}} &= \frac{1}{\frac{1}{2^k} + 1} \begin{pmatrix} 1 & \frac{1}{2^{k/2}} \\ \frac{1}{2^{k/2}} & \frac{1}{2^k} \end{pmatrix}. \end{aligned} \tag{B.14}$$

The probability for applying the Lindblad operators V_1 and V_2 is given by

$$\begin{aligned} \mathcal{P}(\pi_{k+1} = 1 \,|\, \Theta_k) &= 1 - \frac{\gamma_2}{\gamma_1 + \gamma_2} \cdot \frac{(\Theta_n)_{22}}{2^k\ (\Theta_n)_{11} + (\Theta_n)_{22}} \\ \mathcal{P}(\pi_{k+1} = 2 \,|\, \Theta_k) &= \frac{\gamma_2}{\gamma_1 + \gamma_2} \cdot \frac{(\Theta_n)_{22}}{2^k\ (\Theta_n)_{11} + (\Theta_n)_{22}}. \end{aligned} \tag{B.15}$$

Then, by the lemma of Borel-Cantelli, the probability for applying the operator V_2 infinitely often is zero (for $\boldsymbol{\rho}_{11}(0) \neq 0$), since $\sum_{k=0}^{\infty} \mathcal{P}(\pi_{k+1} = 2 \,|\, \Theta_k) < \infty$.
Hence we know almost surely that from a certain time step $k \in \mathbb{N}$ onward only the Lindblad operator V_1 will be applied (so the system will eventually follow the black arrow

of Figure B.1). This means that all states are transient states and hence no stationary distribution for this Markov chain exists.

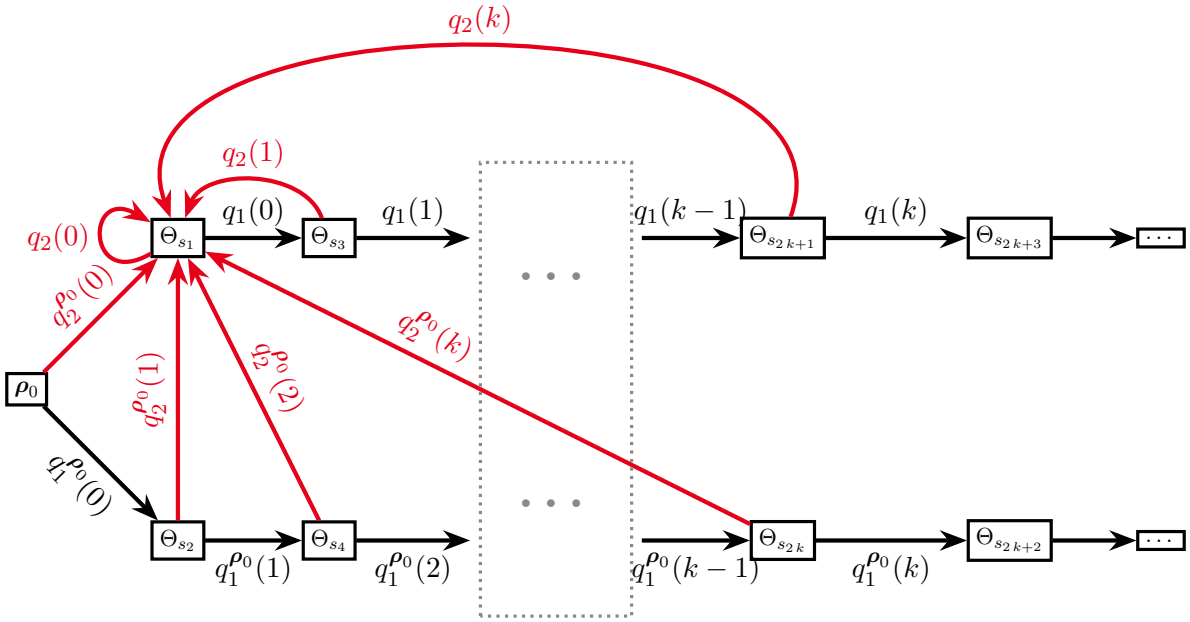


Figure B.1.: State transition network for the the Lindblad operators of example B.12. All states are transient states and the system will eventually follow only the black transitions, indicating the application of operator V_1.

Clearly, this can not happen, for finitely many states. The fact that there is no stationary solution of the corresponding Markov chain in the previous example, does not mean, that no stationary solution for the original Lindblad equation exist. And indeed, the stationary solution is given by $\boldsymbol{\rho}_\infty = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$

The reason for this discrepancy is the fact that the sequence of states in the discrete quantum trajectory $(\Theta_n)_{n\in\mathbb{N}}$ converges, while the classical states are unit vectors $(\boldsymbol{e}_n)_{n\in\mathbb{N}} \in l^2(\mathbb{N})$, which do not converge (with $\boldsymbol{e}_n = (0,\dots,0,1,0,\dots)$).

B.4. Example of a state transition network with an infinite state space, where the number of recurrent states is finite

Let the Hamiltonian and the Lindblad operators be given by

$$
\begin{aligned}
H &= \begin{pmatrix} E_1 & 0 & 0 & 0 \\ 0 & E_2 & 0 & 0 \\ 0 & 0 & E_3 & 0 \\ 0 & 0 & 0 & E_3 \end{pmatrix}, V_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, V_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \\
V_3 &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \Lambda = \begin{pmatrix} 2\,\gamma_1 & 0 & 0 & 0 \\ 0 & 2\,\gamma_1 & 0 & 0 \\ 0 & 0 & \gamma_2 & 0 \\ 0 & 0 & 0 & \gamma_3 \end{pmatrix}, \\
H_c &= \begin{pmatrix} E_1 - i\gamma_1 & 0 & 0 & 0 \\ 0 & E_2 - i\gamma_1 & 0 & 0 \\ 0 & 0 & E_3 - \frac{i\gamma_2}{2} & 0 \\ 0 & 0 & 0 & E_3 - \frac{i\gamma_3}{2} \end{pmatrix}.
\end{aligned}
\tag{B.16}
$$

The set of possible states appearing in the Markov consists of the five states

$$
\Omega_{\boldsymbol{\rho}_0} = \{\Theta_{s_1} = \boldsymbol{\rho}_0, \Theta_{s_2}, \Theta_{s_3}, \Theta_{s_4}, \Theta_{s_5}\}. \tag{B.17}
$$

with

$$
\Theta_{s_1} = \begin{pmatrix} \rho_{11}(0) & \rho_{12}(0) & \rho_{13}(0) & \rho_{14}(0) \\ \rho_{21}(0) & \rho_{22}(0) & \rho_{23}(0) & \rho_{24}(0) \\ \rho_{31}(0) & \rho_{32}(0) & \rho_{33}(0) & \rho_{34}(0) \\ \rho_{41}(0) & \rho_{42}(0) & \rho_{43}(0) & \rho_{44}(0) \end{pmatrix},
$$

$$
\Theta_{s_2}(\tau) = \begin{pmatrix} \frac{1}{2} \\ \hline \rho_{11}(0) + \rho_{22}(0) \end{pmatrix} \begin{pmatrix} \rho_{22}(0) & \rho_{21}(0)\,\mathrm{e}^{i\tau(E_1-E_2)} & \rho_{21}(0)\,\mathrm{e}^{i\tau(E_1-E_2)} & \rho_{22}(0) \\ \rho_{12}(0)\,\mathrm{e}^{-i\tau(E_1-E_2)} & \rho_{11}(0) & \rho_{11}(0) & \rho_{12}(0)\,\mathrm{e}^{-i\tau(E_1-E_2)} \\ \rho_{12}(0)\,\mathrm{e}^{-i\tau(E_1-E_2)} & \rho_{11}(0) & \rho_{11}(0) & \rho_{12}(0)\,\mathrm{e}^{-i\tau(E_1-E_2)} \\ \rho_{22}(0) & \rho_{21}(0)\,\mathrm{e}^{i\tau(E_1-E_2)} & \rho_{21}(0)\,\mathrm{e}^{i\tau(E_1-E_2)} & \rho_{22}(0) \end{pmatrix},
$$

$$
\Theta_{s_3} = \begin{pmatrix} \frac{1}{2} \\ \hline \rho_{11}(0) + \rho_{22}(0) \end{pmatrix} \begin{pmatrix} \rho_{11}(0) & \rho_{12}(0) & \rho_{12}(0) & \rho_{11}(0) \\ \rho_{21}(0) & \rho_{22}(0) & \rho_{22}(0) & \rho_{21}(0) \\ \rho_{21}(0) & \rho_{22}(0) & \rho_{22}(0) & \rho_{21}(0) \\ \rho_{11}(0) & \rho_{12}(0) & \rho_{12}(0) & \rho_{11}(0) \end{pmatrix},
$$

$$
\Theta_{s_4} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \text{ and } \Theta_{s_5} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{B.18}
$$

and the transition probabilities

$$
\begin{aligned}
q_{s_1\to s_1} &= 0, \\
q_{s_1\to s_2} &= \rho_{11}(0) + \rho_{22}(0), \\
q_{s_1\to s_3} &= 0, \\
q_{s_1\to s_4} &= \rho_{33}(0), \\
q_{s_1\to s_5} &= \rho_{44}(0),
\end{aligned}
\qquad
\begin{aligned}
q_{s_2\to s_1} &= 0, \\
q_{s_2\to s_2} &= 0, \\
q_{s_2\to s_3} &= \frac{1}{2}, \\
q_{s_2\to s_4} &= \frac{1}{2}\,\frac{\rho_{11}(0)}{\rho_{11}(0) + \rho_{22}(0)}, \\
q_{s_2\to s_5} &= \frac{1}{2}\,\frac{\rho_{22}(0)}{\rho_{11}(0) + \rho_{22}(0)},
\end{aligned}
\qquad
\begin{aligned}
q_{s_3\to s_1} &= 0, \\
q_{s_3\to s_2} &= \frac{1}{2}, \\
q_{s_3\to s_3} &= 0, \\
q_{s_3\to s_4} &= \frac{1}{2}\,\frac{\rho_{22}(0)}{\rho_{11}(0) + \rho_{22}(0)}, \\
q_{s_3\to s_5} &= \frac{1}{2}\,\frac{\rho_{11}(0)}{\rho_{11}(0) + \rho_{22}(0)},
\end{aligned}
$$

$$
\begin{aligned}
q_{s_4\to s_1} &= 0, \\
q_{s_4\to s_2} &= 0, \\
q_{s_4\to s_3} &= 0, \\
q_{s_4\to s_4} &= 0, \\
q_{s_4\to s_5} &= 1,
\end{aligned}
\qquad
\begin{aligned}
q_{s_5\to s_1} &= 0, \\
q_{s_5\to s_2} &= 0, \\
q_{s_5\to s_3} &= 0, \\
q_{s_5\to s_4} &= 1, \\
q_{s_5\to s_5} &= 0.
\end{aligned}
$$

This results in the following state transition network, which consists of an infinite number of states:

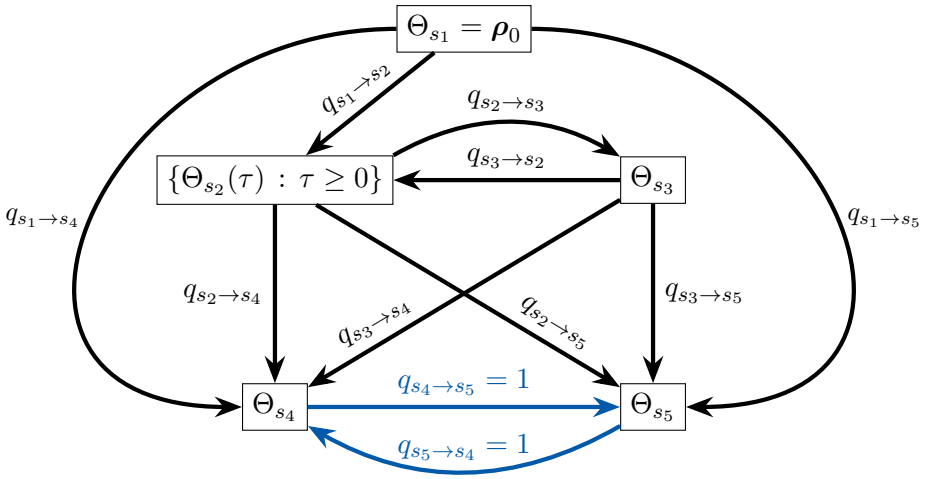


Figure B.2.: State transition network for a system of dimension $N = 4$, with an infinite number of states due to the set of states $\{\Theta_{s_2}(\tau) : \tau \geq 0\}$, but only finitely many states within a minimal absorbing set $B = \{\Theta_{s_4}, \Theta_{s_5}\}$, between which every quantum trajectory oscillates.

Since the number of states in each quantum trajectory remains finite, the stationary solution can still be computed, according to:

$$\boldsymbol{\rho}_\infty(\boldsymbol{\rho}_0) = \sum_{B\in\mathcal{B}} \mathcal{P}(B\,|\,\boldsymbol{\rho}_0) \sum_{s\in B} \frac{q_\infty(\Theta_s\,|\,B)\ \ \langle \boldsymbol{\tau}\cdot\langle \boldsymbol{\Theta}_s(t)\rangle_{t\in[0,\boldsymbol{\tau}]}\rangle_{\boldsymbol{\tau}\sim f(\cdot\,|\,\Theta_s)}}{\mathrm{Tr}[\dots]}$$

$$\overset{\mathcal{B}=\{B_1\}=\{\{\Theta_{s_4},\Theta_{s_5}\}\}}{=\!=\!=\!=\!=\!=\!=\!=\!=\!=\!=} \underbrace{\mathcal{P}(B_1\,|\,\boldsymbol{\rho}_0)}_{1} \frac{\sum_{i\in\{4,5\}} \overbrace{q_\infty(\Theta_{s_i}\,|\,B_1)}^{\frac{1}{2}}\ \ \overline{\boldsymbol{\tau}_{s_i}}\ \ \boldsymbol{\Theta}_{s_i}}{\mathrm{Tr}[\dots]}$$

$$= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{\gamma_2}{\gamma_2+\gamma_3} & 0 \\ 0 & 0 & 0 & \frac{\gamma_3}{\gamma_2+\gamma_3} \end{pmatrix}.$$

B.5. The limit of two fractions

Lemma 47. When the number of quantum jumps in a quantum trajectory is not bounded, then the time point $t_{J(T)}$ of the latest time event divided by the time T converges in probability to one.

Proof. Fix $\epsilon > 0$, then we have:

$$\mathcal{P}\left(\left|\frac{t_{J(T)}}{T} - 1\right| > \epsilon\right) = \mathcal{P}\left(T - t_{J(T)} > \epsilon T\right) = \mathrm{Tr}[U_{\epsilon T}\left(\Theta_{J(T)}\right)] \xrightarrow{T\to\infty} 0, \tag{B.19}$$

where we used the fact that $\Theta_{J(T)}$ can never be a possible trapping state, since the number of quantum jumps is by assumption not bounded, so the quantum trajectory will almost surely jump at some point. □

B.6. Auxiliary calculations for Section 3.1

Let the quantum state of the bath and the time evolution operator be given by

$$\boldsymbol{\rho}_B = \sum_{\beta=1}^{D_B} \lambda_\beta\,|\varphi_\beta\rangle\,\langle\varphi_\beta| \in \mathbb{C}^{D_B\times D_B}. \tag{B.20}$$

Then we get for the time evolution:

$$
\begin{aligned}
\boldsymbol{\rho}_S(t) &= \mathrm{Tr}_B\left[\mathcal{U}(t)\left(\boldsymbol{\rho}_S(0)\otimes\boldsymbol{\rho}_B\right)\mathcal{U}(t)^\dagger\right] \\
&= \sum_{\alpha=1}^{D_B}\Big(\mathbb{1}_{D_S}\otimes\langle\varphi_\alpha|\Big)\,\mathcal{U}(t)\qquad\underbrace{\left(\boldsymbol{\rho}_S(0)\otimes\overbrace{\sum_{\beta=1}^{D_B}\lambda_\beta\,|\varphi_\beta\rangle\,\langle\varphi_\beta|}^{\boldsymbol{\rho}_B}\right)}_{\sqrt{\lambda_\beta}\left(\mathbb{1}_{D_S}\otimes|\varphi_\beta\rangle\right)\boldsymbol{\rho}_S(0)\left(\mathbb{1}_{D_S}\otimes\langle\varphi_\beta|\right)\sqrt{\lambda_\beta}}\qquad\mathcal{U}^\dagger(t)\Big(\mathbb{1}_{D_S}\otimes|\varphi_\alpha\rangle\Big) \\
&= \sum_{\alpha,\beta=1}^{D_B}\underbrace{\sqrt{\lambda_\beta}\Big(\mathbb{1}_{D_S}\otimes\langle\varphi_\alpha|\Big)\,\mathcal{U}(t)\left(\mathbb{1}_{D_S}\otimes|\varphi_\beta\rangle\right)}_{W_{\alpha,\beta}(t)}\boldsymbol{\rho}_S(0)\underbrace{\left(\mathbb{1}_{D_S}\otimes\langle\varphi_\beta|\right)\mathcal{U}(t)^\dagger\left(\mathbb{1}_{D_S}\otimes|\varphi_\alpha\rangle\right)}_{\left[\left(\mathbb{1}_{D_S}\otimes\langle\varphi_\alpha|\right)\mathcal{U}(t)\left(\mathbb{1}_{D_S}\otimes|\varphi_\beta\rangle\right)\right]^\dagger\sqrt{\lambda_\beta}} \\
&= \sum_{\alpha,\beta=1}^{D_B} W_{\alpha,\beta}(t)\,\boldsymbol{\rho}_S(0)\left(W_{\alpha,\beta}(t)\right)^\dagger,
\end{aligned}
$$

with

$$
W_{\alpha,\beta}(t) = \sqrt{\lambda_\beta}\left(\mathbb{1}_{D_S}\otimes\langle\varphi_\alpha|\right)\mathcal{U}(t)\left(\mathbb{1}_{D_S}\otimes|\varphi_\beta\rangle\right)\in\mathbb{C}^{D_S\times D_S} \tag{B.21}
$$

For the Kraus-operators $W_{\alpha,\beta}(t)$ the following relation holds:

$$
\begin{aligned}
&\sum_{\alpha,\beta=1}^{D_B} W_{\alpha,\beta}(t)^\dagger\, W_{\alpha,\beta}(t) \\
&\qquad = \sum_{\alpha,\beta=1}^{D_B} \sqrt{\lambda_\beta}\left(\mathbb{1}_{D_S}\otimes\langle\varphi_\beta|\right)\mathcal{U}(t)^\dagger \underbrace{\left(\mathbb{1}_{D_S}\otimes|\varphi_\alpha\rangle\right)\left(\mathbb{1}_{D_S}\otimes\langle\varphi_\alpha|\right)}_{\mathbb{1}_{D_S}\otimes|\varphi_\alpha\rangle\langle\varphi_\alpha|}\mathcal{U}(t)\left(\mathbb{1}_{D_S}\otimes|\varphi_\beta\rangle\right)\sqrt{\lambda_\beta} \\
&\qquad = \sum_{\beta=1}^{D_B}\lambda_\beta\left(\mathbb{1}_{D_S}\otimes\langle\varphi_\beta|\right)\underbrace{\mathcal{U}(t)^\dagger\left(\mathbb{1}_{D_S}\otimes\overbrace{\sum_{\alpha=1}^{D_B}|\varphi_\alpha\rangle\langle\varphi_\alpha|}^{\mathbb{1}_{D_S}}\right)\mathcal{U}(t)}_{\mathbb{1}_{D_S}}\left(\mathbb{1}_{D_S}\otimes|\varphi_\beta\rangle\right) \\
&\qquad = \sum_{\beta=1}^{D_B}\lambda_\beta\,\mathbb{1}_{D_S}\underbrace{\langle\varphi_\beta\,|\,\varphi_\beta\rangle}_{1} \\
&\qquad = \underbrace{\left(\sum_{\beta=1}^{D_B}\lambda_\beta\right)}_{1}\mathbb{1}_{D_S}.
\end{aligned}
\tag{B.22}
$$

Let $\left(F_i\right)_{i\in\{1,\dots,D_S^2\}}$ be an ordered basis of $\mathbb{C}^{D_S\times D_S}$ with

$$
\begin{aligned}
F_{D_S^2} &= \frac{1}{\sqrt{D_S}}\,\mathbb{1}_{D_S} \text{ and} \\
W_{\alpha,\beta}(t) &= \sum_{i=1}^{D_S^2} F_i\,\mathrm{Tr}[F_i\,W_{\alpha,\beta}(t)].
\end{aligned}
$$

Then we can write $\boldsymbol{\rho}_S(t)$ as follows

$$
\begin{aligned}
\boldsymbol{\rho}_S(t) &= \sum_{\alpha,\beta=1}^{D_B} \underbrace{W_{\alpha,\beta}(t)}_{\sum_{i=1}^{D_S^2} F_i \operatorname{Tr}[F_i\, W_{\alpha,\beta}(t)]} \boldsymbol{\rho}_S(0) \underbrace{\left(W_{\alpha,\beta}(t)\right)^\dagger}_{\sum_{j=1}^{D_S^2} F_j^\dagger \overline{\operatorname{Tr}[F_j\, W_{\alpha,\beta}(t)]}} \\
&= \sum_{i,j=1}^{D_S} \underbrace{\left(\sum_{\alpha,\beta=1}^{D_B} \operatorname{Tr}[F_i\, W_{\alpha,\beta}(t)]\, \overline{\operatorname{Tr}[F_j\, W_{\alpha,\beta}(t)]} \right)}_{c_{i\,j}(t)} F_i\, \boldsymbol{\rho}_S(0)\, F_j^\dagger,
\end{aligned}
\tag{B.23}
$$

where the coefficient matrix $(c_{ij}(t))_{i,j\in\{1,\ldots,D_S^2\}}$ is self-adjoint and positive semi-definite:

$$
\begin{aligned}
\overline{c_{ij}(t)} &= c_{ji}(t) \text{ and} \\
\langle \boldsymbol{v}, c\, \boldsymbol{v} \rangle &= \sum_{i,j=1}^{D_S^2} \overline{v_i}\, c_{ij}(t)\, v_j \\
&= \sum_{\alpha,\beta=1}^{D_B} \underbrace{\left(\operatorname{Tr}\left[\sum_{i=1}^{D_S^2} \overline{v_i} F_i\, W_{\alpha,\beta}(t) \right] \overline{\operatorname{Tr}\left[\sum_{j=1}^{D_S^2} v_j F_j\, W_{\alpha,\beta}(t) \right]} \right)}_{\left| \operatorname{Tr}\left[\sum_{j=1}^{D_S^2} v_j F_j\, W_{\alpha,\beta}(t) \right] \right|^2} \\
&\geq 0, \\
&\text{for all } \boldsymbol{v} \in \mathbb{C}^{D_S^2}
\end{aligned}
\tag{B.24}
$$

In order to compute the generator of the semi-group, we need to compute the time derivative at $t = 0$.
with

$$
\begin{aligned}
a_{i\,j} &:= \frac{\mathrm{d}}{\mathrm{d}t} c_{ij}(t)\Big|_{t=0} = c'_{ij}(t=0) = \lim_{\epsilon \to 0^+} \frac{c_{ij}(\epsilon) - \delta_{i,D_S^2}\, \delta_{j,D_S^2}\, N}{\epsilon} \text{ and} \\
F &:= \frac{1}{\sqrt{N}} \sum_{i=1}^{D_S^2-1} a_{i\,D_S^2}\, F_i
\end{aligned}
\tag{B.25}
$$

we get:

$$
\begin{aligned}
\mathcal{L}(\boldsymbol{\rho}(t)) &= \frac{\mathrm{d}}{\mathrm{d}\,t}\boldsymbol{\rho}(t)\Big|_{t=0} = \lim_{\epsilon\to 0^+}\frac{\boldsymbol{\rho}(\epsilon)-\boldsymbol{\rho}(0)}{\epsilon}\lim_{\epsilon\to 0^+}\frac{\sum_{i,j=1}^{D_S^2} c_{ij}(\epsilon)\,F_i\,\boldsymbol{\rho}_0\,F_j^\dagger - \boldsymbol{\rho}_0}{\epsilon} \\
&= a_{D_S^2\,D_S^2}\,\boldsymbol{\rho}_0 + \sum_{i,j=1}^{D_S-1} a_{i\,j}\,F_i\,\boldsymbol{\rho}_0\,F_j^\dagger \\
&+ \underbrace{\left(\underbrace{\left(\frac{1}{\sqrt{D_S}}\sum_{i=1}^{D_S-1} a_{i\,D_S^2}\,F_i\right)}_{F}\boldsymbol{\rho}_0 + \boldsymbol{\rho}_0\,\underbrace{\left(\frac{1}{\sqrt{D_S}}\sum_{j=1}^{D_S-1} a_{D_S^2\,j}\,F_j^\dagger\right)}_{F^\dagger}\right)}_{\frac{(F-F^\dagger)\,\boldsymbol{\rho}_0-\boldsymbol{\rho}_0\,(F-F^\dagger)}{2}+\frac{(F+F^\dagger)\,\boldsymbol{\rho}_0+\boldsymbol{\rho}_0\,(F+F^\dagger)}{2}} \\
&= -i\left[\underbrace{\left(\frac{F-F^\dagger}{-2\,i}\right)}_{H},\,\boldsymbol{\rho}\right] + \left\{\underbrace{\left(\frac{a_{D_S^2\,D_S^2}}{2}+\frac{F+F^\dagger}{2}\right)}_{=:\,G},\,\boldsymbol{\rho}\right\} + \sum_{i,j=1}^{D_S-1} a_{i\,j}\,F_i\,\boldsymbol{\rho}_0\,F_j^\dagger \\
&= -i[H,\boldsymbol{\rho}(t)] + \sum_{i,j=1}^{D_S-1} a_{ij}\,F_i\,\boldsymbol{\rho}(t)\,F_j^\dagger + \Big\{\underbrace{\left(a_{D_S^2,D_S^2}\,\mathbb{1}_{D_S}+\frac{F^\dagger+F}{2}\right)}_{G},\boldsymbol{\rho}\Big\}.
\end{aligned}
\tag{B.26}
$$

For the time evolution to be trace preserving, the trace of $\mathcal{L}(\boldsymbol{\rho})$ must vanish:

$$
\begin{aligned}
0 \overset{!}{=} \operatorname{Tr}[\mathcal{L}(\boldsymbol{\rho}(t))] &= 2\operatorname{Tr}[G\,\boldsymbol{\rho}(t)] + \sum_{i,j=1}^{D_S-1} a_{ij} \underbrace{\operatorname{Tr}[F_i\,\boldsymbol{\rho}(t)\,F_j^\dagger]}_{\operatorname{Tr}[F_j^\dagger\,F_i\,\boldsymbol{\rho}(t)]} \\
&= \operatorname{Tr}\left[\underbrace{\left(2\,G + \sum_{i,j=1}^{D_S-1} a_{ij}\,F_j^\dagger\,F_i\right)}_{0}\boldsymbol{\rho}(t)\right] \\
\Longrightarrow G &= \frac{1}{2}\sum_{i,j=1}^{D_S-1} a_{ij}\,F_j^\dagger\,F_i. \\
\Longrightarrow \mathcal{L}(\boldsymbol{\rho}(t)) &= -i[H,\boldsymbol{\rho}(t)] + \sum_{i,j=1}^{D_S-1} \underbrace{(a_{ij})}_{\sum_{k=1}^{D_S^2-1} S_{ik}\,\gamma_k\,(S^\dagger)_{kj}} \left(F_i\,\boldsymbol{\rho}(t)\,F_j^\dagger + \left\{F_j^\dagger\,F_i, \boldsymbol{\rho}\right\}\right) \\
&= -i[H,\boldsymbol{\rho}(t)] \\
&\quad + \sum_{k=1}^{D_S^2-1}\gamma_k \left(\underbrace{\left(\sum_{i=1}^{D_S^2-1} F_i\,S_{ik}\right)}_{V_k}\boldsymbol{\rho}(t)\,\underbrace{\left(\sum_{j=1}^{D_S^2-1} F_j^\dagger\,S_{kj}^\dagger\right)}_{V_k^\dagger} + \left\{\left(\underbrace{\sum_{j=1}^{D_S^2-1} F_j^\dagger\,S_{kj}^\dagger}_{V_k^\dagger}\cdot\underbrace{\sum_{i=1}^{D_S^2-1} F_i\,S_{ik}}_{V_k}\right), \boldsymbol{\rho}(t)\right\}\right) \\
&= -i[H,\boldsymbol{\rho}(t)] + \sum_{k=1}^{D_S^2-1}\gamma_k\left(V_k\,\boldsymbol{\rho}(t)\,V_k^\dagger + \left\{V_k^\dagger\,V_k,\,\boldsymbol{\rho}\right\}\right).
\end{aligned}
\tag{B.27}
$$

In the last step, we used the fact that the coefficient matrix $(a_{ij})_{i,j\in\{1,\dots,D_S^2-1\}}$ is positive semi-definite, hence there exists a unitary coefficient matrix $(S_{ij})_{i,j\in\{1,\dots,D_S^2-1\}}$ and non-negative eigenvalues $(\gamma_k)_{k\in\{1,\dots,D_S^2-1\}}$ such that $a_{ij} = \sum_{k=1}^{D_S^2-1} S_{ik}\gamma_k\,(S^\dagger)_{kj}$.

C. Appendix C : Nomenclature

$\mathcal{S} = (\Omega, \mathcal{E}, \substack{\gamma \\ q})$	Directed, weighted graph with weight function $\substack{\gamma:\mathcal{E}\to\mathbb{R}_{\geq 0} \\ q:\mathcal{E}\to[0,1]}$
$\boldsymbol{p}_0 = \left(p_0^{(1)}, \ldots, p_0^{(\lvert\Omega\rvert)}\right)$ $\boldsymbol{q}_0 = \left(q_0^{(1)}, \ldots, q_0^{(\lvert\Omega\rvert)}\right)$	Initial state for a $\substack{\text{continuous-time} \\ \text{discrete-time}}$ Markov chain
$\boldsymbol{p}(t \mid \boldsymbol{p}_0) = \left(p^{(1)}(t \mid \boldsymbol{p}_0), \ldots, p^{(\lvert\Omega\rvert)}(t \mid \boldsymbol{p}_0)\right)$ $\boldsymbol{q}(n \mid \boldsymbol{q}_0) = \left(q^{(1)}(n \mid \boldsymbol{q}_0), \ldots, q^{(\lvert\Omega\rvert)}(n \mid \boldsymbol{q}_0)\right)$	probability vector of the $\substack{\text{continuous-time} \\ \text{discrete-time}}$ at time $\substack{t\geq 0 \\ n\in\mathbb{N}}$, given the initial state $\substack{\boldsymbol{p}_0 \\ \boldsymbol{q}_0}$.
$\boldsymbol{p}_\infty(\boldsymbol{p}_0) = \left(p_\infty^{(1)}(\boldsymbol{p}_0), \ldots, p_\infty^{(\lvert\Omega\rvert)}(\boldsymbol{p}_0)\right)$ $\boldsymbol{q}_\infty(\boldsymbol{q}_0) = \left(q_\infty^{(1)}(\boldsymbol{q}_0), \ldots, p_\infty^{(\lvert\Omega\rvert)}(\boldsymbol{p}_0)\right)$	The stationary solution of the $\substack{\text{continuous-time} \\ \text{discrete-time}}$ Markov chain corresponding to the initial state $\substack{\boldsymbol{p}_0 \\ \boldsymbol{q}_0}$.
Q	transition matrix for a discrete-time Markov chain
$Q(t)$	transition matrix for a continuous-time Markov chain
Γ	generator of the transition matrix for a continuous-time Markov chain

$\gamma_{j\to i}$	probability rate of a continuous-time Markov chain between the states $j \in \Omega$ and the state $i \in \Omega$
$\gamma_{i\to} := \sum_{\substack{j\in\Omega \\ j\neq i}} \gamma_{i\to j}$ outgoing rates	The sum of the from state $i \in \Omega$
$q_{i\to} := \sum_{\substack{j\in\Omega \\ j\neq i}} q_{i\to j}$	The sum of the outgoing probabilities from state $i \in \Omega$
$\langle\rangle_{t\geq 0} := \lim_{T\to\infty} \frac{1}{T} \int_0^T \mathrm{d}t$	time average
$\mathcal{U}$	set of all possible unravellings
$\langle\rangle_{\omega\in\mathcal{U}} := \int_{\mathcal{U}} \mathrm{d}\,\mathcal{P}(\omega)$	ensemble average
gcd	greatest common divisor
a_λ and g_λ	algebraic and geometric multiplicity of the eigenvalue λ.
Subnetworks	see: definition 4
$\mathrm{Perm}(J)$	the set of all permutation of the set J
$\mathrm{Fix}(\pi)$	the set of all fixed points of the permutation π

Pow (Ω)	The power set of the set Ω
$\boldsymbol{h}_{\lambda,s} := \boldsymbol{h}_{\lambda,s}(A) \in$ kern $(A - \lambda\mathbb{1})^s \backslash$kern $(A - \lambda\mathbb{1})^{s-1}$	The generalized eigenvector to the eigenvalue λ of the matrix A of step $s \in \mathbb{N}$
in-trees	see definition 34
in-forests	see definition 35
$\mathcal{T}(\to j, \mathcal{S})$ and $\mathcal{F}(\to J, \mathcal{S})$	see definition 36
$\frac{A}{\mathrm{Tr}[]} := \frac{A}{\mathrm{Tr}[A]}$	for every matrix A
$(\Theta_n)_{n\in\mathbb{N}}$	discrete quantum trajectory
H	Hamiltonian operator
$\{V_k \,:\, k \in I\}$	set of Lindblad operators with $\lvert I\rvert < \infty$
$\Lambda := \sum_{k\in I} \gamma_k \, V_k^\dagger \, V_k$	
$H_c := H - \frac{i}{2}\,\Lambda$	conditional Hamiltonian not self-adjoint, $H_c^\dagger \neq H_c$ determines the time evolution between two quantum jumps
$\Theta_n := \Theta(\omega, t_n)$	n-th state in the discrete quantum trajectory
$U_t(\Theta) := \mathrm{e}^{-i\,tH_c}\,\Theta\,\left(\mathrm{e}^{-i\,tH_c}\right)^\dagger$	time evolution

	without normalization
$\Theta(t) := \frac{U_t(\Theta)}{\mathrm{Tr}[\dots]}$	time evolution according to the conditional Hamiltonian H_c.
$\dot{\bigcup_{i\in I}} A_i$	disjoint union of sets, that is $\bigcup_{i\in I}$ with $A_i \cap A_j = \emptyset$, for $i \neq j$.

Bibliography

[Aga73] GS Agarwal. "I Master equation methods in quantum optics". In: Progress in optics. Vol. 11. Elsevier, 1973, pp. 1–76.

[And+22] Alexander A Andrianov et al. "The Franke–Gorini–Kossakowski–Lindblad–Sudarshan (FGKLS) Equation for Two-Dimensional Systems". In: Symmetry 14.4 (2022), p. 754.

[BP02] Heinz-Peter Breuer and Francesco Petruccione. "The theory of open quantum systems". Oxford University Press on Demand, 2002.

[BP12] Berislav Bua and Toma Prosen. "A note on symmetry reductions of the Lindblad equation: transport in constrained open spin chains". In: New Journal of Physics 14.7 (2012), p. 073007.

[Bre14] Paul C Bressloff. "Stochastic processes in cell biology". Vol. 41. Springer, 2014.

[Bré20] Pierre Brémaud. "Markov chains: Gibbs fields, Monte Carlo simulation, and queues". Vol. 31. Springer Science & Business Media, 2020.

[CK78] Seth Chaiken and Daniel J Kleitman. "Matrix tree theorems". In: Journal of combinatorial theory, Series A 24.3 (1978), pp. 377–381.

[Dav84] Mark H. A. Davis. "Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models". In: Journal of the Royal Statistical Society: Series B (Methodological) 46.3 (1984), pp. 353–388.

[Dou+18] Randal Douc et al. "Markov chains". Springer, 2018.

[EN00] Klaus-Jochen Engel and Rainer Nagel et al. "One-parameter semigroups for linear evolution equations". Vol. 194. Springer, 2000.

[FD22] Bernd Michael Fernengel and Barbara Drossel. "Obtaining the long-term behavior of master equations with finite state space from the structure of the associated state transition network". In: Journal of Physics A: Mathematical and Theoretical 55.11 (2022), p. 115201.

[Fro+12] Georg Frobenius et al. "Über Matrizen aus nicht negativen Elementen". In: (1912).

[Ger31] Semyon Aranovich Gershgorin. “Über die Abgrenzung der Eigenwerte einer Matrix”. In: Bulletin de lAcadémie des Sciences de lURSS 6 (1931), pp. 749–754.

[GKS76] Vittorio Gorini, Andrzej Kossakowski, and Ennackal Chandy George Sudarshan. “Completely positive dynamical semigroups of N-level systems”. In: Journal of Mathematical Physics 17.5 (1976), pp. 821–825.

[GL05] Yakov Goltser and Elena Litsyn. “Volterra integro-differential equations and infinite systems of ordinary differential equations”. In: Mathematical and computer Modelling 42.1-2 (2005), pp. 221–233.

[GM78] Harold N Gabow and Eugene W Myers. “Finding all spanning trees of directed and undirected graphs”. In: SIAM Journal on Computing 7.3 (1978), pp. 280–287.

[GP92] Nicolas Gisin and Ian C Percival. “The quantum-state diffusion model applied to open systems”. In: Journal of Physics A: Mathematical and General 25.21 (1992), p. 5677.

[GR74] NS Goel and N Richter-Dyn. “Stochastic Models in Biology”. 1974.

[Haa78] G Haag. “Transition factor method for discrete master equations; and application to chemical reactions”. In: Zeitschrift für Physik B Condensed Matter 29.2 (1978), pp. 153–159.

[HD83] Gunter Haag and Dimitrios S Dendrinos. “Toward a stochastic dynamical theory of location: A nonlinear migration process”. In: Geographical Analysis 15.4 (1983), pp. 269–286.

[Hil66] Terrell L Hill. “Studies in irreversible thermodynamics IV. Diagrammatic representation of steady state fluxes for unimolecular systems”. In: Journal of theoretical biology 10.3 (1966), pp. 442–459.

[HJ12] Roger A Horn and Charles R Johnson. “Matrix analysis”. Cambridge university press, 2012.

[Hoh98] Walter Hoh. “Pseudo differential operators generating Markov processes”. PhD thesis. Universität Bielefeld, 1998.

[Hon12] Josef Honerkamp. “Statistical physics: an advanced approach with applications”. Springer Science & Business Media, 2012.

[HW06] Bertram Huppert and Wolfgang Willems. “Lineare Algebra”. Vol. 2. Springer, 2006.

[Kel11] Frank P Kelly. “Reversibility and stochastic networks”. Cambridge University Press, 2011.

[KM04] Burkhard Kümmerer and Hans Maassen. “A pathwise ergodic theorem for quantum trajectories”. In: Journal of Physics A: Mathematical and General 37.49 (2004), p. 11889.

[Lid19] Daniel A Lidar. “Lecture notes on the theory of open quantum systems”. In: arXiv preprint arXiv:1902.00967 (2019).

[Lin76] Goran Lindblad. “On the generators of quantum dynamical semigroups”. In: Communications in Mathematical Physics 48.2 (1976), pp. 119–130.

[LR86] F Leighton and Ronald Rivest. “Estimating a probability using finite memory”. In: IEEE Transactions on Information Theory 32.6 (1986), pp. 733–742.

[MG13] Inomzhon Mirzaev and Jeremy Gunawardena. “Laplacian dynamics on general graphs”. In: Bulletin of mathematical biology 75.11 (2013), pp. 2118–2149.

[Per07] Oskar Perron. “Zur Theorie der Matrices”. In: Mathematische Annalen 64.2 (1907), pp. 248–263.

[Pri13] Nicolas Privault. “Understanding markov chains”. In: Examples and Applications, Publisher Springer-Verlag Singapore 357 (2013), p. 358.

[QG21] Hong Qian and Hao Ge. “Stochastic Chemical Reaction Systems in Biology”. Springer, 2021.

[Spo77] Herbert Spohn. “An algebraic condition for the approach to equilibrium of an open N-level system”. In: Letters in Mathematical Physics 2.1 (1977), pp. 33–38.

[TM21] Juzar Thingna and Daniel Manzano. “Degenerated Liouvillians and steady-state reduced density matrices”. In: Chaos: An Interdisciplinary Journal of Nonlinear Science 31.7 (2021), p. 073114.

[Tru18] Anton Sergeevich Trushechkin. “Finding stationary solutions of the Lindblad equation by analyzing the entropy production functional”. In: Proceedings of the Steklov Institute of Mathematics 301.1 (2018), pp. 262–271.

[Tut48] WT Tutte. “The dissection of equilateral triangles into equilateral triangles”. In: Mathematical Proceedings of the Cambridge Philosophical Society. Vol. 44. 4. Cambridge University Press. 1948, pp. 463–482.

[Van92] Nicolaas Godfried Van Kampen. “Stochastic processes in physics and chemistry”. Vol. 1. Elsevier, 1992.

[Wil22] Lauren K Williams. “The combinatorics of hopping particles and positivity in Markov chains”. In: arXiv preprint arXiv:2202.00214 (2022).

[Wüs13] Gisbert Wüstholz. “Algebra: Für Studierende der Mathematik, Physik, Informatik”. Springer-Verlag, 2013.

[Xun21] Jing Xun. “The Research of Markov Chain Application under Two Common Real World Examples”. In: Journal of Physics: Conference Series. Vol. 1936. 1. IOP Publishing. 2021, p. 012004.

Erklärungen laut Promotionsordnung

§ 8 Abs. 1 lit. c PromO

Ich versichere hiermit, dass die elektronische Version meiner Dissertation mit der schriftlichen Version übereinstimmt.

§ 8 Abs. 1 lit. d PromO

Ich versichere hiermit, dass zu einem vorherigen Zeitpunkt noch keine Promotion versucht wurde. In diesem Fall sind nähere Angaben über Zeitpunkt, Hochschule, Dissertationsthema und Ergebnis dieses Versuchs mitzuteilen.

§ 9 Abs. 1 PromO

Ich versichere hiermit, dass die vorliegende Dissertation selbstständig und nur unter Verwendung der angegebenen Quellen verfasst wurde.

§ 9 Abs. 2 PromO

Die Arbeit hat bisher noch nicht zu Prüfungszwecken gedient.

Darmstadt, 11.04.2023

Bernd Michael Fernengel