LIMIT THEOREMS FOR LAYERED MARKOV PROCESSES AND APPLICATIONS IN RNA TRANSCRIPTION MODELS

IN HOMOGENEOUS AND RANDOM ENVIRONMENTS

by

Mark VAN DER POST

to obtain the degree of Bachelor of Science
at the Delft University of Technology
to be defended publicly on Friday 5 July, 2024 at 10:00

Thesis committee

prof. dr. F.H.J. Redig	TU Delft	supervisor Applied Mathematics
dr. T. Idema	TU Delft	supervisor Applied Physics
dr. S.W.H. Eijt	TU Delft	faculty Applied Sciences
dr. J.L.A. Dubbeldam	TU Delft	faculty Applied Mathematics

Abstract

RNA transcription is an integral stage in the process cells undertake to synthesise proteins based on the DNA. For a long time, transcription was seen as a rather trivial procedure, but in recent years it has gained more recognition as a complex physical process with many external influences and interacting particles that are called RNA polymerase (RNAp). As a result, transcription has attracted a vast amount of scientific research and many stochastic models based on the mathematical theory of Markov processes have been proposed to emulate the motion of RNAps. A common assumption in most of these models is the use of uniform DNA. However, this assumption is not entirely justified as the DNA is in reality highly variable due to its changing nucleotide sequence and numerous proteins that bind to the DNA in specific positions. As such, the question arises whether results obtained in homogeneous models are still valid in inhomogeneous systems as well.

We will investigate a particular RNAp model proposed by (Klumpp, 2011). We establish the validity of the model by performing several numerical simulations to show that it exhibits various features typical to the motion of RNAps, such as clustering and a particular phenomenon known as "RNAp cooperation", thus reproducing results obtained by (Klumpp, 2011). Additionally, we will place the model in a more general, mathematical setting called the class of layered Markov / jump processes. Aside from transcription, this class has a wide range of physical applications, such as run-and-tumble particles and electron spin flip dynamics. We perform a thorough mathematical analysis of this class and present several theorems on the limiting behaviour of particles in these systems, such as the drift of the particle and the variation around this drift. These limit theorems can consequently be applied and interpreted in the relevant physical systems.

Additionally, we rigorously prove that the limit theorems carry over from homogeneous to inhomogeneous environments, which provides evidence that the original assumption of uniform DNA is justified in RNA transcription models. We complement this mathematical evidence by conducting numerical simulations in inhomogeneous environments and confirm that this leads to similar results as for the homogeneous model.

Contents

INTRODUCTION

The modern theory of DNA has proven instrumental in the understanding of evolution theory, heredity and the biological processes that constitute life. The basic principles are simple: DNA encodes the information needed for cells to produce proteins and these proteins play a fundamental role in all processes in the body. For examples, proteins are responsible for the contraction of muscles, the transportation of nutrients across the body and the catalyzation of chemical reactions such as metabolism.
Upon cell division, the DNA is replicated, such that offspring inherits genetic information from parents. For some time, many scientists believed that the genetic information alone fully encapsulates all characteristics and traits of the individual (van den Berg, 2017). Perhaps unsurprisingly, this view is not the complete picture, as some experimental observations cannot be explained solely from genetics. For instance, (Nakata et al., 2021) points out that children born to famished mothers are more susceptible to disease later in life.

Such phenomena can only be explained by looking further than the static DNA code. Even though DNA acts as a blueprint for protein synthesis, the actual synthesis is highly dependent on dynamic, external factors. This synthesis consists of two main stages: transcription and translation. During transcription, the DNA is copied onto an RNA strand, which serves as the protein template. In translation, this template is moved to a 'cellular factory' called ribosome, which assembles proteins. In this thesis, we will focus exclusively on transcription. Cells use various techniques to regulate this stage. For example, the cell can alter the structural organization of the DNA, which has a major impact on the rate of transcription. Any influence on protein synthesis that does not stem directly from the DNA sequence is called epigenetic. Epigenetic information is said to be as crucial as genetic information (Nakata et al., 2021).

It should therefore come as no surprise that the epigenetic influences and transcription have become "one of the most intensely studied areas of all of science" (Sims et al., 2004). Transcription consists of three phases: initiation, elongation and termination. Although elongation was long considered trivial, it is now seen as a complicated process with many external influences and dozens of interacting particles, called RNA polymerase (RNAp) (Sims et al., 2004). Several models have been proposed to emulate the motion of RNAps (Jülicher and Bruinsma, 1998; Tripathi and Chowdhury, 2008). Many of these models are stochastic in nature and model the RNAp as jumping randomly between DNA sites. The models are, as is common in physical stochastic processes, fully based on Markov theory.

In this thesis, we will investigate a particular RNAp model, presented by (Klumpp, 2011; Klumpp and Hwa, 2008). In essence, this model is as a one-dimensional process with multiple internal states or "layers". Each layer is associated with a different type of motion. In transcription, the RNAp may enter a paused state, a backtracking state or an active transcribing state. Although the model is relatively simple, it exhibits some features typical to RNAp, such as clustering and cooperation.

The RNAp model is part of a much more general class of Markov processes that we name 'layered jump processes'. In this class, we consider all Markov processes with various different internal layers, each having a specific type of diffusive or active motion. The main benefit of generalising the model's framework to such a degree, is that the class is highly applicable to other physical systems as well. For instance, layered jump processes are also used to model other molecular motors in nanobiology, such as cytoskeletal walkers (Akhmanova and Kapitein, 2022) and Run and Tumble particles (RTPs) (Bechinger et al., 2016). Additionally, layered jump processes have found applications in a wide range of fields outside of nano-biology. For instance, they can serve as a model for heat conduction (Krámli et al., 1986), a model for spin flip dynamics of an electron (Demaerel and Maes, 2018) or even as a model for the movements of groups of animals or humans (Großmann et al., 2016; Fodor and Marchetti, 2018).

A thing all these applications and many other models based on Markov processes have in common is that they are usually assumed to take place in homogeneous environments. In the case of transcription, this is especially unrealistic. Not only does the DNA sequence itself vary, it is also covered with obstacles, that make the RNAp's movements highly position-dependent. More generally, the considered physical media are uniform on large scales, but do exhibit drastic fluctuations locally.
It would be significantly more realistic and thus highly beneficial to consider the models in inhomogeneous environments as well. Of course, we expect that any small fluctuations in the medium will average out as the particle moves through space. We call this process of evening out "homogenisation". It is difficult to support the expectation of homogenisation with rigorous evidence when working with Markov processes.

We can extract two primary problems from the discussion above.

1. What results can be derived for the general class of layered jump processes and how are these applied to specific applications such as RNA transcription?
2. Do these results still hold when moving from homogeneous to inhomogeneous environments and how can we verify this?

Of course, the above problems are extremely broad. To answer such questions, different, more specific avenues of attack need to be explored. The first problem can already be narrowed down: a particularly useful result for the applications would be the classification of the movements of particles in systems that behave as layered jump processes. The mathematical theory of Markov processes is tailor-made for this purpose and has a highly developed framework for deriving results on long time scales, particularly in the case of single-particle systems. In the late 20th century, many advances were made on central limit scaling theorems for particles in both homogeneous and inhomogeneous media, for instance by the works of (Kipnis and Varadhan, 1986 and Landim et al., 2012). We aim to apply these results to the layered jump processes and find the exact scaling limit for single-particles in these systems. We will consider both homogeneous and inhomogeneous systems to see if homogenisation takes place. Subsequently, these limiting behaviours can be interpreted and used in the related physical applications such as RNA transcription.

While the single-particle systems are intriguing on their own, real physical systems often operate with many interacting particles simultaneously. Although multi-particle analysis is possible, it is far more difficult. A better understanding of the single particle system can serve a stepping stone for further research into multi-particle systems and can certainly aid our understanding of more complex systems. Furthermore, to expand on the mathematical single-particle results, we will perform numerical simulations of multi-particle systems. Results from these simulations can be compared against the theoretical predictions for single-particle systems, but can also be more directly applied to the relevant physical systems. In particular, we will conduct simulations to investigate the effect of RNAp interactions and DNA inhomogeneity on transcription.
Although the basic layered system is already a reasonable model for transcription, it lacks some additional components. For example, RNAps encounter a few insurmountable barriers on the tracks which are not included in the mathematical analysis. Our numerical simulations indicate that the barriers lead to a special phenomenon known as "RNAp cooperation", presented by (Klumpp, 2011; Nudler, 2012).
From the discussion presented above, we can summarise our research on the two primary problems into three main research questions:

1. What is the exact scaling limit for layered jump processes in both homogeneous and inhomogeneous environments?
2. What conclusions can we draw on RNA transcription and the motion of RNAps, based on our mathematical and numerical analysis of the layered jump processes?
3. What additional phenomena do we observe when adding the extra, more realistic components, such as RNAp interactions and diffusion barriers, to our Markov model and how can this be interpreted in the context of RNA transcription?

The outline of this thesis to answer these three questions is as follows.
We shall begin by outlining the preliminary mathematical theory on Markov processes in chapter 1. The definitions and theorems presented there are prerequisites to all subsequent mathematical analysis and should mostly be familiar to the reader. For instance, we will define the exact meaning of stochastic processes, Markov processes and layered jump processes. Secondly, we will introduce the formalism of semi-groups and generators. These extra tools will greatly aid the remainder of the mathematical analysis.

We will follow up on this fundamental mathematical theory by going into further detail on physical applications in chapter 2. We begin this chapter by discussing why Markov processes are widely used to model physical systems. Afterwards, we investigate applications, of which several have already been brought up in this introduction. A particularly archetypal application is Brownian motion. Brownian motion will prove to be the continuous-time counterpart of the Gaussian distribution in the context of central limit theorems and is therefore crucial in all scaling limit theorems of the subsequent chapters. After introducing Brownian motion, we will delve deeper into actual applications of jump processes, such as the molecular motors and RNA transcription. Additionally, we establish a connection between our microscopic model and the modelling of vehicular traffic. This surprising link is important as some of our results from numerical simulations are easiest to interpret in the context of vehicles travelling on a highway.

After familiarising ourselves with the physical systems, we move on to a mathematical analysis of layered jump processes in chapter 3. In particular, we will be considering the limit behaviour of single particle systems, by using ergodic theory and martingales. For both homogeneous and inhomogeneous systems, we prove analogous theorems to the law of large numbers (LLN) and the central limit theorem (CLT). These type of limit theorems are generally referred to as invariance principles.

In the homogeneous case, it is in fact possible to derive exact formulas for the drift velocity of the LLN and the diffusion coefficient of the CLT. The main results concerning this are theorems 24, 30 and 31. These equations can be used to determine the drift of a particle and to estimate the variation around the drift. In the inhomogeneous case, we will consider two types of random environments; namely the random-waiting-time model and the random conductance model. In both cases, it becomes infeasible to actually find closed-form expressions for the diffusion constants, but it is mostly still possible to show abstract limit theorems. This rigorously confirms our expectation of homogenisation. The main results are analogous to the ones for homogeneous environments and are shown in theorems 33 and 36.

These abstract theorems can be practically applied to the physical systems. For instance, the theorems confirm that RTPs are comparable to Brownian particles and that homogenisation is valid for RTPs. For RNA transcription, the drift theorem is more important for practical use, as it provides estimates and bounds for the transcription speed. Chapter 3 concludes the mathematical analysis of our (single-particle) systems and provides answers to the first main research question and partially to the second.

To add to the results of single-particle systems, we conduct simulations on multiple-particle systems as well. In chapter 4, we introduce a general numerical algorithm that can simulate jump-processes. We implement this algorithm in Python and verify the validity of this implementation by confirming our derived mathematical results.

Having introduced a general method of computing layered jump processes, we can finally apply this to our main process of interest, RNA transcription. In chapter 5, we will reproduce results obtained by (Klumpp, 2011) in both homogeneous and inhomogeneous environments. In particular, we find that the phenomenon of clustering occurs in our simple model. Additionally, we find that the addition of certain barriers to our simple model leads to the experimentally observed RNAp cooperation. We also explain how this effect naturally arises due to this added component, thus investigating the third research question. Lastly, we will compare our numerical results against the theoretical results in both sparse and dense conditions. This analysis will provide the main answer to the second research question.

Lastly, in the conclusion 6, we will summarise our findings and discuss any potential shortcomings which indicate a potential direction for future study.

Chapter 1

Mathematical foundation of Markov processes

In the introduction, we outlined the main relevant physical systems. The models for these systems are all stochastic in nature. Even to be able to formulate statements regarding stochasticity, an adequate mathematical foundation is needed. More specifically, the processes will turn out to be Markov processes, for which a vast and rich mathematical framework is readily available. In this chapter, we will therefore lay the groundwork for further analytical investigations and we begin exploring the theory on general Markov processes so that we can apply these results to the physical systems later.

1.1. Probability theory

A first step is to clarify the notion of "stochastic processes". The word "stochastic" belongs to the mathematical domain of probability theory, probability spaces and random variables, which we take as prerequisite concepts. The second part, "processes", refers to time dependent phenomena. We shall now combine these two concepts more formally.

Definition 1. *Stochastic process:*
A stochastic process on probability space $(\Omega, \mathscr{A}, \mathbb{P})$ is a collection of random variables $(X_t)_{t\in T}$ on $(\Omega, \mathscr{A}, \mathbb{P})$. Here, T is a linearly ordered index set (normally $\mathbb{N}, \mathbb{Z}$ or $\mathbb{R}_{\geq 0}$). Each X_t maps to the same pair $(F, \mathscr{F})$ and may depend on previous random variables in the collection. The pair $(F, \mathscr{F})$ is usually the physical space in which the process takes place.
For any $\omega \in \Omega$, the function $X^\omega : T \to S$ with $X^\omega(t) = X_t(\omega)$ is called a sample path.
Furthermore, we call a stochastic process stationary if $\mathbb{P}_{X_t}$ does not depend on t. $\mathbb{P}_{X_t}$ is the push-forward measure on F with $\mathbb{P}_{X_t}(f) = \mathbb{P}\left(X_t^{-1}(f)\right)$ for each $f \in \mathscr{F}$.

From this definition, we find the surprising consequence that the underlying sample space $(\Omega, \mathscr{A}, \mathbb{P})$ is not all that interesting. Since we are more interested in the outcomes of the process and the probabilities of these outcomes occurring, the more important aspects are the index set T, the outcome space $(F, \mathscr{F})$ and the measures $\mathbb{P}_{X_t}$. In fact, the process is often defined directly using the random variables and the push-forward measures themselves, omitting $(\Omega, \mathscr{A}, \mathbb{P})$ altogether.

Definition 2. *Discrete and continuous stochastic processes.*
We say that a stochastic process on $(\Omega, \mathscr{A}, \mathbb{P}) \to (F, \mathscr{F})$ is in discrete space if F is countable. Commonly $F \subseteq \mathbb{N}$ or $F \subseteq \mathbb{Z}$ in this case.
Secondly, a stochastic process is in discrete time if T is countable. On the other hand, it is in continuous time if T is continuous (for our applications, this means $T = \mathbb{R}_{\geq 0}$).

Our main focus will be on continuous time, discrete space stochastic processes. Nevertheless, discrete time processes are often simpler and hence it is insightful to provide a discrete example first.

Example 1. *The simple discrete time random walk:*
We let $(S_n)_{n\in\mathbb{N}}$ *be a stochastic process in discrete space and time, defined by* $S_n = \sum_{i=1}^{n} X_i$. *Here,* $(X_i)_{i\in\mathbb{N}}$ *is a sequence of i.i.d. random variables mapping to* $\{-1,1\}$ *with* $\mathbb{P}_{X_1}(\{1\}) =: p$. *The process can easily be visualised by viewing* S_n *as the position of a particle at time* n. *After each time step, the particle jumps to a new position, making it a random walk. We call the random walk simple since the particle can only jump to neighbouring positions. If* $p = \frac{1}{2}$, *the process is a simple, symmetric random walk. On the complete opposite end of the spectrum, for* $p = 0, p = 1$, *the walk is completely deterministic in one direction.*
This process is defined without referring to a sample space Ω. *In principle,* Ω *can be constructed, but this is cumbersome. The usual method is to take a product space of the underlying sample spaces of the increments* X_i *(Saeki, 1996; Hewitt and Stromberg, 1965).*

1.2. DISCRETE SPACE MARKOV CHAIN BASICS

We are now equipped with the necessary ingredients to turn to Markov chains. These are stochastic processes which additionally satisfy the Markov property. This property entails that the processes only depend on the present, but not the past. This constraint narrows down the scope of stochastic processes to allow for a deep mathematical analysis. However, the property is also not too restrictive, rendering it useful for applications in a wide range of scientific fields. In fact, many physical processes can be modelled as Markov processes (van Kampen, 2007) and we provide several examples in chapter 2.

1.2.1. DISCRETE TIME MARKOV PROCESSES

Definition 3. *Discrete space and time Markov chain:*
Take a stochastic process X_n *in discrete time on* $(\Omega, \mathcal{A}, \mathbb{P})$ *mapping to a discrete space* S.
We call X_n *a Markov process if it additionally satisfies the Markov property:*
If $\forall n \in \mathbb{N}, i_0, i_1, \ldots, i_n, j \in S$ *we have* $\mathbb{P}(X_n = i_n, \ldots, X_0 = i_0) > 0$, *then*

$$\mathbb{P}(X_{n+1} = j | X_n = i_n, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) = \mathbb{P}(X_{n+1} = j | X_n = i_n). \tag{1.1}$$

Note that the condition $\mathbb{P}(X_n = i_n, \ldots, X_0 = i_0) > 0$ is only there to guarantee that equation 1.1 is conditioned on events with positive probability.

Without knowing it, we have already encountered our first discrete time Markov process in the form of the random walk described in example 1. All we need to do, is simply verify that the Markov property is satisfied for this process. Consider any path $i_0, i_1, \ldots, i_n \in \mathbb{Z}$ that satisfies the conditions of equation 1.1 and take $i_{n+1} \in S$ arbitarily. Then,

$$\mathbb{P}(S_{n+1} = i_{n+1} | S_n = i_n, \ldots, S_0 = i_0) = \mathbb{P}(S_n + X_{n+1} = i_{n+1} | S_n = i_n, \ldots, S_0 = i_0) =$$
$$\mathbb{P}(X_{n+1} = i_{n+1} - i_n | S_n = i_n, \ldots, S_0 = i_0) = \mathbb{P}(X_{n+1} = i_{n+1} - i_n | S_n = i_n) = \mathbb{P}(S_{n+1} = i_{n+1} | S_n = i_n).$$

where we use the independence of X_{n+1} to any X_k for $k < n+1$ in the third equality.
The simple random walk also exhibits an important property known as time homogeneity. Time homogeneous Markov processes form an important class of Markov chains.

Definition 4. *(Time) Homogeneous Markov chain:*
A chain for which the transition probabilities $\mathbb{P}(X_{n+1} = j | X_n = i)$ *do not depend on time. By this we mean,* $\forall n, j, i : \mathbb{P}(X_{n+1} = j | X_n = i) = \mathbb{P}(X_1 = j | X_0 = i)$.

Corollary 1. *Chapman-Kolmogorov equation (G. R. Grimmett and Stirzaker, 2001):*
Take a time-homogeneous Markov chain on a finite state space S. Then, the transition probabilities can be written in a matrix as $P = [p_{i,j}] = [\mathbb{P}(X_1 = j|X_0 = i)]$. The n-step probabilities are $p^n_{i,j} = \mathbb{P}(X_n = j|X_0 = i)$ with corresponding n-step matrix $[p^n_{i,j}] = P_n$. Let $\mu = (\mathbb{P}(X_0 = i))_{i\in S}$ be the initial condition. We can compute P_n and $\mathbb{P}(X_n = j)$ by:

$$P_n = P^n, \quad \mathbb{P}(X_n = j) = \left(\mu P_n\right)_j = \left(\mu P^n\right)_j. \tag{1.2}$$

1.2.2. CONTINUOUS TIME MARKOV PROCESSES

Although much more can be said about discrete time Markov processes, we move on to continuous time Markov processes as all our applications will be in continuous time. We will use our discrete time definitions as a basis for the counterparts in the continuous time setting.

Definition 5. *The continuous time Markov property:*
Let $(X_t)_{t\in T}$ be a stochastic process mapping to countable S with $T = [0,\infty)$. X_t satisfies the Markov property if for any sequence $i_1,\dots,i_{n-1},i_n \in S$, $t_1 < t_2 < \cdots < t_n \in T$ we have that:

$$\mathbb{P}(X(t_n) = i_n|X(t_{n-1}) = i_{n-1},\dots,X(t_1) = i_1) = \mathbb{P}(X(t_n) = i_n|X(t_{n-1}) = i_{n-1}) \tag{1.3}$$

In discrete time, we could describe the process completely by its transition matrix P. In continuous time, this becomes more difficult as there is no unit step of time. The role of the transition matrix is assumed by the semigroup and generators in this context.

THE SEMIGROUP AND GENERATORS

Firstly, we shall **assume that the state space S is finite** to illustrate the concepts.

Definition 6. *Transition probability:*
The transition probability $p_{i,j}(s,t)$ for $s \le t \in T$ and $i,j \in S$ is

$$p_{i,j}(s,t) = \mathbb{P}(X_t = j|X_s = i). \tag{1.4}$$

The process is time-homogeneous if the transitions only depend on the time increment:

$$p_{i,j}(s,t) = p_{i,j}(0,t-s) =: p_{i,j}(t-s). \tag{1.5}$$

Definition 7. *Semi-group:*
Take a time-homogeneous Markov chain and let $S_t = [p_{i,j}(t)]_{i,j\in S}$ be a transition matrix. Then, the family $\{S_t : t \ge 0\}$ is called a stochastic semi-group.

Proposition 2.
The semi-group satisfies three basic properties (G. R. Grimmett and Stirzaker, 2001):
1) $S_0 = I$
2) S_t is stochastic, meaning that $\forall i,j : p_{i,j}(t) \ge 0$ and $\forall i : \sum_{j=1}^{|S|} p_{i,j}(t) = 1$.
3) The Chapman-Kolmogorov equation: $S_tS_s = S_{t+s} = S_{s+t} = S_sS_t$ for $s,t \ge 0$.

Proposition 3. *The stochastic semi-group $\{S_t : t \ge 0\}$ is a family of operators on the space of continuous functions mapping from Ω to $\mathbb{R}$, $\mathscr{C}(\Omega)$. For $f \in \mathscr{C}(\Omega)$ define:*

$$S_tf(x) = \mathbb{E}(f(X_t)|X_0 = x) = S_t \cdot (f(x_0), f(x_1),\dots,f(x_n))^T \tag{1.6}$$

where the indexing $S = \{x_0,\dots,x_n\}$ is used on both S and the ordering of elements of S_t.

1

For any process we can easily find its corresponding semi-group $\{S_t : t \geq 0\}$. But, it is especially useful that one can start with a semi-group and guarantee that there actually exists a corresponding process. In example 1, we saw that working with the definitions of stochastic processes can become quite unwieldy. The semi-group simplifies working with this formalism. Additionally, semi-groups have generators associated to them, which track the transitions of X_t on small time scales. In turn, these generators simplify the formalism even further.
In the following sections, we will derive a general form for generators of so-called "jump processes". Using the general form of these generators, all one needs to do to figure out the generator of a specific jump process, is to draw out a kinetic scheme of the process.

Definition 8. *Generator:*

$$L = \lim_{t \to 0} \frac{S_t - I}{t} \tag{1.7}$$

The existence of L is not at all obvious, but we shall only analyze the specific class of jump processes in which L can be directly computed and thus we can safely assume existence.

Theorem 4. *Kolmogorov's Forward equation:*

$$\frac{d}{dt} S_t = S_t L. \tag{1.8}$$

Proof. The forward equation can be derived from proposition 2 (van Kampen, 2007).

$$\frac{d}{dt} S_t = \lim_{h \to 0} \frac{S_{t+h} - S_t}{h} = \lim_{h \to 0} S_t \frac{S_h - I}{h} = S_t L. \qquad \square$$

Kolmogorov's forward equation also became known as the Master equation in the context of discrete space processes in applied sciences. In the case of infinite state space S, the step $\frac{d}{dt} S_t = \lim_{h \to 0} \frac{S_{t+h} - S_t}{h}$ becomes problematic. Under some extra assumptions that we will not delve further into, the Kolmogorov equations still hold (G. R. Grimmett and Stirzaker, 2001). For jump processes, these extra conditions are satisfied.

Theorem 5. *Kolmogorov's Backward equation:*

$$\frac{d}{dt} S_t = L S_t \tag{1.9}$$

Kolmogorov's backward equation is also called the Fokker-Planck equation in applied sciences and can similarly be proven using proposition 2. Note that both Kolmogorov equations describe the evolution of the transition probabilities in terms of the generator. This allows for direct computation of S_t; for instance, by using the following corollary.

Corollary 6. *(Bhattacharya and Waymire, 2023):*

$$S_t = e^{tL} = \sum_{n=0}^{\infty} \frac{t^n}{n!} L^n \tag{1.10}$$

is the solution to the differential equations 1.8 and 1.9 with initial condition $P_0 = I$.

In principle, e^{tL} can be computed by diagonalising L, but this is usually cumbersome and often even impossible. We will provide alternative methods of solving the master equation that are based on Fourier and Laplace transforms in chapter 3.

The generator and semi-group formalism is especially useful when working with irreducible processes, for which further deep and intriguing results can be derived.

Definition 9. *Irreducible chain:*
A Markov chain is irreducible if for any $i, j \in S$, there exists some $t \in T$ such that $p_{ij}(t) > 0$.

Definition 10. *Stationary distribution:*
A vector $\pi = (\pi_1, \ldots, \pi_n)$ with $\forall j : \pi_j \geq 0$ and $\sum_{j \in S} \pi_j = 1$ can be seen as an initial distribution of the chain. Such a distribution is unchanging in time, or stationary, if:

$$\pi = \pi S_t. \tag{1.11}$$

Theorem 7. *Stationarity and the generator (G. R. Grimmett and Stirzaker, 2001):*
For finite state space, π is stationary if and only if $\pi \cdot L = \vec{0}^T$.

Theorem 8. *Ergodic theorem (G. R. Grimmett and Stirzaker, 2001):*
Let X_t be an irreducible Markov chain with associated semi-group $\{S_t\}$, then either:
a) If there exists a stationary distribution π, then it is unique and $p_{ij}(t) \to \pi_j$ as $t \to \infty$.
b) Otherwise, $p_{ij}(t) \to 0$ as $t \to \infty$.

So, no matter the initial distribution of the Markov chain, the eventual distribution converges to the stationary distribution. Theorem 8 can can be further generalised and the main ergodic theorem is given in equation 3.3.

We now turn to the case of more general state spaces, namely countable S. In this case, these processes belong to a class of processes called Feller processes. In particular, we shall be working with $S = \mathbb{Z} \times I$ where I is some finite set of "layers". We can once again define the semi-group, based on the intuition gained from the case of finite state space. This time however, S_t and L cannot be interpreted as matrices, but rather only as operators on function spaces.

Definition 11. *Let S_t be an operator on $\mathscr{C}(\Omega)$ defined by*

$$S_t f(x) = \mathbb{E}(f(X_t) | X_0 = x) = \sum_{y \in S} \mathbb{P}(X_t = y | X_0 = x) f(y). \tag{1.12}$$

Proposition 9. *Once again, as consequences of the definition of S_t we get:*
1) $S_0 = I$
2) $S_{t+s} = S_t S_s$ (Semi-group property / Chapman-Kolmogorov equations)
3) $\forall f : t \mapsto S_t f$ is a right continuous map. (This point is an assumption on the process.)

Corollary 10. *As a consequence of point (3), we once again obtain that $S_t = e^{tL}$ for some generator L. The generator can be computed by using definition 8 and is an operator on the space of functions for which this limit exists.*

1.2.3. JUMP PROCESSES

In the previous sections, we laid out some general results of continuous time Markov processes. In doing so, we have already announced several times that we will only be considering "jump processes". In this section, we will finally clarify what we mean by jump processes. We shall introduce the concept by means of two useful examples and afterwards finally extract a complete definition.

1

A FINITE STATE SPACE EXAMPLE

Let $S = \{0, 1\}$ and let $H_0(0), H_1(0)$ be two positive, continuous random variables. We refer to H_0, H_1 as waiting/holding times. When the process X_t enters state $i \in S$ at time t, we take a new $H_i(t)$, i.i.d. to $H_i(0)$. The particle waits for a time $H_i(t)$ until transitioning to the opposite state. If we demand that X_t is a time-homogeneous Markov chain, it automatically follows that H_i must have an exponential distribution (Sigman, 2022):

$$\mathbb{P}(H_i > x + y | H_i > x) = \mathbb{P}(H_i > x + y | \forall s \in [0, x] : X_s = i)$$
$$\mathbb{P}(H_i > x + y | \forall s \in [0, x] : X_s = i) = \mathbb{P}(H_i > x + y | X_x = i) \quad \text{(Markov property)}$$
$$\mathbb{P}(H_i > x + y | X_x = i) = \mathbb{P}(H_i > y | X_0 = i) \quad \text{(Demand of time homogeneity)}$$
$$\mathbb{P}(H_i > y | X_0 = i) = \mathbb{P}(H_i > y) \quad \text{(by definition of } H_i\text{)}$$

We see that H_i must be memoryless. The only memoryless continuous distributions are exponential distributions with a certain rate parameter λ_i (Sigman, 2022).
Now, let us compute the generator of the process. We first need to establish what the probability of making at least 2 jumps in a time interval $[0, t]$ is. If the process starts from $X_0 = 0$, the probability is of order $O(t^2)$, since:

$$\mathbb{P}(H_0 \le t, H_1 \le t - H_0) = \int_0^t \int_0^{t-h_0} \lambda_1 \lambda_2 e^{-\lambda_0 h_0 - \lambda_1 h_1} \, dh_1 \, dh_0 = O\left(t^2\right).$$

The latter equality follows easily by considering the cases $\lambda_0 = \lambda_1$ and $\lambda_0 \neq \lambda_1$ separately. Here, $O\left(t^2\right)$ denotes the Landau big O, where $f(t) = O(g(t))$ means $\lim_{t \to 0} \frac{f(t)}{g(t)} < \infty$.
The probability of making at least 2 jumps starting from $X_0 = 1$ yields the same expression with λ_0 and λ_1 exchanged. We let K_t denote the amount of jumps up to time t and compute the generator by using the tower rule (Jacod and Protter, 2004):

$$S_t f(i) = \mathbb{E}(f(X_t) | X_0 = i) = \mathbb{E}(\mathbb{E}(f(X_t) | K_t = k) | X_0 = i). \tag{1.13}$$

Note, $\mathbb{P}(K_t \ge 2 | X_0 = i) = O(t^2)$ and $\mathbb{P}(K_t = 0 | X_0 = i) = e^{-\lambda_i t}$.
Therefore, $\mathbb{P}(K_t = 1 | X_0 = i) = 1 - e^{-\lambda_i t} + O(t^2)$. We conclude that:

$$\begin{aligned}
\mathbb{E}(f(X_t) | X_0 = i) &= \mathbb{E}(f(X_t) | K_t = 0, X_0 = i) e^{-\lambda_i t} + \mathbb{E}(f(X_t) | K_t = 1, X_0 = i)(1 - e^{-\lambda_i t}) + O(t^2) \\
&= f(X_i) e^{-\lambda_i t} + f(X_{1-i})(1 - e^{-\lambda_i t}) + O(t^2) \\
&= f(X_i)(1 - \lambda_i t) + f(X_{1-i}) \lambda_i t + O(t^2).
\end{aligned}$$

Consequently, the generator is

$$L = \lim_{t \to 0} \frac{1}{t} \left(\begin{bmatrix} 1 - \lambda_0 t & \lambda_0 t \\ \lambda_1 t & 1 - \lambda_1 t \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + O(t^2) \right) = \begin{bmatrix} -\lambda_0 & \lambda_0 \\ \lambda_1 & -\lambda_1 \end{bmatrix} : \lambda_0, \lambda_1 > 0. \tag{1.14}$$

We can repeat the procedure on a space with three states, $S = \{0, 1, 2\}$, with holding times H_0, H_1, H_2. We distinguish between the possible transitions by letting $\frac{\lambda_{i,j}}{\lambda_i}$ for $i \neq j$ be the probability of transitioning from state i to state j upon a jump. Once again, the result will be that all H_i are exponentially distributed with parameter λ_i. This yields the generator:

$$L = \begin{bmatrix} -\lambda_0 & \lambda_{0,1} & \lambda_{0,2} \\ \lambda_{1,0} & -\lambda_1 & \lambda_{1,2} \\ \lambda_{2,0} & \lambda_{2,1} & -\lambda_2 \end{bmatrix}. \tag{1.15}$$

The tower rule method can in fact be used to compute L on any finite space S. Again, let λ_i, $\frac{\lambda_{ij}}{\lambda_i}$ be the holding time rate and the transition probabilities at state i respectively.

$$(S_t - 1)f(i) = (e^{-\lambda_i t} - 1)f(i) - (1 - e^{-\lambda_i t}) \sum_{j \neq i} \frac{\lambda_{ij}}{\lambda_i} f(j) + O(t^2) = \lambda_i t f(i) + \sum_{j \neq i} \lambda_{ij} t f(j) + O(t^2)$$

$$\implies Lf(i) = \sum_{j \neq i} \lambda_{ij}(f(j) - f(i)). \tag{1.16}$$

where we used that $\sum_{j \neq i} \lambda_{i,j} = \lambda_i$. It is easily verified that the generators in equations 1.14 and 1.15 also satisfy 1.16 when interpreting f as a vector $(f(1), f(2), f(3))^T$.
We can re-interpret this generator by using two properties of exponential distributions:

Theorem 11. *Minimum of exponential distributions:*
Take an arbitrary state i. Let $H_{i,1}, \ldots, H_{i,i-1}, H_{i,i+1}, \ldots, H_{i,n}$ be independent, exponentially distributed variables with respective rates $\lambda_{i,1}, \ldots, \lambda_{i,n} \geq 0$. Let $H = min\{H_{i,j} : j \neq i\}$.
Then H is also an exponentially distributed random variable with rate $\lambda_i := \sum_{j=1, j \neq i}^{n} \lambda_{i,j}$.

The shorthand notation $\lambda_{i,j} = 0$ denotes an impossible transition, i.e. "$H_{i,j} = \infty$".

Theorem 12. *Relating minimum of exponential distributions to components:*
Let $i, H_{i,j}, H$ be defined as in the previous theorem.
Then, $\mathbb{P}(H = H_{i,j}) = \frac{\lambda_{i,j}}{\lambda_i}$ for any $j \neq i$ such that $\lambda_i \neq 0$. If $\lambda_i = 0$, there are no transitions possible exiting state i (i.e. "$H = \infty$") and the process will remain in i indefinitely.

These two results are standard results from probability theory (G. Grimmett and Welsh, 2014) and therefore we do not include the proofs. Equipped with these two properties, the nature of the jumping process can be seen in a different light.

Proposition 13. *We can re-interpret the finite state jumping process as follows:*
Every state has a finite set of neighbours. A transition from state i is always to one of its neighbours j. For each neighbour, we set a exponentially distributed clock with rate $\lambda_{i,j}$ (i.e., a holding time). Then, we transition to a new state as soon as one of the clocks "goes off" and move to the state corresponding to this first clock.
Consequently, the process exits state i with total rate λ_i by theorem 11 and transitions to state j with probability $\frac{\lambda_{i,j}}{\lambda_i}$ by theorem 12. Hence, this new interpretation is completely equivalent to the earlier presented formalism and thus yields the same generator L.

The benefits of this second interpretation will become apparent when we explain the Gillespie algorithm in chapter 4. Another advantage of the second interpretation is that it can be related to actual physical systems, in which particles transition similarly to the "clock formulation" given above.

Before we continue to the more general Feller processes on countable state space S, we derive some final results for finite state spaces. In particular, we need to find the stationary distributions of the 2-state and 3-state systems. These will be used to calculate the drift of physical jump processes in section 3.3.1.

1

Theorem 14. *Stationary distribution of two state system:*
The stationary distribution π of processes on $S = \{0,1\}$ with generator 1.14, is given by:

$$\pi = \left(\frac{\lambda_1}{\lambda_0 + \lambda_1}, \frac{\lambda_0}{\lambda_0 + \lambda_1} \right) \tag{1.17}$$

Proof. If $\lambda_1, \lambda_2 > 0$, the process is irreducible (any state can be reached by any other state). By theorem 7, the stationary distribution can be found by solving $\pi \cdot L = \vec{0}^T$. The proposed vector π of equation 1.17 is a solution to the resulting system of equations:

$$\pi_0 + \pi_1 = 1, \qquad \pi_1 \lambda_1 - \pi_0 \lambda_0 = 0. \qquad \square$$

The assumption $\lambda_0, \lambda_1 > 0$ is necessary in this theorem. If either of the two are in fact 0, the process will eventually remain stuck in one of the states forever, which is automatically a trivial stationary distribution.

Theorem 15. *Stationary distribution of three state system:*
The stationary distribution π of irreducible processes with generator 1.15 on state space $S = \{0,1,2\}$ is given by:

$$\pi = K(\lambda_{1,2}\lambda_{2,0} + \lambda_{2,0}\lambda_{1,0} + \lambda_{1,0}\lambda_{2,1}; \lambda_{2,0}\lambda_{0,1} + \lambda_{2,1}\lambda_{0,2} + \lambda_{2,1}\lambda_{0,1}; \lambda_{0,2}\lambda_{1,0} + \lambda_{1,2}\lambda_{0,2} + \lambda_{1,2}\lambda_{0,1}) \tag{1.18}$$

with K being a normalization constant.

Proof. Again, this is the solution to the system $\sum_j \pi_j = 1, \pi \cdot L = \vec{0}^T$. $\square$

Once again, we need the assumption of irreducibility in this theorem. For non-irreducible processes, the process will eventually reduce to the 2-state system.

Note that by theorem 8, any irreducible process in the 2- and the 3-state systems will converge to the stationary distribution, no matter the initial condition $\mu = (\mathbb{P}(X_0 = i)_{i \in S}$.

Proof.

$$p_{ij}(t) \to \pi_j \implies (\mu S_t)_{i \in S} \to \pi_i \sum_{j \in S} \mu_j. \qquad \square$$

THE RANDOM WALK FELLER PROCESS

We have just established the general form of generators in finite state spaces, by defining time homogeneous Markov processes and their transitions using holding times. For countable spaces, we can work our way backwards; starting with the generator to define the process and its transitions.
First, we consider a continuous time variant of the simple random walk (example 1).
Let $S = \mathbb{Z}$ and let take symmetric transition rates $\lambda_{x,x+1} = 1 = \lambda_{x,x-1}$. The generator

$$Lf(x) = 1 \cdot (f(x+1) - f(x)) + 1 \cdot (f(x-1) - f(x)) = (f(x+1) + f(x-1) - 2f(x))$$

defines a simple, symmetric continuous time random walk. Particularly intriguing here is the striking similarity to a second-order central difference, which is the discrete analogue to the second derivative of a function. Any term in the generator that has this central difference form is referred to as a diffusive component of the process.

This name is not a coincidence, as there is a somewhat surprising connection between these generators and diffusive processes. For instance, (Schwarz, 2022) shows that an actual diffusion equation in the form of Fick's second law of diffusion can be derived from such difference equations. The probabilistic process most closely related to diffusive motion is Brownian motion. This process has an actual second derivative in the generator which will be of significance in section 2.2. The random walk above can be seen as the discrete time variant of this process due to the similarities in the generator. As opposed to central differences, forward differences may also appear in the generator. These parts are called active components and we will provide an example below.

GENERAL JUMP PROCESSES

Definition 12. *Let S be a countable state space. For each $x \in S$, assign a finite set of neighbours N_x. For each $y \in N_x$ take a rate $\lambda_{x,y} > 0$ and let $\lambda_x := \sum_{y\in N_x} \lambda_{x,y}$. When a process starts in x, we make a jump to any neighbour y with probability $\frac{\lambda_{x,y}}{\lambda_x}$ and total rate λ_x. We call any process based on these transitions rates a **jump process**. Jump processes have a generator operating on the functions with bounded supremum norm:*

$$Lf(x) = \sum_{y\in N_x} \lambda_{x,y}(f(y) - f(x)). \tag{1.19}$$

Furthermore, we call the process a layered jump process if $S = \mathbb{Z} \times I$ for some finite I. A particularly useful layered jump process is described by the following generator:

$$\begin{aligned} Lf(x,\sigma) =& \kappa_\sigma(f(x+1,\sigma) + f(x-1,\sigma) - 2f(x,\sigma)) + f(x+v_\sigma,\sigma) - f(x,\sigma) \\ &+ \sum_{\tau\neq\lambda} \lambda_{\sigma,\tau}(f(x,\tau) - f(x,\sigma)) \end{aligned} \tag{1.20}$$

*where $\sigma \in I$ denotes the layer. So, this is a process in which each layer has a distinct diffusive term with rate κ_σ and an active component of size v_i and rate 1. **From now on, we will only be considering layered jump processes.** An example is given in figure 1.1.*

Proof. The derivation of the generator is the same as in equation 1.16. Again, we use that the probability of more than one jump in a time interval $[0, t]$ is $O(t^2)$. Furthermore, from the fact that $|N_x|, ||f||_\infty < \infty$, we can safely conclude that the sum converges. □

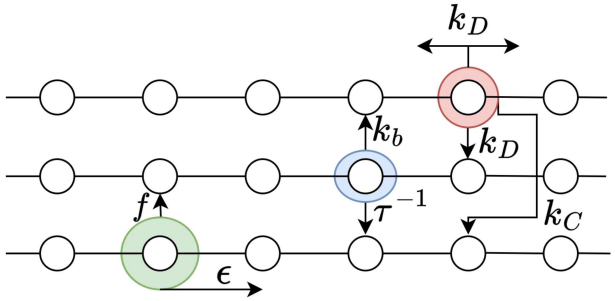


Figure 1.1: A three-layer system used for modelling RNA transcription in chapter 5. We have diffusive motion in the upper layer, active motion of rate ϵ in the lowest layer and the second layer is a pause layer. The parameters for this specific model are given by $v_0 = 1, \kappa_0 = \kappa_1 = v_1 = v_2 = 0, \kappa_2 = k_D$ and switching rates $\lambda_{0,1} = f, \lambda_{1,0} = \tau^{-1}, \lambda_{1,2} = k_b, \lambda_{2,1} = k_D, \lambda_{2,0} = k_C, \lambda_{0,2} = 0$.

CHAPTER 2

MARKOV PROCESSES IN PHYSICS

Now that we have covered the necessary background in probability theory, it is about time that we actually defend why we consider Markov processes at all. We will begin with a justification of the use of stochastic processes in physical applications. Afterwards, we show that Markov processes are widely used in systems of all scales, such as in the fields of nanobiology and life/behavioural sciences. The particular application we focus on is RNA transcription and we introduce the basic theory of transcription in this chapter.

2.1. WHY DO WE NEED STOCHASTIC PROCESSES IN PHYSICS?

One of the fundamental principles associated with classical physics is determinism (Van Strien, 2021). The exact future trajectory of a particle can, in theory, be determined using Newton's laws, if its position, velocity and all forces acting upon it at a certain time are known. Inherently, there is no randomness about such motion, but determinism does trivially guarantee the Markov property.

However, for complicated systems with many interacting particles, actually analytically solving the equations of motion becomes impossible. In fact, for physical systems such as the ideal gas model, even numerically solving the trajectories of the enormous amount of particles involved is infeasible. As a result, we need to forfeit the idea of keeping track of the exact motion of each and every particle and instead work with only a few macroscopic variables. The justification for being able to filter out so much information, is that we expect the movements of the particles to average out amongst each other, resulting in measurable realisations of macroscopic quantities such as pressure and temperature. The transition from microscopic to macroscopic often comes hand in hand with introducing stochasticity into the model (van Kampen, 2007). Instead of following what state the particle *is in*, we only speak of what state the particle *is likely to be in* and assume that all particles follow the same probability distribution. Another benefit of this probabilistic formulation is its aptness to modern physics and quantum indeterminancy.

A problem that arises when restricting ourselves to macroscopic variables is that the Markov property may become invalidated. Indeed, this property may have depended on some of the relinquished information. We illustrate this idea by an example:

Example 2. *Consider a particle in a simple harmonic potential* $U(x) = \frac{1}{2}x^2$*. The trajectory* $(x(t), x'(t))$ *is a Markov process, as two initial conditions at each time* t *are sufficient to solve the equation of motion. However, if we restrict ourselves to the process* $(x(t))$*, we lose the Markov property. Indeed, for a particle at* $x(0) = 0$*, the future trajectory is highly dependent on the history. For instance, the histories* $\forall\, t \in [-2\pi, 0] : x(t) = 0$ *and* $x(t) = sin(t)$ *lead to vastly different future trajectories.*

Hence, the validity of Markov property is an assumption in our models. Nevertheless, when transitions are caused by molecular collisions (such as in chemical reactions), the transitions can be approximated as being exponentially distributed (Gillespie, 2007) and are thus Markovian by our discussion in proposition 13.

For the case of RNA transcription, approximating the transitional process by exponential distributions is further justified in (van Kampen, 2007; van den Berg, 2017). During transcription, the RNA polymerase traverses an energy landscape that contains several energetically favourable positions on the DNA. RNA is transcribed at these DNA sites. To move between two adjacent sites, RNAp has to cross a certain free energy barrier of size ΔG. For instance, the particle has to perform work to open the double-stranded DNA and to overcome viscous drag due to sliding along the DNA (Jülicher and Bruinsma, 1998). Note that ΔG may depend on the DNA sequence. In fact, this dependence may be a way for RNAp to identify and copy the code (Jülicher and Bruinsma, 1998).
To overcome the energy barrier, energy is attained through thermal noise and the energy-releasing binding of NTP to the RNA transcript (Jülicher and Bruinsma, 1998).
We can view the thermal noise as the polymerase being bombarded by numerous smaller molecules. Upon collision, the polymerase transitions to excited energy states. After several collisions, the polymerase gains enough energy to cross the energy barrier and travels to the next site on the DNA chain. The transitions to excited states approximately follow the simple master equation $\frac{dp}{dt} = -kp$ (van den Berg, 2017), which has an exponential as solution. The Arrhenius-Boltzmann factor, $e^{\frac{-\Delta G}{k_B T}}$, relates to the probability of an excited particle having a change of ΔG in energy (van Kampen, 2007). By combining the transition master equation with the probability of the transition having enough energy to cross the barrier, it follows that the total transition of RNAp to the next site approximately occurs at an exponential rate.

We have thus shown that Markov processes are suitable for modelling physical systems. We will now show examples of such models and introduce the model of transcription.

2.2. BROWNIAN MOTION

Brownian motion is an archetypal example of continuous time Markov processes in physics. It is particularly of interest to us, since our objective is to prove that the rescaled movements of particles in several systems converge to Brownian motion.

BROWNIAN MOTION IN PHYSICS

The discovery of Brownian motion is often attributed to Robert Brown. In 1827, Brown investigated the movement of pollen suspended in water and noticed tiny particles moving in an erratic, unpredictable fashion (Schwarz, 2022). The movement of a heavy particle, such as the pollen, immersed in a fluid of light particles is now called Brownian motion (van Kampen, 2007). In our applications, we will only work with one-dimensional Brownian motion. However, we emphasise that the following discussion generalises to higher dimensions in a straightforward manner by letting each coordinate behave independently as a one-dimensional Brownian motion.

As explained in (van Kampen, 2007), when the heavy particle has a positive velocity V, it will collide with more molecules in front than in the back. The amount of collisions and the subsequent values of velocity depends on the present V, but not on the past. The process $\{V(t)\}$ is therefore Markovian. We can also observe the position at some

time steps, $X_{t_0}, X_{t_1}, \ldots$, for times that are not too close. By this we mean that the time increments are larger than the self-correlation time of the process (van Kampen, 2007). For such time increments, the differences $X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, \ldots$ will all be independent. Additionally, each difference only depends on the time duration $t_{i+1} - t_i$ but not on its history. Hence, the position at these time steps is also Markovian.

Famously, Albert Einstein published in his "miracle year", 1905, a theoretical investigation on the displacement and the diffusion coefficient D of Brownian particles, using a molecular collision model as basis. Recall that we are dealing with stochastic processes rather than exact motions. As such, Einstein provides an equation for the (probability) density $\rho(x,t)$ of particles at each position and time (x,t) (Einstein, 1905):

$$\frac{\partial \rho(x,t)}{\partial t} = D\frac{\partial^2 \rho(x,t)}{\partial x^2}. \tag{2.1}$$

Of course, this equation is a diffusion equation and a particular solution with initial condition $\rho(x,0) = \delta_x$ is given by (Einstein, 1905):

$$\rho(x,t) = \frac{1}{\sqrt{4\pi D t}} e^{\frac{-x^2}{4Dt}}. \tag{2.2}$$

Here, we recognise a Gaussian distribution function with expectation 0 and variance $2Dt =: \sigma^2 t$. Einstein concludes that the average displacement of a single particle is proportional to $\sqrt{t}$. This conclusion relies on an ergodic property: the behaviour of the same particle in different time periods is interchangeable with the behaviour of different independent particles at the same time. The ergodic property allows for exchanging time averages with sample/ensemble averages, as we will see in theorem 18.

BROWNIAN MOTION IN MATHEMATICS

Brownian motion also holds a central place in mathematics and is known as the Wiener process in this context. We will later see that in continuous-time central limit theorems (CLTs), the Wiener process assumes the same role the Gaussian has in discrete-time CLTs. Therefore, it is important to formalise the discussion presented above.

Definition 13. *Wiener process (G. R. Grimmett and Stirzaker, 2001):*
The Wiener process $\{W(t) : t \geq 0\}$ *is defined by three basic properties:*
1) The trajectories $t \to W(t)$ *are continuous.*
2) The initial condition is given by $W(0) = 0$
3) The increments in position are ***independent*** *and* ***normally distributed****, i.e.*

$$\forall (t_i)_{i=1}^n : 0 \leq t_1 < t_2 < \ldots, t_n, \forall i \leq n : W(t_i) - W(t_{i-1}) \sim \mathcal{N}\left(0, \sigma^2(t_i - t_{i-1})\right).$$

The Wiener process satisfies some basic, desirable properties such as existence, well-definedness and the Markov property (G. R. Grimmett and Stirzaker, 2001). Additionally, the physically derived equation 2.2 follows directly from property (3). Thirdly, the generator of the process can be found by comparing equations 1.8 and 2.1. We conclude that $L = \frac{1}{2}\sigma^2 \frac{\partial^2}{\partial x^2}$. Once again, this confirms the diffusive nature of this process, as this is the continuous space analogue of the discrete diffusion generator given in section 1.2.3.

Lastly, we need to establish some computational properties of the Wiener process that will be used later in the central limit theorem proofs of chapter 3.

Proposition 16. *For a certain time t, the characteristic function (CF) of a random variable $W(t)$ from the Wiener process $\{W(t)\}$ is given by:*

$$\phi_t(s) := \mathbb{E}\left(e^{iW(t)s}\right) = e^{-Dts^2}. \tag{2.3}$$

This can be directly found from the CF of Gaussian distributed random variables (Jacod and Protter, 2004). Consequently, the Laplace transform of these CFs is given by:

$$\mathscr{L}\left(\phi_t(s)\right)(z) = \int_0^\infty e^{-zt} e^{-Dts^2}\, dt = \frac{1}{z + Ds^2}. \tag{2.4}$$

On the other hand, the Wiener process also has some unexpected properties. For example, almost all sample particle paths of the process are nowhere differentiable, which seems contradictory to the motion it is trying to model. The nowhere-differentiability is most clearly exemplified by the scaling property of Wiener processes (Lowther, 2009).

Proposition 17. *Scale invariance (Lowther, 2009):*
If $W(t)$ is a Wiener process, then $\forall a \in (0,\infty) : W_a(t) := aW(a^{-2}t)$ is also a Wiener process.

By Taylor's theorem, differentiable processes look approximately linear when zoomed in close enough. By property 17, it is clear that this is not true for the Wiener process.
The apparent contradiction can partially be attributed to the fact that in the model, the molecular collisions are assumed to have an instantaneous effect on the heavy particle. In reality, the collisions act for a nonzero, albeit small duration.

2.3. LAYERED JUMP PROCESSES IN PHYSICAL SYSTEMS

In the previous section, we presented Brownian motion as an example of a continuous time and space process. However, the main focus of this report lies on discrete space Feller processes. In section 2.1, we already expanded on the idea that exponential transitions often arise in molecular systems and chemical reactions. Therefore, we can expect that microbiological systems are the main application of Feller processes in physics. Interestingly enough, macroscopic applications also exist. In particular, jump processes have found a remarkable extent of applicability in life and behavioural sciences (Schwarz, 2022). Before we present the main matter on applications in microscopic biology, we will therefore first discuss some applications on larger scales.

2.3.1. JUMP PROCESSES IN LIFE AND BEHAVIOURAL SCIENCES

(Schwarz, 2022) presents some particularly striking macroscopic examples, like modelling the movements of animals such as reindeer, the length of a hospital stay and the decision making process of a bee colony as a random walk.
More specifically tied to the Feller processes of section 1.2.3 is the notion of discrete and continuous time stochastic cellular automaton models (Klumpp and Hwa, 2008). These automaton models appear in all sort of macroscopic contexts, most notably in modelling vehicular traffic. The study (Chowdhury et al., 2000) provides an extensive overview of

several traffic models, including "particle-hopping models" that are comparable to jump processes. In fact, viewing traffic from a statistical physics perspective is actually common practice (Chowdhury et al., 2000). Theories on kinetic gas models, interacting particle systems and compressible fluids lend themselves to use in traffic models and vice versa. We emphasise the strong and surprising connection to this application, since we will later see phenomena in chapter 5 that one would also expect from a traffic model.

2.3.2. RUN AND TUMBLE PARTICLES

Now, we move on to microscopic applications. The Feller processes studied in section 1.2.3 focus on layered jump processes. The different layers of the process can be used to represent the internal state of a particle. A notable example of particles with different internal states are so-called run and tumble particles (RTPs). These are a class of "self-propelled Brownian particles that exhibit an interplay between random fluctuations and active swimming" (Bechinger et al., 2016). Examples of RTPs are bacteria such as motile *E. coli* and *M. xanthus* and algae such as *C. reinhardtii* (Bechinger et al., 2016; Großmann et al., 2016). Aside from biological swimmers, the development of artificial RTPs has recently progressed rapidly too (Bechinger et al., 2016).

The random fluctuations in the motion of RTPs are diffusive in nature and caused by interactions with the medium, similar to the Brownian particle. The active part is performed by the particle itself, by converting ATP from the environment into (kinetic) energy. The active motion causes the particle to drift linearly in a certain direction. From time to time, the particle reorients itself (tumble) and starts drifting in a different direction. The different internal states of a layered jump process lend themselves perfectly for modelling the different possible drift directions of the RTPs, as each layer has its own type of active and diffusive motion. Especially the *M. xanthus* bacteria behaves similarly to a one-dimensional layer process, as the bacteria completely flips direction when reorienting (Mignot, 2007). However, when considering a system with many bacteria, the movement will still be higher-dimensional due to collisions.

There are also some differences between the RTPs and the layer processes. The motion of RTPs mostly takes place in 2 or 3 dimensional continuous space, while the mathematical theory under consideration is in 1-dimensional discrete space. As the continuous space can suitably be discretised by a lattice, the fact that jump processes are in discrete space is not problematic. In principle, the mathematical theory can also be expanded to 2 and 3-dimensional systems, but this would be highly case-dependent. For example, the way in which the particle reorients itself needs to be incorporated into the layers and this may lead to different models for different RTPs. In (Großmann et al., 2016; Fodor and Marchetti, 2018), some higher-dimensional mathematical models of RTPs and further examples of bacteria, cells and even groups of animals or humans that move like RTPs are presented. Since the focus of this thesis lies on the 1-dimensional process of RNA transcription, we will not generalise the theory to higher dimensions here.

2.3.3. RNA TRANSCRIPTION

In the introduction, we outlined the process of protein synthesis in a very rudimentary form: cells produce proteins based on the DNA via the stages of transcription and translation. We shall now describe each step in transcription in more depth.

DNA AND RNA

As is commonly known, DNA encodes the genetic information of any individual. The DNA is a long chain of sequences of nucleotides, with 4 possible constituents: guanine (G), cytosine (C), adenine (A) and thymine (T). These nucleotides provide the blueprint for the synthesis of proteins, although only a small portion of the DNA (1.5 % in humans) is actually protein code (Nakata et al., 2021). The specific protein coding sequences are called genes. Even inside a gene, not all nucleotides are used for protein generation. The parts that do encode protein information are called exons, while the parts that do not go into the eventual RNA are introns.

The DNA has a distinct double helix structure, in which two nucleotide sequences are placed in complementary form. By this, we mean that the nucleotides in direction opposition always come in base pairs, that are either G and C or A and T. The double helix structure provides chemical stability through hydrogen bonds but also allows for error-correction as the same information is stored twice in the DNA. Both strands of the DNA have a polarity related to their internal structure, which run in opposite directions compared to each other. During the replication of DNA, both strands are copied simultaneously by a family of proteins called DNA polymerases.
Although we have presented DNA as a sort of long chain, it can actually spatially bend and twist too. To efficiently compactify the DNA, and to protect DNA against tangling and damage, it is wrapped around packaging proteins called histones. A structure of the DNA wrapped around a histone is called a nucleosome. In turn, a compact structure of multiple nucleosomes is called a chromatin.

In transcription, the DNA is used to construct a corresponding RNA strand. An RNA strand is similar to a single DNA strand but with thymine replaced with the base uracil (U). Although there are in fact multiple types of RNA, the messenger RNA or mRNA is the most important for transcription, as the mRNA actually encodes the information needed for protein synthesis. Similarly to the replication of DNA, the creation of RNA from DNA is done by RNA polymerases and RNA polymerase (RNAp) II is responsible for mRNA.

RNA TRANSCRIPTION

The transcription of RNA from DNA consists of three stages; initiation, elongation and termination. Out of these three, the elongation phase will be of main interest to us as this is a direct application of jump processes in a physical system.

During initiation, RNAp binds to one DNA strand in specific “promoter regions”. The availability of promoter regions is controlled by proteins called transcription factors (TFs). As such, TFs regulate which genes are transcribed and the rate at which this occurs.

After initiation, the RNAp becomes active through use of NTP and ATP and can begin transcribing the gene (Sims et al., 2004). This phase is called elongation. Due to the polarity of the DNA strand, the RNAp moves in a specific direction along the chain. The double helix of the DNA is “zipped open” in the direction of operation, after which the RNAp copies the nucleotides onto an RNA strand in a stepwise manner. Once the DNA is transcribed, the DNA behind the RNAp is “zipped closed”. The RNAp’s forward motion

is interspersed with pauses, which are for example used to benefit genome stability (van den Berg, 2017; Jonkers and Lis, 2015) and to ensure proper pre-mRNA modifications are made (Sims et al., 2004). The RNAp can return from a pause to the active elongation state or switch to a backtracking pause. During a backtracking pause, the RNAp moves in a diffusive fashion while being detached from the RNA transcript (Klumpp, 2011). This backtracking movement is theorised to serve the function of correcting any potential earlier mistakes (Nudler, 2012). The site where an RNAp starts a backtrack is called the initiation site. The RNAp can exit a backtracking state via two mechanisms. The RNAp can return to the initiation site and switch back to the pause state. Otherwise, the RNAp can cleave the transcript and re-engage transcription from any earlier position on the DNA chain. Note that the RNAp can never move past the initiation site in a backtrack, as the transcript and zipped open DNA prevent it from moving further (van den Berg, 2017). A schematic drawing of the process of elongation can be found in figure 5.1.

The RNAp completes gene transcription once it reaches the end of the chain, which is signaled by specific nucelotide sequences called termination regions. When the RNAp terminates, it detaches itself from the DNA chain and finishes the pre-mRNA transcript is finished. This transcript is called pre-mRNA, as it still has to go through a splicing process of clearing out the excess intron regions before it is entirely mRNA.

REGULATING TRANSCRIPTION

The cell needs to be able to adjust to the need of specific proteins at any moment. Therefore, control of the rate of transcription of different genes is required.
Transcription is a highly complex process subject to many influencing factors, but there are two main mechanisms the cell can use to regulate transcription. Firstly, the amount of polymerase binding to specific genes can be controlled. Secondly, the speed at which the RNAp traverses the DNA strand can be regulated as well. The latter is for example important to synchronise the transcription with the speed of translation (Klumpp, 2011).

We shall now introduce some of the influencing factors on which the rate of transcription is dependent. These influences can be divided into two categories: influences that stem directly from the genetic information of the DNA and influences that are more from external, which we call epigenetic. The epigenetic information therefore involves all features that cannot be directly seen from the nucleotide sequence, such as the spatial organization of the DNA and the surrounding environment.

RNAp binding is mostly controlled by epigenetic influences such as the TFs. Additionally, the spatial structure of the DNA and the manner in which it is wrapped around histones have a major impact on binding. If the gene is wrapped tightly around the histone, it becomes less accessible to RNAps and less RNAps will bind to it (Nakata et al., 2021).

The transcription speed itself is influenced by both genetic and epigenetic information. The main epigenomic influences are again histones and TFs. Even when the RNAp is on the DNA chain, it can still be blocked or hindered by histones in specific positions on the DNA. As a result, the RNAp slows down or even halts completely when encountering histones on the track (van den Berg, 2017; Li et al., 2007).

TFs can also somewhat control the dynamics of RNAp as they can attach to specific DNA sites and form so-called regulatory regions (Merkenschlager and Odom, 2013). For example, the protein CTCF can bind to the DNA and slow down transcription (Merkenschlager and Odom, 2013). Consequently, the RNAp behaves differently at different positions on the DNA chain. TFs can even influence the rate at which the RNAp enters pauses at specific sites (van den Berg, 2017). "Indeed, upon collisions with different DNA-bound proteins, ..., RNAp backtracks in vitro and in vivo" - (Nudler, 2012).
(Sims et al., 2004) aptly summarises the effect of genetic and epigenomic properties on pausing as follows: "Transcriptional pause and arrest in vivo are most likely caused by a combination of identifiable DNA sequences, protein factors, and the nascent transcript." Lastly, the manner in which RNAp moves on the DNA strand is also directly dependent on the genetic nucleotide sequence. For instance, it has been experimentally observed that the RNAp moves differently on exon and intron parts of the sequence (Nakata et al., 2021; Kolasinska-Zwierz et al., 2009; Sneppen et al., 2005).

INTERACTING POLYMERASE

A gene is typically thousands of nucleotides in length. On the other hand, an RNAp protein is about 30-50nt long and copies only one nucleotide at a time (Nakata et al., 2021; Klumpp, 2011). The RNAp operates at a velocity of approximately 25-65 nt per second (Sneppen et al., 2005). This means that when many proteins are required at a certain moment, a single transcribing RNAp may not be sufficient. Fortunately, multiple RNAps can work the same track simultaneously. These RNAps can interact with each other. We shall assume a simple form of RNAp interaction, namely exclusion interaction.

Definition 14. *Exclusion interaction:*
Two particles cannot occupy the same position at the same time. Hence particles cannot move to a position that is already taken by another particle.

In reality, the RNAp has more complex interactions, such as pushing forces and collisions. In fact, RNAp may even collide with RNAp that is working on the other DNA strand in the opposite direction (Sneppen et al., 2005). Exclusion is only an approximation of these interactions, but still leads to interesting phenomena. For example, (Klumpp, 2011) shows that RNAps can block each other from making long backtracks, thus leading to shorter backtracking pauses in general. As a result, multiple RNAps can work together to increase the rate of transcription along the gene and we call this effect "RNAp cooperation". This does come at the cost of accuracy though, as transcript cleaving occurs less frequently. The cooperation between RNAps has also been observed in experiments on *E. coli* bacteria and yeast (Nudler, 2012). A main goal of this thesis is to replicate this result obtained by (Klumpp, 2011) and to extend its validity to inhomogeneous media as well.

One of the consequences of the exclusion interaction is that RNAps cannot overtake each other during transcription along the chain. In reality, DNA has a spatial structure, which allows for nucleotides that are far removed in the DNA sequence to be spatially close. Consequently, RNAp shortcutting a part of the DNA and thus potentially overtaking other polymerase is possible (Nakata et al., 2021). However, we shall neglect this phenomenon and work with the simplifying assumption that RNAp does not shortcut.

In similar vain, a process that arises due to the spatial structure of DNA is the re-entering of terminated RNAp to the start of the gene. This might occur if both ends of the DNA chain are spatially close and allows for modelling transcription with periodic boundary conditions (Nakata et al., 2021). We also do not include this process in our model and will instead work with open boundary conditions. By this, we mean that RNAp is free to enter at the start of the chain and terminates after reaching the end of the chain, but these processes are not directly dependent. We will assume that new RNAp binds to the chain at rate α. This parameter is also called the initiation attempt rate.

RELATING TRANSCRIPTION TO JUMP PROCESSES

The main aspects of RNAp elongation are excellently encapsulated by the jump process formulation of section 1.2.3. First of all, the RNAp also transitions from state to state with exponential rates by our discussion in section 2.1. Secondly, the state space the RNAp operates on is the DNA, which is naturally one-dimensional and discrete. Thirdly, the RNAp operates with 3 internal states that can be modelled as layers in the jump process.

Despite the common basis to both RNAp elongation and jump processes, there are some important differences as well. In particular, the motion of RNAp is more complicated and the basic jump process formulation simplifies transcription in a couple of ways. We already discussed that the transitions of RNAp are sequence-dependent. For instance, this is due to histones and TFs being located at specific DNA sites. The nucleotide sequence itself and its introns and exons also add to the sequence-dependence. Therefore, the spatially homogeneous generator of equation 1.20 is not completely adequate. In this thesis, we will also introduce space-inhomogeneity to the mathematical analysis by investigating random environments, thus allowing for a more realistic modelling of RNAp.

The main distinction between RNAp and a jump process however is the fact that the polymerase cannot pass the state where backtracking was initiated in the diffusive layer. This barrier adds glaring issues when trying to combine it with the Markov theory. Since the barrier is placed at the location where the polymerase entered a backtrack, the polymerase's past influences future motion and thus the Markov property is violated.
We will regard the barrier as an added component to the jump process formulation, and in chapter 5, we will investigate numerically the new phenomena this additional component introduces into the system.
Again, we also discussed that pausing of the polymerase may only occur at specific sites on the DNA (for example due to the TFs). We will also view these "pause sites" as an extra component to the simple jump process model. We will conduct both simulations with and without sites and compare the results between the two.

CHAPTER 3

INVARIANCE PRINCIPLES FOR LAYERED MARKOV PROCESSES

In this chapter, we conclude the mathematical analysis of single-particle layered Markov processes by deriving invariance principles in homogeneous and inhomogeneous environments. In the next chapter, we will follow up on this by conducting numerical simulations on multi-particle systems and comparing the numerical and analytical results.

To be able to derive the results on single particle systems, we need to expand further on the theory of Markov processes. For instance, we introduce martingales, which act as a standard tool in proving central limit theorems/invariance principles. After introducing this new theory, we apply it to investigate the limiting behaviour of particles in the layered systems of section 1.2.3. In particular, we prove that the particles, on average, drift off with a certain drift velocity and that the spread around this drift is diffusive in nature. The drift part of the limiting behaviour can be seen as a sort of continuous time generalisation of the law of large numbers, while the diffusive motion is akin to a central limit theorem (CLT), which is also known as an invariance principle in continuous time. In homogeneous environments, we find exact expressions for the drift velocity and diffusion coefficient. However, we will not provide these estimates in inhomogeneous media; these lie outside the scope of this thesis, and in a sense also outside of the scope of available mathematical theory. Diffusion coefficients in random environments are more suited to analysis in less general, more application-specific settings.

But first, as promised, we need some more advanced Markov theory to be able to establish all these limit theorems in homogeneous and inhomogeneous environments.

3.1. EXPANDING ON THE BASIC MARKOV THEORY

3.1.1. INTRODUCING NEW TYPES OF PROCESSES

In the context of finite state space, we introduced several notions to distinguish between different types of Markov processes. We will now classify several types of Feller processes as well. By classifying the processes, we allow for new category-specific consequences to arise. Therefore, we will also investigate what new properties each category brings and apply these results later in this chapter.

STATIONARY DISTRIBUTIONS

In finite state space, we introduced stationary distributions as distributions that remain unchanged in time. The defining equations were $\pi S_t = \pi$ and $\pi L = \vec{0}^T$, which can be found in definition 10 and theorem 7. We will now generalise this idea to an arbitrary countable state space S.

Definition 15. *Stationary measure:*
Let S_t be the transition operator of a Feller process with generator L on state space S.
We call a measure μ stationary if, for any bounded and measurable f:

$$\mathbb{E}_\mu(S_t f) = \mathbb{E}_\mu(f). \tag{3.1}$$

which can be viewed as the countable space version of $\pi S_t = \pi$.
Equivalently, μ is stationary if for any f in the domain of L we have that:

$$\mathbb{E}_\mu(Lf) = 0 \tag{3.2}$$

which, in turn relates to the finite space equation $\pi L = \vec{0}^T$.

ERGODIC PROCESSES

Ergodicity entails that the process will eventually explore every state in the system. We briefly touched on the notion in our discussion of theorem 8 and in the interchanging of sample and time averages in section 2.2. We will now define ergodicity more formally.

Definition 16. *Invariant set:*
With the same set-up as in definition 15, take a set of states $A \subset S$.
We call A invariant if $S_t \mathbb{1}_A = \mathbb{1}_A$, i.e. $X_0 \in A \implies X_t \in A$.

The invariant sets are closed regions in the state space. Once the Markov process enters such a set, it will remain there and can never exit the region.

Definition 17. *Ergodic measure:*
A stationary measure μ is called ergodic if all invariant sets have μ-measure one or zero.
Equivalently, all invariant bounded measurable functions are μ-almost surely constant.

Take an ergodic measure μ as initial condition and consider the region in S where all possible processes start; i.e. the set with measure 1. As μ is stationary, the process cannot leave this region. So, from the particle's perspective, this might as well be the entire space. Hence, ergodicity says that the only stationary sets are the empty set and the entire space/starting region. Consequently, there is no strict subset of space in which the process will stay forever, which corresponds with our intuitive idea of ergodicity.
With the complete definition of ergodicity at hand, we can finally present the full ergodic theorem, which builds further on theorem 8.

Theorem 18. *Let X_t be an irreducible Markov chain on finite state space S. By theorem 8 and the fact that S is finite, there exists a stationary distribution π.*
Firstly, note that π will be ergodic, since the only invariant set is S due to the irreducibility.
Secondly, we have that for any function $f : S \to \mathbb{R}$:

$$\frac{1}{T}\int_0^T f(X_t)\,dt \to \mathbb{E}_\pi(f) \text{ a.s.} \tag{3.3}$$

Corollary 19. *By taking $f(x) = \mathbb{1}_{x=i}$ in the previous theorem, we see that the mean time spent in state i converges to π_i. This also holds for the layer processes of equation 1.20. Indeed, we can write $X_t = (x_t, \sigma_t)$ where σ_t denotes the layer at time t and x_t denotes the position in the layer. Note that x_t and σ_t are independent and that (σ_t) itself describes a Markov process on finite state space. Therefore, we can conclude that the mean time spent in each layer is given by π_i with π being the stationary distribution over the layers.*

3.2. MARTINGALES

Most of the theory on central limit theorems relies on a concept known as martingales. In this section, we introduce the definition of martingales and show why they are especially useful in proving invariance principles by stating the martingale central limit theorem.

Definition 18. *Filtration:*
Let $(\Omega, \mathscr{A}, \mathbb{P})$ be a probability space and let $T = [0, \infty)$. Suppose that we have a collection of sub-sigma algebras of $\mathscr{A}$: $(\mathscr{F}_t)_{t \in T}$. If this collection satisfies $q \leq r \implies \mathscr{F}_q \subset \mathscr{F}_r$, we call the collection of sigma algebras a filtration. Any stochastic process X_t on $(\Omega, \mathscr{A}, \mathbb{P})$ has a natural filtration given by $\mathscr{F}_t = \sigma\{X_r : r \leq t\}$.

Filtrations entail the information that is available to the process at each moment in time and thus also the history of the process. Martingales are the probabilistic counterpart of a constant sequence: given the position at present, martingales have an expectation to stay in the same place. We elaborate this idea in the following definition.

Definition 19. *Martingale:*
Let M_t be a stochastic process on $(\Omega, \mathscr{A}, \mathbb{P})$.
Let $(\mathscr{F}_t)_{t \in [0,\infty)}$ be the associated natural filtration. We say that M_t is a martingale if:

$$\forall s \leq t : \mathbb{E}(M_t | \mathscr{F}_s) = M_s. \tag{3.4}$$

That is, we expect M_t to retain its current value in the future. Equivalently, we have:

$$\forall s \leq t : \mathbb{E}(M_t - M_s | \mathscr{F}_s) = 0.$$

Example 3. *Dynkin martingales (Landim et al., 2012).*
The most important martingales in our work are the ones naturally associated to Markov processes: the so-called "Dynkin martingales". Let L be the generator and let $f \in D(L)$, the domain of L. We have that the following process, M_t, is a martingale (Landim et al., 2012):

$$M_t = f(X_t) - f(X_0) - \int_0^t (Lf)(X_s)\, ds. \tag{3.5}$$

Additionally, we will see that the associated process M_t^2 is of significance in estimating the variance of the limiting process of X_t. For calculations on M_t^2, we turn to the concept known as the quadratic variation of a martingale.

Definition 20. *The quadratic variation $[M, M]_t$:*
Let M_t be a martingale. $[M, M]_t$ is defined as the unique increasing process such that:
1) $[M, M]_0 = M_0^2$.
2) $M_t^2 - [M, M]_t = \widetilde{M}_t$ where $\widetilde{M}_t$ is also a martingale.

Corollary 20. *By properties 1 and 2, we have that $M_t^2 - [M, M]_t$ is a martingale with expectation zero. It follows that $\mathbb{E}(M_t^2) = \mathbb{E}([M, M]_t)$ for any t.*

Theorem 21. *Quadratic variation for Dynkin martingales (Landim et al., 2012):*
If M_t is a Dynkin martingale defined as in equation 3.5, then we have:

$$[M, M]_t = \int_0^t (Lf^2)(X_s) - 2f(X_s)(Lf)(X_s)\, ds. \tag{3.6}$$

Additionally, if π is a stationary measure of the process X_t, then: $\mathbb{E}_\pi([M, M]_t) = 2t\langle f, -Lf \rangle_\pi$.

Theorem 22. *Martingale Central Limit Theorem (Landim et al., 2012):*
Let M_t be a martingale with stationary increments on $(\Omega, \mathscr{A}, \mathbb{P})$.
Suppose that $M_0 = 0$ and that for some $\sigma^2 < \infty$:

$$\mathbb{E}_{\mathbb{P}}\left(\left|\frac{1}{t}[M,M]_t - \sigma^2\right|\right) \to 0 \text{ as } t \to \infty.$$

Then, the distribution of $\frac{M_t}{\sqrt{t}}$ converges in probability to a Gaussian with mean zero and variance σ^2.

3

Proposition 23. *Let M_t be a Dynkin martingale for a process X_t with stationary and ergodic measure π. Then, the Martingale Central Limit theorem holds if $\mathbb{E}_{\mathbb{P}_\pi}(\frac{1}{t}[M,M]_t) < \infty$.*

Proof. The assumption $[M,M]_0 = M_0^2$ is clearly satisfied. The other assumptions such as convergence in $L_1(\mathbb{P}_\pi)$ and stationary increments are proven in (Landim et al., 2012) based on a general ergodic theorem and on the fact that X_t is stationary.
Furthermore, for Dynkin martingales, theorem 21 states that $\mathbb{E}_\pi(\frac{1}{t}\langle M,M\rangle_t) = 2\langle f, -Lf\rangle_\pi$.
Thus, equivalently, we could check that $\sigma^2 = 2\langle f, -Lf\rangle_\pi < \infty$.

In practice, we will often only verify that $\frac{1}{t}\langle M,M\rangle_t$ is bounded. Then, $\mathbb{E}_\pi(\frac{1}{t}\langle M,M\rangle_t)$ must automatically be finite too, as the expectation is taken over a probability measure. Of course, this last step only holds if the ergodic probability measure π exists in the first place. However, we shall work with irreducible processes, such that any stationary measure is automatically ergodic, as the only invariant set is the entire space. Generally, the stationary measure can be found quite easily by combining the known stationary measure of the switching process with symmetry considerations for the spatial process. □

3.3. LIMIT BEHAVIOUR IN HOMOGENEOUS ENVIRONMENT

For sequences of random variables, the reader should already be familiar with standard convergence theorems such as the law of large numbers (LLN) and the CLT.
The LLN and CLT take on the form $\frac{S_n}{n} \xrightarrow{n\to\infty} \mathbb{E}(S_1)$ and $\frac{S_n}{\sqrt{n}} \xrightarrow{n\to\infty} \mathcal{N}(\mathbb{E}[S_1], \mathbb{V}\text{ar}[S_1])$ respectively, where S_n is the sum of n i.i.d. random variables X_i.
These theorems apply naturally in the context of discrete random walks, where we take S_n to be the position at time n and X_i to be the increments. If we want to generalise these results to continuous time random walks and processes, we will have to change the way in which we define the sequences $\frac{S_n}{n}$ and $\frac{S_n}{\sqrt{n}}$.
We postulate that the scaling should be the same as in the discrete case. So, for the LLN we will use the scaling $\frac{x_t}{t}$ with $t \to \infty$ and x_t the position of the particle. For the CLT, we take the scaling $\frac{x_{kt}}{\sqrt{k}}$ as $k \to \infty$, which is equivalent to taking $\epsilon x_{\epsilon^{-2}t}$ with $\epsilon \to 0$ since

$$\lim_{\epsilon\to 0} \epsilon x_{\epsilon^{-2}t} = \lim_{\epsilon^{-2}\to\infty} \epsilon x_{\epsilon^{-2}t} = \lim_{k\to\infty} \frac{x_{kt}}{\sqrt{k}}$$

Our main objective will be to derive that, for the layered jump processes of generator 1.20, $\frac{x_t}{t} \to c$ and $\frac{x_{kt}}{\sqrt{k}} \to B(Dt)$, with c being the drift velocity and with $B(Dt)$ denoting Brownian motion with diffusion constant D. The former of these two is our continuous-time analogue to the LLN, while the latter is an invariance principle similar to the CLT.

3.3.1. WEAK LAW OF LARGE NUMBERS

Theorem 24. *Drift in general layer processes:*
Let X_t be an $n+1$-layered jump process, similar to equation 1.20. So, $X_t = (x_t, \sigma_t)$ walks on $\mathbb{Z} \times \{0, 1, \dots, n\}$. Each layer i has a diffusive walk with rate κ_i and an active component making jumps of size v_i with rate 1. Note that for a system with active components with arbitrary rates, we can always reduce it to rate 1 by rescaling time.
Additionally, we shall assume irreducibility over the layers, by which we mean that for any pair of layers i, j, we have that $p_{i,j}(t) > 0$ for some t onwards.
Let π be the stationary distribution over the layers, then the drift of the process is given by:

$$\frac{x_t}{t} \xrightarrow{\mathbb{P}} \sum_{i=0}^{n} \pi_i v_i \text{ as } t \to \infty \tag{3.7}$$

Proof. Firstly note that π exists by irreducibility over the layers.
Our approach will now be to show that $\frac{x_t}{t} \to \frac{1}{t}\sum_{i=0}^{n} T_i v_i$, where T_i is the time spent in layer i, i.e. $T_i = \int_0^t \mathbb{1}_{\sigma_s = i} ds$.
We take $f(x, \sigma) := x$ and its corresponding Dynkin martingale M_t (equation 3.5):

$$\frac{x_t}{t} = \frac{x_0}{t} + \frac{1}{t}\int_0^t (Lx)(x_s, \sigma_s) ds + \frac{M_t}{t}$$

All diffusive components and switching terms trivially cancel out in Lx. So, all that remains in Lx are the active terms $\mathbb{1}_{\sigma=i} f(x + v_i, i) - f(x, i) = \mathbb{1}_{\sigma=i} v_i$. We conclude that

$$\frac{x_t}{t} = \frac{x_0}{t} + \frac{1}{t}\int_0^t \sum_{i=1}^{n} \mathbb{1}_{\sigma_s = i} v_i ds + \frac{M_t}{t}.$$

Since the layer process is irreducible, we may use the ergodic theorem (equation 3.3) and we find $\frac{1}{t}\int_0^t \sum_{i=1}^{n} \mathbb{1}_{\sigma_s=i} v_i ds \to \sum \pi_i v_i$.
Therefore, all that remains to finish the proof is to get rid of the x_0 and M_t terms. Firstly, it is clear that $\frac{x_0}{t} \to 0$ since x_0 is a finite constant. The M_t term vanishes as a consequence of its linear quadratic variation and Chebyshev's inequality. Indeed, we have that:

$$\begin{aligned}
[M, M]_t &= \int_0^t (Lx^2)(x_s, \sigma_s) - 2x_s (Lx)(x_s, \sigma_s) ds \\
&= \int_0^t \sum_{i=0}^{n} \mathbb{1}_{\sigma_s = i} \left((x_s + v_i)^2 - x_s^2 + \kappa_i ((x_s + 1)^2 + (x_s - 1)^2 - 2x_s^2) \right) - \left(2x_s \sum_{i=0}^{n} v_i \right) ds \\
&= \int_0^t \sum_{i=0}^{n} \mathbb{1}_{\sigma_s = i} \left(v_i^2 + 2\kappa_i \right) ds \leq \left(\sum_{i=0}^{n} v_i^2 + 2\kappa_i \right) t.
\end{aligned}$$

Therefore,

$$0 \leq \mathbb{E}\left(\frac{M_t^2}{t^2}\right) = \mathbb{E}\left(\frac{[M, M]_t}{t^2}\right) \leq \frac{\sum_{i=0}^{n} v_i^2 + 2\kappa_i}{t} \implies \lim_{t \to \infty} \frac{\mathbb{E}(M_t^2)}{t^2} = 0.$$

Hence, by Chebyshev's inequality:

$$\mathbb{P}\left(\left| \frac{M_t}{t} \right| \geq \epsilon \right) \leq \frac{\mathbb{E}(M_t^2)}{t^2 \epsilon^2} \to 0.$$

From this last expression, it follows that the convergence of $\frac{x_t}{t}$ is in probability. □

3.3.2. INVARIANCE PRINCIPLES/CENTRAL LIMIT THEOREMS

Our goal is to prove an invariance principle for the continuous-time layered jump processes under our study. In the previous section 3.3.1, we found the drift of such processes. Therefore, we can reasonably expect that the scaling should be

$$\frac{x_t - c\cdot t}{\sqrt{t}}, \frac{x_{kt} - c\cdot(kt)}{\sqrt{k}} \text{ or equivalently } \epsilon\left(x_{\epsilon^{-2}t} - c\cdot(\epsilon^{-2}t)\right).$$

Here, we centered the process by subtracting the drift velocity c multiplied by the time.

THE STANDARD METHOD

The standard method of proving central limit theorems for these layered random walks is outlined by (Landim et al., 2012). We employ the following procedure:

1. Formulate the Dynkin martingale $x_t - x_0 - \int_0^t (Lx)(X_s)ds = M_t$. Since we need to center the process, it also makes sense to subtract ct from both sides, hence:
$$x_t - ct = x_0 + \int_0^t (Lx)(X_s) - c\,ds + M_t =: x_0 + \int_0^t V(X_s)ds + M_t.$$
2. $\frac{x_0}{\sqrt{t}}$ automatically vanishes due to x_0 being a finite constant.
3. Apply the Martingale CLT (theorem 22) to $\frac{M_t}{\sqrt{t}}$.
4. All that remains is to deal with the term $\frac{1}{\sqrt{t}}\int_0^t V(X_s)ds$. This term is referred to as an additive functional of X_s. A large framework of results that provide conditions for its convergence to Brownian motion has been established (Landim et al., 2012, Kipnis and Varadhan, 1986). Admittedly, these results are not well-suited to the case $c \neq 0$ as they mostly operate under the assumption that $V(X_s) = Lf$ for some $f \in D(L)$. Therefore, the cases $c \neq 0$ are in general problematic.
However, if Lx happens to depend only on the layer σ_s, but not on the position, we can introduce a second Dynkin martingale to resolve this term as well. This idea will be illustrated in the upcoming section.

CONVERGENCE TO BROWNIAN MOTION FOR LAYER PROCESSES

Lemma 25. *The drift additive functional lemma.*
Let (x_t, σ_t) be a jump process in $n+1$ layers with generator L. Let L_σ be the corresponding generator of the switching process between the layers. As always, assume that the layer process is irreducible and has stationary distribution π.
Then, the centered functional $\int_0^t \mathbb{1}_{\sigma_s=i} v_i - \sum_{i=1}^n v_i\pi_i\,ds$ satisfies a central limit theorem.

Proof. Note that the lemma is trivial for the case $n = 0$. Indeed, for $n = 0$, we see that $\mathbb{1}_{\sigma_s=0} = 1$ and $\pi_0 = 1$. Therefore, the centered functional vanishes.
For $n \geq 1$, the plan is to relate the integrand to another Dynkin martingale and to use its quadratic variation. If we can solve the equation $(L_\sigma h)(\sigma) = \mathbb{1}_{\sigma=i} v_i - \sum_{i=1}^n v_i\pi_i =: V(\sigma)$ for some function h, then we have that for some martingale $\widetilde{M}_t$:

$$\int_0^t \mathbb{1}_{\sigma_s} v_i - \sum_{i=1}^n v_i\pi_i\,ds = \int_0^t (Lh)(\sigma_s)ds = h(\sigma_t) - h(\sigma_0) - \widetilde{M}_t.$$

If such a function h exists, the quadratic variation of $\widetilde{M}_t$ is certainly going to be bounded, as h takes only finitely many values. Therefore, by the Martingale CLT, we can conclude that $\widetilde{M}_t$ satisfies a central limit type theorem. Additionally, the terms $h(\sigma_t) - h(\sigma_0)$ trivially fall out as h is bounded. Therefore, the centered functional will follow a central limit theorem, if the equation $(L_\sigma h)(\sigma) = V(\sigma)$ has a solution.
We can borrow some theory from linear algebra to show that $\exists h : (L_\sigma h)(\sigma) = V(\sigma)$. Note that this problem is equivalent to showing $V \in \text{Range}(L) = \text{Col}(L)$. A well known result from linear algebra is that the row space and the null space of a matrix are orthogonal (Fraleigh and Beauregard, 1995). Therefore, $\text{Col}(L) = \text{Row}(L^T) = \text{Null}(L^T)^\perp$. By theorems 7 and 8, we know that π is the only probability measure in $\text{Null}(L^T)$. Hence, $\text{Null}(L^T) = \{C\pi : C \in \mathbb{R}\}$. For any function g it is easy to see that $g - \mathbb{E}_\pi(g) \in \text{Null}(L^T)^\perp$, as $(g - \mathbb{E}_\pi(g)) \cdot \pi = \mathbb{E}_\pi(g) - \sum_{i=0}^n \pi_i \mathbb{E}_\pi(g) = 0$. Therefore, $V \in \text{Null}(L^T)^\perp = \text{Range}(L)$. □

Theorem 26. *Central limit theorem for layer processes:*
Let $X_t = (x_t, \sigma_t)$ be a layered jump process on $n+1$ layers. Again, we shall assume irreducibility over the layers. Hence there exists a stationary distribution π over the layers. The process has a drift velocity c given by theorem 24.
Then, the rescaled process $\frac{x_{kt} - ckt}{\sqrt{k}}$ converges to Brownian motion $B(Dt)$ in probability, for some diffusion constant D. The exact value of D will be extracted in the next section.

Proof. We will apply the standard method of proving CLTs. We have already computed the Dynkin martingale in theorem 24. Writing out the scaling fully yields:

$$\frac{x_{kt} - ckt}{\sqrt{k}} = \frac{x_0}{\sqrt{k}} + \frac{1}{\sqrt{k}} \int_0^{kt} \sum_{i=1}^n \mathbb{1}_{\sigma_s = i} v_i - c\, ds + \frac{M_{kt}}{\sqrt{k}}.$$

Now, we can move on to step 3. Both conditions of the martingale CLT are met for M_t. The stationary increments of M_t follow directly from the homogeneity of space and the stationary layer process. To establish L^1-convergence of the quadratic variation, we remark that by the ergodic theorem 18, we already have the following convergence:

$$\frac{1}{t}[M,M]_t = \frac{1}{t} \int_0^t \sum_{i=0}^n \mathbb{1}_{\sigma_s = i} (v_i^2 + 2\kappa_i)\, ds \to \sum_{i=0}^n \pi_i (v_i^2 + 2\kappa_i) = \sigma_1^2 \text{ a.s.}$$

By boundedness of $\frac{1}{t}[M,M]_t$ (see the proof of 24), we may apply the dominated convergence theorem to conclude $\mathbb{E}(\frac{1}{t}[M,M]_t) \to \sigma^2$ as well and hence the L^1-convergence follows from Scheffe's lemma. Therefore, by the martingale CLT:

$$\frac{M_{kt}}{\sqrt{kt}} \xrightarrow{\mathbb{P}} \mathcal{N}\left(0, \frac{\sigma_1^2}{2}\right) \implies \frac{M_{kt}}{\sqrt{k}} \xrightarrow{\mathbb{P}} \mathcal{N}\left(0, \frac{\sigma_1^2 t}{2}\right) = B(D_1 t).$$

Lastly, we can perform step 4. We can now employ the drift functional lemma 25 to resolve the additive functional. As a consequence of this lemma, $\int_0^t V(X_s)\, ds \to B(D_2 t)$ for another constant D_2. It is clear that processes giving rise to $B(D_1 t)$ and $B(D_2 t)$ are completely independent: the first is caused by the movements within each layer, the second by the movement between layers. Therefore, we can safely conclude that:

$$\frac{x_{kt} - ckt}{\sqrt{k}} \xrightarrow{\mathbb{P}} B(D_1 t) + B(D_2 t) \sim B((D_1 + D_2) t).$$ □

CALCULATING THE LIMITING DIFFUSION CONSTANTS

In this section, we will derive the actual value of the constant D for the limiting Brownian motion of 2- and 3-layer systems by solving the master equation 1.8. Chapter 1 introduced an approach based on matrix exponentials, but other solving techniques also exist. A particular method that comes to mind is based on Fourier-Laplace transforms (FLTs), which turn differential equations into polynomial ones. We follow this approach based on (van Ginkel et al., 2021) to solve master equations for layered jump processes. Let $\mu_t(x,\sigma)$ be the probability to be in (x,σ) at t and $\widehat{\mu}_t(q)$ be its Fourier transform (FT):

$$\widehat{\mu}_t(q,\sigma) := \mathscr{F}(\mu_t(x,\sigma)) = \sum_x e^{iqx}\mu_t(x,\sigma). \tag{3.8}$$

Master equation $\frac{d}{dt}\mu_t^T(x,\sigma) = L^T\mu_t^T(x,\sigma)$, becomes $\frac{d}{dt}\widehat{\mu}_t(q) = M(q)\widehat{\mu}_t(q)$ for some $M(q)$. The FT is convenient as it expresses all positional information in terms of phase shifts based on transition increments $y-x$. This trick works if the generator itself only depends on increment; that is, N_x and $\lambda_{x,y}$ both only depend on $y-x$ for $y\in N_x$. We will refer to this property as space homogeneity of the generator. In section 3.4, we will also cover an inhomogeneous model that cannot be treated with the FT.
To eliminate the differential equation, we find the Laplace transform of the expression:

$$\begin{aligned}\overline{\mu}(q,z) &:= \mathscr{L}(\widehat{\mu}_t(q)) = \int_0^\infty \widehat{\mu}_t(q)e^{-zt}dt = \left[\frac{-1}{z}I\widehat{\mu}_t(q)e^{-zt}\right]_0^\infty + \frac{1}{z}\int_0^\infty e^{-zt}\frac{d}{dt}\widehat{\mu}_t(q)dt \\ &= \frac{1}{z}I\widehat{\mu}_0(q) + \frac{1}{z}\int_0^\infty M(q)\widehat{\mu}_t(q)e^{-zt}dt = \bar{\mu}(q,z) \implies \bar{\mu}(q,z) = \big(zI - M(q)\big)^{-1}\widehat{\mu}_0(q).\end{aligned} \tag{3.9}$$

We also compute the Laplace transform of the characteristic function (CF) of the process:

$$\begin{aligned}S(q,z) :&= \int_0^\infty \mathbb{E}(e^{iqX_t})e^{-zt}dt = \int_0^\infty \sum_\sigma\sum_x e^{iqx}\mu_t(x,\sigma)e^{-zt}dt = \sum_\sigma\int_0^\infty\sum_x e^{iqx}\mu_t(x,\sigma)e^{-zt}dt \\ &= \sum_\sigma \bar{\mu}(q,z,\sigma) = (1,\ldots,1)\bar{\mu}(q,z) = (1,\ldots,1)\big(zI - M(q)\big)^{-1}\widehat{\mu}_0(q)\end{aligned} \tag{3.10}$$

Using all these transforms, we can finally start computing actual diffusion coefficients. In the lemmas below, we show that the FLT of the limit of X_t is the limit of the transforms of X_t. Note that we know that the limit of X_t is Brownian motion $B(Dt)$ with FLT given by equation 2.4. However, since it is easier to compute the limit of the FLTs, we can find the exact value of D from the sequence of FLTs instead of the process itself.

Theorem 27. *Lévy's continuity theorem (Jacod and Protter, 2004):*
Let $(x_n)_{n\in\mathbb{N}}$ be a sequence of real-valued random variables with CFs/FTs: ψ_{x_n}, defined by $\psi_{x_n} = \mathbb{E}(e^{iux_n})$ for $u\in\mathbb{R}$. Suppose that $x_n \xrightarrow{d} x$ for some random variable x.
Then, $\forall u : \psi_{X_n}(u) \to \phi(u)$ as $n\to\infty$, where ϕ is the CF of x.

Lemma 28. *The Fourier-Laplace Convergence Lemma:*
Let x_t be a positional process, $(\epsilon_n)_{n\in\mathbb{N}}$ be any decreasing sequence converging to 0 and label $x_n(t) = \epsilon_n x_{\epsilon_n^{-2}t}$. Then, the diffusion constant D can be found from:

$$\frac{(x_{kt}-ckt)}{\sqrt{k}} \xrightarrow[k\to\infty]{\mathbb{P}} B(Dt) \implies \frac{(x_{kt}-ckt)}{\sqrt{k}} \xrightarrow[k\to\infty]{d} B(Dt) \implies \mathscr{L}(\psi_{x_n(t)}) \xrightarrow[t\to\infty]{} \frac{1}{z+Ds^2}.$$

Proof. Convergence in probability is stronger than convergence in distribution, so we may use Lévy's continuity theorem 27. Hence, $\psi_{x_n(t)} \to \phi_{Dt}$ where $\psi_{x_n(t)}$ is the CF of $x_n(t)$ and $\phi_{Dt}(s) = e^{-Dts^2}$ is the CF of Brownian motion with constant D by equation 2.3. By definition, we know that the modulus of any CF for a real-valued stochastic process is at most 1. Consequently, we get a sequence of Laplace transforms $(\psi_{X_n(t)}(s)e^{-zt})$, which satisfies two important properties. Firstly, $\psi_{X(\epsilon_n,t)}e^{-zt} \to \phi_D e^{-zt}$ almost everywhere. Secondly, $|\psi_{X_n(t)}e^{-zt}| \leq e^{-zt}$, which is an integrable function on our time domain $[0,\infty)$. Therefore, the sequence of Laplace transforms satisfies all the conditions of the Dominated Convergence Theorem for complex-valued functions. We conclude:

$$\lim_{n\to\infty}\int_0^\infty \psi_{X_n(t)}e^{-zt}dt = \int_0^\infty \phi_{Dt}e^{-zt}dt = \frac{1}{z+Ds^2} \qquad \square$$

We can now apply this result to the transform $S(q,z)$, defined in equation 3.10. Hence, we see that we can compute the diffusion constant D by finding the limit of the transforms instead of the transform of the limit of the process itself. Of course, the main difficulty in using $S(q,z)$ lies in inverting the matrix. Nevertheless, it can be used to calculate the diffusion constant for the 2- and 3-layer systems of equation 1.20. But first, we begin with a 1-layer system, to give a simple example of the Fourier-Laplace method in action.

Theorem 29. ***The main homogeneous environment 1-layer invariance principle:*** *Let x_t be a 1-layered jump process with initial condition $x_0 = 0$, as defined by the generator: $Lf(x) = \kappa(f(x+1)+f(x-1)-2f(x))+f(x+v_0)-f(x)$. Then,*

$$\epsilon\left(x_{\epsilon^{-2}t} - c\epsilon^{-2}t\right) \xrightarrow{d} B(Dt) \text{ as } \epsilon \to 0 \text{ with } D = \kappa_0 + \frac{v_0^2}{2}. \tag{3.11}$$

Proof. The master equation results in:

$$\frac{d}{dt}\widehat{\mu}_t(q) = \kappa_0(e^{iq}+e^{-iq}-2)+e^{iv_0q}-1 \implies S(q,z) = \left(z+(2-2\cos(q))+(1-e^{iv_0q})\right)^{-1}.$$

$S(q,z)$ is the FLT of the process x_t without any rescaling or centering. The actual process under consideration is: $\epsilon(x_{\epsilon^{-2}t} - \epsilon^{-2}ct)$ and has its own Fourier-Laplace transform:

$$\begin{aligned}\hat{S}(q,z) &= \int_0^\infty \mathbb{E}\left(e^{iq\epsilon(x_{\epsilon^{-2}t}-c\epsilon^{-2}t)}\right)e^{-zt}dt = \int_0^\infty \mathbb{E}\left(e^{i(q\epsilon)x_{\epsilon^{-2}t}}\right)e^{-t(z+iqc\epsilon^{-1})}dt \\ &= \epsilon^2\int_0^\infty \mathbb{E}\left(e^{i(q\epsilon)x_{\epsilon^{-2}t}}\right)e^{-\epsilon^{-2}t(\epsilon^2 z+iqc\epsilon)}d(\epsilon^{-2}t) = \epsilon^2 S\left(q\epsilon,\epsilon^2 z+iqc\epsilon\right).\end{aligned}$$

Hence, $\hat{S}(q,z)$ can be obtained by taking $S(q,z)$ under the transformation $z \to z+iqc$ and afterwards substituting $q \to q\epsilon, z \to \epsilon^2 z$. We will be brief with these calculations, as many of the steps are repeated in more detail for the 2- and 3-layer systems.

$$\hat{S}(q,z) = \epsilon^2\left(\epsilon^2 z + \left(\frac{v_0^2q^2}{2}\epsilon^2 + iq\epsilon v_0 - iq\epsilon v_0 + \kappa_0 q^2\epsilon^2 + O(\epsilon^3)\right)\right)^{-1} \to \left(z+\left(\kappa_0+\frac{v_0^2}{2}q^2\right)\right)^{-1}.$$

In lemma 28 we showed that $\hat{S}_n(q,z) \xrightarrow{n\to\infty} (z+q^2D)^{-1}$ for discrete sequences $\hat{S}_n$, with D the diffusion constant of the limiting Brownian motion. The continuous limit $\epsilon \downarrow 0$ given above, shows in particular that any sequence will also converge to this limit. Therefore, the diffusion constant of the 1-layered jump process can be read off: $D = \kappa_0 + \frac{v_0^2}{2}$. $\square$

Theorem 30. ***The main homogeneous environment 2-layer central limit theorem:*** *Let X_t be a 2-layered jump process as defined by the generator 1.20. So, both layers have diffusive rates κ_i and active jumps of size v_i and rate 1. We assume that $\lambda_{1,2}, \lambda_{2,1} > 0$ to guarantee irreducibility. Let $Y_t = X_t - ct$ be the centered process with:*

$$c = \frac{\lambda_{1,0} v_0 - \lambda_{0,1} v_1}{\lambda_{1,0} + \lambda_{0,1}}. \tag{3.12}$$

Then, for any initial distribution $\mu_0(x,\sigma) = \alpha\delta_{(0,0)}(x,\sigma) + (1-\alpha)\delta_{(0,1)}(x,\sigma)$ with $\alpha \in [0,1]$:

$$\epsilon Y_{\epsilon^{-2}t} \xrightarrow{d} B(Dt) \text{ as } \epsilon \to 0 \tag{3.13}$$

Here, $B(Dt)$ denotes Brownian motion with diffusion constant

$$D = \frac{\lambda_{1,0}\lambda_{0,1}(v_0+v_1)^2}{(\lambda_{1,0}+\lambda_{0,1})^3} + \frac{(\kappa_0 + \frac{v_0^2}{2})\lambda_{1,0} + (\kappa_1 + \frac{v_1^2}{2})\lambda_{0,1}}{\lambda_{1,0}+\lambda_{0,1}}. \tag{3.14}$$

Proof. Note that, since we have to transpose L, the transformed master equation states:

$$\frac{d}{dt}\widehat{\mu}_t(q,\sigma) = \begin{bmatrix} \kappa_0(2\cos(q)-2) - \lambda_{0,1} + e^{iv_0 q} - 1 & \lambda_{1,0} \\ \lambda_{0,1} & \kappa_1(2\cos(q)-2) + e^{-iv_1 q} - 1 - \lambda_{1,0} \end{bmatrix} \widehat{\mu}_t(q,\sigma) \tag{3.15}$$

The cosine term that arises in the diffusive component is a typical result for Fourier transform of equations with nearest-neighbour interactions. For instance, the term also appears in the context of atomic chain models in solid state physics (Simon, 2019). In these chains, atom bonds are modelled as springs between nearest neighbours.
On the off-diagonals, we recognise the rate of transitions between the two states. In physics, this phenomenon is also apparent in multiple different fields. For example, in quantum perturbation theory, an overlap matrix W is used to compute perturbed quantum states (Griffiths and Schroeter, 2018). In this matrix, the off-diagonals provide a measure for the overlap between different states which is similar to the off-diagonals measuring transitions between states in our example. Another striking similarity is again in the field of solid state physics, where the off-diagonals of energy-transition matrices represent the "hopping" between different atomic orbitals (Simon, 2019).
Now, to continue the calculation we need to compute $(zI - M(q))^{-1}$:

$$M(q) = \begin{bmatrix} \kappa_0(2\cos(q)-2) - \lambda_{0,1} + e^{iv_0 q} - 1 & \lambda_{1,0} \\ \lambda_{0,1} & \kappa_1(2\cos(q)-2) + e^{-iv_1 q} - 1 - \lambda_{1,0} \end{bmatrix}$$

$$\big(zI - M(q)\big)^{-1} =: \begin{bmatrix} z+\varphi_0+\lambda_{0,1} & -\lambda_{1,0} \\ -\lambda_{0,1} & z+\varphi_1+\lambda_{1,0} \end{bmatrix}^{-1} =$$

$$\frac{1}{z^2 + z(\varphi_0+\varphi_1+\lambda_{0,1}+\lambda_{1,0}) + \varphi_0\varphi_1 + \varphi_0\lambda_{1,0} + \varphi_1\lambda_{0,1}} \cdot \begin{bmatrix} z+\varphi_1+\lambda_{1,0} & \lambda_{1,0} \\ \lambda_{0,1} & \varphi_0+z+\lambda_{1,0} \end{bmatrix}.$$

Recall that $S(q,z) = (1,1)\big(zI - M(q)\big)^{-1}\widehat{\mu}_0(q)$. So, we still need to choose an appropriate initial condition $\mu_0(x,\sigma)$. By theorems 8 and 18 the limit of the process cannot depend on the layer distributions $\sum_x \mu_0(x,0)$ and $\sum_x \mu_0(x,1)$ as the distributions will converge to the stationary distribution anyways.

Consequently, any initial distribution of the form $X_0(x,\sigma) = \alpha\delta_{(0,0)}(1-\alpha)(x,\sigma)+\delta_{(0,1)}(x,\sigma)$ must have the same limit.
The limit process for a more general initial distribution such as:

$$X_0(x,\sigma) = \sum_{(x',\sigma')\in F} \delta_{x',\sigma'}(x,\sigma)\cdot\alpha_{x',\sigma'}.$$

for some finite F can be found by using symmetry and superposition arguments.
Indeed, recall that the process is space-homogeneous. Therefore, if the process has a limit when starting with $\mu_0(x,\sigma) = \delta_{(0,0)}$, any process starting from $\mu_0(x,\sigma) = \delta_{(x_0,0)}(x,\sigma)$ must have the same limit shifted by x_0, as it is the exact same process translated to a new coordinate system. Again, $\delta_{(x_0,1)}(x,\sigma)$ and $\delta_{(x_0,0)}(x,\sigma)$ or any linear combination of the two must give the exact same limit. Therefore, the total limiting process will be a weighted sum of $|F|$ shifted, independent identically distributed Brownian motions.
We therefore might as well just start with the entire density concentrated at one point $(x,\sigma) = (0,0)$ to compute the limit. So, we take $\mu_0(x,\sigma) = \delta_{(0,0)}(x,\sigma)$ which has the Fourier transform $\mathbb{1}_{\sigma=0} = (1,0)^T$. Consequently,

$$\begin{aligned} S(q,z) &= (1,1)\left(zI - M(q)\right)^{-1}(1,0)^T \\ &= \frac{z+\varphi_1+\lambda_{1,0}+\lambda_{0,1}}{z^2+z(\varphi_0+\varphi_1+\lambda_{0,1}+\lambda_{1,0})+\varphi_0\varphi_1+\varphi_0\lambda_{1,0}+\varphi_1\lambda_{0,1}} \\ &= \left(z+\frac{z\phi_0+\phi_0\phi_1+\phi_0\lambda_{1,0}+\phi_1\lambda_{0,1}}{z+\phi_1+\lambda_{1,0}+\lambda_{0,1}}\right)^{-1}, \end{aligned} \tag{3.16}$$

$$\begin{aligned} \hat{S}(q,z) &= \left(\epsilon^2 S(q,z)\big|_{z\to z+iqc}\right)\Big|_{z\to\epsilon^2 z,\, q\to\epsilon q} \\ &= \epsilon^2\left(z+iqc+\frac{(z+iqc)\phi_0+\phi_0\phi_1+\phi_0\lambda_{1,0}+\phi_1\lambda_{0,1}}{(z+iqc)+\phi_1+\lambda_{1,0}+\lambda_{0,1}}\right)^{-1}\Big|_{z\to\epsilon^2 z,\, q\to\epsilon q} \\ &= \epsilon^2\left(z+\frac{z\phi_0+iqc(\phi_0+\phi_1+\lambda_{1,0}+\lambda_{0,1}+z)-q^2c^2+\phi_0\phi_1+\phi_0\lambda_{1,0}+\phi_1\lambda_{0,1}}{(z+iqc)+\phi_1+\lambda_{1,0}+\lambda_{0,1}}\right)^{-1}\Big|_{z\to\epsilon^2 z,\, q\to\epsilon q} \end{aligned}$$

To compute the limit, we need to collect terms of the same order of ϵ. In particular, we look at the numerator:

$$z\phi_0 + iqc(\phi_0+\phi_1+\lambda_{1,0}+\lambda_{0,1}+z) - q^2c^2+\phi_0\phi_1+\phi_0\lambda_{1,0}+\phi_1\lambda_{0,1}.$$

Any term with q^n will be of order ϵ^n and any term with z^m will be of order ϵ^{2m}. Also note that the numerator will be multiplied by ϵ^{-2}, so any higher order terms will vanish when taking the limit $\epsilon\to 0$.
By Taylor series:

$$\phi_0 = \kappa_0(2-2\cos(q)) + 1 - e^{iv_0 q} = \kappa_0 q^2 + \frac{v_0^2 q^2}{2} - iv_0 q + O(q^3).$$

$$\phi_1 = \kappa_1(2-2\cos(q)) + 1 - e^{-iv_1 q} = \kappa_1 q^2 + \frac{v_1^2 q^2}{2} + iv_1 q + O(q^3).$$

Collecting terms of order ϵ we get:

$$iqc(\lambda_{1,0}+\lambda_{0,1})iv_0q\lambda_{1,0}+iv_1q\lambda_{0,1}.$$

This term needs to vanish, otherwise the expression will certainly not converge as it would be of the form $(z+K\epsilon^{-1})^{-1}$ for some nonzero constant K. Therefore, we demand:

$$\begin{aligned} &iqc(\lambda_{1,0}+\lambda_{0,1})-iv_0q\lambda_{1,0}+iv_1q\lambda_{0,1}=0 \iff \\ &c=\frac{v_0\lambda_{1,0}-v_1\lambda_{0,1}}{\lambda_{0,1}+\lambda_{1,0}} \end{aligned}$$

which is the given expression for the drift of the process! Of course, we already knew that c should equal this value by theorem 26 and equation 1.17. It is however reassuring that the calculation yields the exact same result.

Now, we collect the terms of order ϵ^2, giving:

$$iqc\cdot(-iv_0q+iv_1q)-q^2c^2+v_0v_1q^2+(\kappa_0q^2+\frac{v_0^2q^2}{2})\lambda_{1,0}+(\kappa_1q^2+\frac{v_1^2q^2}{2})\lambda_{0,1}$$

Note that we can already recognise the terms $(\kappa_iq^2+\frac{v_i^2q^2}{2})\pi_i$ from the proof of theorem 26 and conclude that these are a result of the Dynkin martingale. The other terms in the diffusion constant are caused by the additive functional.

We conclude that:

$$\begin{aligned} \hat{S}(q,z)&=\epsilon^2\left(z+\frac{z\phi_0+iqc\left(\phi_0+\phi_1+\lambda_{1,0}+\lambda_{0,1}+z\right)-q^2c^2+\phi_0\phi_1+\phi_0\lambda_{1,0}+\phi_1\lambda_{0,1}}{\left(z+iqc\right)+\phi_1+\lambda_{1,0}+\lambda_{0,1}}\right)^{-1}\Bigg|_{z\to\epsilon^2z}^{q\to\epsilon q} \\ &=\left(z+\frac{iqc(-iv_0q+iv_1q)-q^2c^2+v_0v_1q^2+(\kappa_0q^2+\frac{v_0^2q^2}{2})\lambda_{1,0}+(\kappa_1q^2+\frac{v_1^2q^2}{2})\lambda_{0,1}+O(\epsilon)}{O(\epsilon)+\lambda_{1,0}+\lambda_{0,1}}\right)^{-1} \\ &\to\left(z+\frac{q^2c(v_0-v_1-c)+v_0v_1q^2+q^2\left((\kappa_0+\frac{v_0^2}{2})\lambda_{1,0}+(\kappa_1+\frac{v_1^2}{2})\lambda_{0,1}\right)}{\lambda_{1,0}+\lambda_{0,1}}\right)^{-1}\quad\text{as }\epsilon\to0 \\ &=\left(z+\frac{q^2\left(\frac{\lambda_{1,0}\lambda_{0,1}(v_0^2+v_1^2)-v_0v_1(\lambda_{0,1}^2+\lambda_{1,0}^2)}{(\lambda_{1,0}+\lambda_{0,1})^2}+v_0v_1\right)+q^2\left((\kappa_0+\frac{v_0^2}{2})\lambda_{1,0}+(\kappa_1+\frac{v_1^2}{2})\lambda_{0,1}\right)}{\lambda_{1,0}+\lambda_{0,1}}\right)^{-1} \\ &=\left(z+q^2\left(\frac{\lambda_{1,0}\lambda_{0,1}(v_0+v_1)^2}{(\lambda_{1,0}+\lambda_{0,1})^3}+\frac{(\kappa_0+\frac{v_0^2}{2})\lambda_{1,0}+(\kappa_1+\frac{v_1^2}{2})\lambda_{0,1}}{\lambda_{1,0}+\lambda_{0,1}}\right)\right)^{-1}=:\left(z+q^2D\right)^{-1}. \end{aligned}$$

□

Theorem 31. ***The main homogeneous environment 3-layer central limit theorem:*** *We can repeat the proof and the result for the general three layer process as well; Let X_t be a 3-layered jump process as defined by the generator 1.20 where we shall assume irreducibility over the layers. We shall assume that layer 1 is a pause layer, while layers 0, 2 both have diffusive rate κ_i and active jumps of size v_i with rate 1. Let $Y_t = X_t - ct$ be the centered process with*

$$c = K\left(v_0(\lambda_{1,0}\lambda_{2,1} + \lambda_{1,0}\lambda_{2,0} + \lambda_{1,2}\lambda_{2,0}) - v_2(\lambda_{0,1}\lambda_{1,2} + \lambda_{0,2}\lambda_{1,0} + \lambda_{0,2}\lambda_{1,2})\right). \tag{3.17}$$

Then,

$$\epsilon Y_{\epsilon^{-2}t} \xrightarrow{d} B(Dt) \text{ as } \epsilon \to 0 \tag{3.18}$$

for any initial distribution $\mu_0(x,\sigma) = \alpha\delta_{(0,0)}(x,\sigma) + (1-\alpha-\beta)\delta_{(0,1)}(x,\sigma) + \beta\delta_{(0,2)}(x,\sigma)$ for some $\alpha, \beta \geq 0 : \alpha + \beta \leq 1$. The diffusion coefficient is given by:

$$\begin{aligned} D = K\Big(&cv_0(\lambda_{1,0} + \lambda_{2,0} + \lambda_{2,1} + \lambda_{1,2}) - cv_2(\lambda_{0,1} + \lambda_{0,2} + \lambda_{1,0} + \lambda_{1,2}) + v_0v_2(\lambda_{1,0} + \lambda_{1,2}) \\ &- c^2(\lambda_{0,1} + \lambda_{0,2} + \lambda_{1,0} + \lambda_{1,2} + \lambda_{2,1} + \lambda_{2,0})\Big) + (\kappa_0 + \frac{v_0^2}{2})\pi_0 + (\kappa_2 + \frac{v_2^2}{2})\pi_2 \end{aligned} \tag{3.19}$$

where K is the normalization constant

$$K = \frac{1}{\lambda_{0,1}(\lambda_{2,0} + \lambda_{2,1} + \lambda_{1,2}) + \lambda_{1,2}(\lambda_{2,0} + \lambda_{0,2}) + \lambda_{1,0}(\lambda_{2,0} + \lambda_{2,1} + \lambda_{0,2}) + \lambda_{2,1}\lambda_{0,2}}.$$

Proof.

$$M(q) = \begin{bmatrix} -\phi_0 - \lambda_{0,1} - \lambda_{0,2} & \lambda_{1,0} & \lambda_{2,0} \\ \lambda_{0,1} & -\lambda_{1,0} - \lambda_{1,2} & \lambda_{2,1} \\ \lambda_{0,2} & \lambda_{1,2} & -\phi_1 - \lambda_{2,1} - \lambda_{2,0} \end{bmatrix} \tag{3.20}$$

By similar reasoning as in theorem 30 we can restrict ourselves to $\mu_0 = \delta_{(0,0)}$. Hence, only the first column of the inverse of $(zI - M(q))$ is needed. The inverse of a 3×3 matrix is:

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}^{-1} = \frac{1}{a(ei - fh) - b(di - fg) + c(dh - eg)} \begin{bmatrix} ei - fh & \dots & \dots \\ fg - di & \dots & \dots \\ dh - eg & \dots & \dots \end{bmatrix} \tag{3.21}$$

if it exists. In our case, this results in a numerator $\mathcal{N}$ and a denominator $\mathcal{D}$:

$$\begin{aligned} &\Big[(z + \lambda_{1,0} + \lambda_{1,2})(z + \phi_1 + \lambda_{2,1} + \lambda_{2,0}) - \lambda_{2,1}\lambda_{1,2}\Big] + \Big[\lambda_{2,1}\lambda_{0,2} + \lambda_{0,1}(z + \phi_1 + \lambda_{2,1} + \lambda_{2,0})\Big] \\ &+ \Big[\lambda_{0,1}\lambda_{1,2} + \lambda_{0,2}(z + \lambda_{1,0} + \lambda_{1,2})\Big] =: [\mathcal{N}_1] + [\mathcal{N}_2] + [\mathcal{N}_3] =: \mathcal{N} \end{aligned}$$

$$\begin{aligned} &(\lambda_{0,1} + \lambda_{0,2})\left((z + \lambda_{1,0} + \lambda_{1,2})(z + \phi_1 + \lambda_{2,1} + \lambda_{2,0}) - \lambda_{2,1}\lambda_{1,2}\right) - \lambda_{2,0}\left(\lambda_{0,1}\lambda_{1,2} + \lambda_{0,2}(z + \lambda_{1,0} + \lambda_{1,2})\right) \\ &+ \lambda_{1,0}\left(-\lambda_{2,1}\lambda_{0,2} - \lambda_{0,1}(z + \phi_1 + \lambda_{2,1} + \lambda_{2,0})\right) =: \mathcal{D} \end{aligned}$$

Therefore, the resulting expression is equivalent to

$$S(q,z) = \left(z + \frac{(\phi_0 + \lambda_{0,1} + \lambda_{0,2})\mathcal{N}_1 + (-\lambda_{1,0} - z)\mathcal{N}_2 - (\lambda_{2,0} + z)\mathcal{N}_3}{\mathcal{N}}\right)^{-1}. \tag{3.22}$$

After the usual coordinate transformation $z \to z + iqc$, we get:

$$\left(z + \frac{(\phi_0 + \lambda_{0,1} + \lambda_{0,2} + iqc)\widetilde{\mathcal{N}_1} + (-\lambda_{1,0} - z - iqc + iqc)\widetilde{\mathcal{N}_2} - (\lambda_{2,0} + z + iqc - iqc)\widetilde{\mathcal{N}_3}}{\mathcal{N}}\right)^{-1}.$$

The notation $\widetilde{\mathcal{N}_i}$ indicates that these terms are equal to the original $\mathcal{N}_i$ but with z substituted by $z + iqc$. Once again, we can collect terms in the new numerator:

3

$$\begin{aligned}
\widetilde{\mathcal{N}_1} &= (\lambda_{1,0} + \lambda_{1,2})(\lambda_{2,1} + \lambda_{2,0}) - \lambda_{2,1}\lambda_{1,2} \\
&\quad + iv_2 q(\lambda_{1,0} + \lambda_{1,2}) + iqc(\lambda_{1,0} + \lambda_{1,2} + \lambda_{2,1} + \lambda_{2,0}) \\
&\quad + z(\lambda_{1,0} + \lambda_{1,2} + \lambda_{2,1} + \lambda_{2,0}) + q^2\left(\kappa_2 + \frac{v_2^2}{2}\right)(\lambda_{1,0} + \lambda_{1,2}) - q^2(c^2 + cv_2) + \text{h.o.t.} \\
\widetilde{\mathcal{N}_2} &= \left[\lambda_{2,1}\lambda_{0,2} + \lambda_{0,1}(\lambda_{2,1} + \lambda_{2,0})\right] + \left[iv_2 q\lambda_{0,1} + iqc\lambda_{0,1}\right] + \left[z\lambda_{0,1} + q^2\left(\kappa_2 + \frac{v_2^2}{2}\right)\lambda_{0,1} + \text{h.o.t.}\right] \\
\widetilde{\mathcal{N}_3} &= \left[\lambda_{0,1}\lambda_{1,2} + \lambda_{0,2}(\lambda_{1,0} + \lambda_{1,2})\right] \qquad + \qquad \left[iqc\lambda_{0,2} + z\lambda_{0,2} + \text{h.o.t.}\right].
\end{aligned}$$

Therefore, the order $O(1)$ terms in the numerator amount to:

$$\begin{aligned}
&(\lambda_{0,1} + \lambda_{0,2})((\lambda_{1,0} + \lambda_{1,2})(\lambda_{2,1} + \lambda_{2,0}) - \lambda_{2,1}\lambda_{1,2}) - \lambda_{1,0}(\lambda_{2,1}\lambda_{0,2} + \lambda_{0,1}(\lambda_{2,1} + \lambda_{2,0})) \\
&- \lambda_{2,0}(\lambda_{0,1}\lambda_{1,2} + \lambda_{0,2}(\lambda_{1,0} + \lambda_{1,2})) = 0
\end{aligned}$$

The order $O(\epsilon)$ terms in the numerator add up to:

$$\begin{aligned}
&iq\Big(v_2(\lambda_{0,1} + \lambda_{0,2})(\lambda_{1,0} + \lambda_{1,2}) - v_2\lambda_{1,0}\lambda_{0,1} + c((\lambda_{1,0} + \lambda_{1,2})(\lambda_{2,1} + \lambda_{2,0}) - \lambda_{2,1}\lambda_{1,2} + \\
&(\lambda_{1,0} + \lambda_{1,2} + \lambda_{2,1} + \lambda_{2,0})(\lambda_{0,1} + \lambda_{0,2}) - \lambda_{1,0}\lambda_{0,1} - \lambda_{0,2}\lambda_{2,0}) - v_0((\lambda_{1,0} + \lambda_{1,2})(\lambda_{2,1} + \lambda_{2,0}) - \lambda_{2,1}\lambda_{1,2})\Big)
\end{aligned}$$

Once again, for convergence we require that this term is equal to zero. Consequently,

$$c = \frac{v_0(\lambda_{1,0}\lambda_{2,1} + \lambda_{1,0}\lambda_{2,0} + \lambda_{1,2}\lambda_{2,0}) - v_2(\lambda_{0,1}\lambda_{1,2} + \lambda_{0,2}\lambda_{1,0} + \lambda_{0,2}\lambda_{1,2})}{\lambda_{1,0}\lambda_{2,1} + \lambda_{1,0}\lambda_{2,0} + \lambda_{1,2}\lambda_{2,0} + \lambda_{1,0}\lambda_{0,2} + \lambda_{1,2}\lambda_{0,1} + \lambda_{1,2}\lambda_{0,2} + \lambda_{0,1}\lambda_{2,1} + \lambda_{0,2}\lambda_{2,1} + \lambda_{2,0}\lambda_{0,1}}. \tag{3.23}$$

Again, we already knew this by theorem 26 combined with equation 1.18.
Lastly, we collect the terms of order $O(\epsilon^2)$:

$$\begin{aligned}
&(\kappa_0 q^2 + \frac{v_0^2 q^2}{2})(\lambda_{1,0}\lambda_{2,1} + \lambda_{1,0}\lambda_{2,0} + \lambda_{1,2}\lambda_{2,0}) + (\kappa_2 q^2 + \frac{v_2^2 q^2}{2})(\lambda_{1,0}\lambda_{0,2} + \lambda_{1,2}\lambda_{0,1} + \lambda_{1,2}\lambda_{0,2}) \\
&+ (\lambda_{0,1} + \lambda_{0,2})(\lambda_{1,0} + \lambda_{1,2} + \lambda_{2,1} + \lambda_{2,0})z \\
&- q^2c^2(\lambda_{0,1} + \lambda_{0,2} + \lambda_{1,0} + \lambda_{1,2} + \lambda_{2,1} + \lambda_{2,0}) \\
&- q^2cv_2(\lambda_{0,1} + \lambda_{0,2} + \lambda_{1,0} + \lambda_{1,2}) + q^2cv_0(\lambda_{1,0} + \lambda_{2,0} + \lambda_{2,1} + \lambda_{1,2}) + v_2v_0q^2(\lambda_{1,0} + \lambda_{1,2}) \\
&- z(\lambda_{1,0}\lambda_{0,1} + \lambda_{2,1}\lambda_{0,2} + \lambda_{0,1}(\lambda_{2,1} + \lambda_{2,0}) + \lambda_{0,2}\lambda_{2,0} + \lambda_{0,1}\lambda_{1,2} + \lambda_{0,2}(\lambda_{1,0} + \lambda_{1,2})).
\end{aligned}$$

From the order $O(\epsilon^2)$ terms, it follows that:

$$D = \Big(c v_0(\lambda_{1,0} + \lambda_{2,0} + \lambda_{2,1} + \lambda_{1,2}) - c v_2(\lambda_{0,1} + \lambda_{0,2} + \lambda_{1,0} + \lambda_{1,2}) + v_0 v_2(\lambda_{1,0} + \lambda_{1,2}) \\ - c^2(\lambda_{0,1} + \lambda_{0,2} + \lambda_{1,0} + \lambda_{1,2} + \lambda_{2,1} + \lambda_{2,0}) \Big) K + \left(\kappa_0 + \frac{v_0^2}{2} \right) \pi_0 + \left(\kappa_2 + \frac{v_2^2}{2} \right) \pi_2$$

with K the same normalization constant as in equation 1.18. In principle, we still need to check that $D > 0$. However, since we know by lemma 28, that it is the diffusion constant of some Brownian motion process, this must automatically be satisfied. □

3.4. INVARIANCE PRINCIPLES IN RANDOM ENVIRONMENTS

In physics, it is commonly assumed that the medium is completely uniform. For example, when modelling a fluid, we tend to ignore small fluctuations in temperature and assume a completely homogeneous liquid with constant viscosity everywhere. Homogeneity is an approximation used for its practical merit in deriving theoretical results, but also in performing numerical simulations. It would be more realistic if thermal fluctuations, or other factors causing the medium to be locally disordered could be incorporated into these models. However, we expect that small local variations will average throughout the environment, i.e. that homogenisation takes place, thus validating the original assumption of complete homogeneity. It is therefore of particular significance to put homogenisation of the environment, on a rigorous footing. On the other hand, it is even more important to outline possible scenarios in which homogenisation does not occur. It might be possible that the local variations cause significant changes in the behaviour of the system at a global scale. Especially for applications in these scenarios, one needs to realise that modelling with the assumption of homogeneity may lead to different results than what happens in reality.
A particular application of a Markov process in a highly inhomogeneous environment is RNA transcription. In section 2.3.3, we already outlined several factors that influence RNAps on the DNA at a local scale. Therefore, for application of RNA transcription in particular it is important to verify that homogenisation takes place.

In light of this discussion, we will expand the mathematical theory of previous sections to inhomogeneous environments as well. In particular, by inhomogeneous environments, we mean random environments, in which the transition rates are taken as position dependent and drawn from independent, identical distributions. As a result, on large scales, the random environment still looks similar to the homogeneous environment, but locally, the medium is highly variable.

For these models, we will attempt to prove analogous limit theorems. However, due to the more complex nature of random environments, the proofs and theorems of previous sections do not directly translate over. In fact, several aspects of the previous section relied heavily on translation-invariance. For example, this caused the additive functional $\int_0^t V(X_s)ds$ to be independent of x_s, allowing use of lemma 25. In our calculations of diffusion constants, we exploited space homogeneity too, as we used Fourier transforms.

Nevertheless, we will see that even in inhomogeneous environments, one can still prove invariance principles for most models, by introducing some extra assumptions such as additional symmetry conditions on the processes. We will also present some models in 3.4.3 for which we were unable to show CLTs. These models indicate that there is either some missing pieces in the available mathematical theory to resolve these cases, or they point at the surprising possibility that homogenisation might not happen in them.

Before we can get into the details of those problematic models, we first need to specify what is meant by a random environment. There are two competing, main formulations of random environments; the random waiting time model and the random conductance model. The random waiting time model has been introduced by (Bouchaud and Georges, 1990) and has also become known as the Bouchaud Trap model. The main principle behind the Bouchaud model is the idea that each site is a trap in which particles remain stuck for some time. Some traps are larger than others, thus forming a disordered medium. The Bouchaud Trap model allows one to model the diffusive motion of charge carriers in a conductor with impurities (Bouchaud and Georges, 1990). However, one can also make the case that the Bouchaud model lends itself for RNA transcription. Earlier, we discussed the presence of histones on the DNA and the sequence-dependent movement of RNA polymerase. As a result, some of the nucleotide sites can be seen as traps; the polymerase is somewhat blocked out of escaping these sites. Since each nucleotide has a slightly different effect on the polymerase, the chain becomes a inhomogeneuos medium of random traps.

At the opposite end of the spectrum, we have the random conductance model. Instead of assigning transition rates to each site, the conductance model assigns the rates to each link between sites. Varying the transition rates between sites allows one to model the conductance of current in a resistor network (Doyle and Snell, 1984, Biskup, 2011). Aside from the connection with electrostatics, the conductance model can also be employed in biochemical applications. In fact, out of the two formalisms, the random conductance model seems best suited for modelling RNA transcription. One of the main drawbacks of the Bouchaud Trap model is the fact that the transitions have no preferential direction. Transitions to the left and right occur equally at each site and this seems incompatible with the sequence-dependent nature of the movement, as the left and right neighbours are not necessarily equal on a DNA chain. On the other hand, the random conductance model does allow for polarity at each site. However, the lack of polarity in the random waiting time model simplifies the mathematical analysis. As the random waiting time model is the easier of the two, we will commence with analysing this model, after outlining the general procedure in the next section.

3.4.1. THE STANDARD METHOD

Once again, we employ the standard scheme of section 3.3.2. Previously, the additive functional $\int_0^t V(X_s)ds$ was relatively easily dealt with using lemma 25. In random environments, this term becomes more difficult to resolve and we will rely on general results on additive functionals obtained by (Landim et al., 2012; Kipnis and Varadhan, 1986). Additionally, we introduce a new formalism for the positional process of the particle.

The associated process is known as the environment process (Biskup, 2011).
In the positional process, we track the movement of the particle from the position of a stationary observer in a fixed coordinate system. For the environment process, we change perspective and follow the trajectory from the particle's point of view. Similarly to changing observers in relativity theory, we now view the particle as stationary and the environment as moving. Therefore, the random walk becomes a process traversing from one environment to another. Due to the fact that the environments now actually vary in space, we can distinguish between different environments as opposed to the case where all transition rates were translation invariant. The environment process is usually notated using shift operators, which we will introduce in the upcoming sections.

3.4.2. BOUCHAUD TRAP MODEL

Just now, we introduced the Bouchaud trap model; the main idea behind this model is that each site operates as a trap, out of which the particle can escape with a certain rate. In the start of this section, we also stated that the random waiting time model is easiest to analyse. We will round off this subsection by discussing why this is the case.

THE FRAMEWORK

Let $\mathscr{S} = [k,K]^{(\mathbb{Z}\times I)}$ be the set of environments of random waiting times for some constants $K \geq k > 0$. Here, $I = \{0,\dots,n\}$ represents the set of possible layers.
The set of particle sites is $S = \mathbb{Z}\times I$, and each site is assigned a waiting time $w(x,\sigma)$, resulting in the environment $(w(x,\sigma))_{(x,\sigma)\in S} \in \mathscr{S}$. The coordinates are assigned waiting times based on a joint distribution ν. We shall assume that ν is independent of x. In some applications, ν will be layer-dependent and written as $\nu = (\nu_0,\nu_1,\dots,\nu_n)$ such that $\forall x,i,\forall A\subset[k,K] : \nu(w(x,i)\in A) = \nu_i(A)$. In other applications, we will demand layer-independence as well, resulting in $\nu_i = \nu_0$.
Given any environment w, the particle will make a diffusive random walk on $\mathbb{Z}\times I$ with transitions based on w. The transitions $(x,\sigma)\to(x\pm1,\sigma)$ both occur with rate $\frac{1}{w(x,\sigma)}$. Transitions between layers are modelled in the same way as always. Again, we assume that the layer-process is irreducible. So, given w, the generator of the process will be:

$$L_w f(x,\sigma) = \frac{1}{w(x,\sigma)}\big(f(x+1,\sigma)+f(x-1,\sigma)-2f(x,\sigma)\big) + \sum_{i\in L, i\neq\sigma}\lambda_{\sigma,i}\big(f(x,i)-f(x,\sigma)\big) \tag{3.24}$$

Definition 21. *We introduce the shift operators τ_e for any $e\in\mathbb{Z}$ as a function on $\mathscr{S}\to\mathscr{S}$.*

$$\tau_e(w(x,\sigma))_{(x,\sigma)\in S} := (w(x+e,\sigma))_{(x,\sigma)\in S} \tag{3.25}$$

The environment process $\{(w_t,\sigma_t) : t\geq 0\}$ is a Markov process on S, defined via its generator:

$$Lf(w,\sigma) = \frac{1}{w(0,\sigma)}\big(f(\tau_1 w,\sigma)+f(\tau_{-1}w,\sigma)-2f(w,\sigma)\big) + \sum_{i\in L, i\neq\sigma}\lambda_{\sigma,i}\big(f(w,i)-f(w,\sigma)\big) \tag{3.26}$$

INVARIANCE PRINCIPLES FOR THE BOUCHAUD TRAP MODEL

Due to the complexity of the model, even if we can prove a central limit theorem, obtaining a closed form formula for the diffusion constant D is often not possible. However, we shall commence with one of the unique setups in which we can actually find a diffusion constant D for the limiting motion.

Theorem 32. *Fix a random environment of waiting times w. Let $X_t = (x_t, \sigma_t)$ be a random walk in 1-layer, with diffusive motion based on w and generator 3.24.*
We demand that the measure ν_0, from which the random environment w is drawn, is ergodic under shifts.
Then, $x_t - x_0$ is a martingale that satisfies the Martingale CLT. That is, $\frac{x_{kt}}{\sqrt{k}} \xrightarrow[k\to\infty]{\mathbb{P}} B(Dt)$.
Furthermore, the limiting diffusion constant is given by

$$D = \frac{1}{\int w(0,0)\,d\nu_0} =: \frac{1}{M}.$$

Proof. Let $M := \int w(0,0)\,d\nu_0$. We will use this as a normalization constant.
Note that $x_t - x_0 - \int_0^t (L_w x)(x_s, \sigma_s)\,ds$ is a Dynkin martingale. Also note that $L_w x = 0$ here, since we only have diffusive motion. Therefore, $x_t - x_0$ is itself a martingale.
We compute the quadratic variation of this martingale:

$$[x,x]_t = \int_0^t L_w x^2 - 2xL_w x\,ds = \int_0^t \frac{1}{w(x_s)} \cdot 2\,ds = \int_0^t \frac{2}{w(x_s)}\,ds$$

We postulate that the reversible and ergodic measure of the environment process is given by $d\mu = \frac{w(0,0)}{M}\,d\nu_0$. Indeed, note that for stationarity we require $\int Lf\,d\mu = 0$:

$$\int Lf(w)\,d\mu = \int \frac{w(0)}{Mw(0)} \big(f(\tau_1 w) + f(\tau_{-1} w) - 2f(w)\big)\,d\nu_0 = 0$$

by shift invariance under ν_0.
For ergodicity, we need to check $Lf = 0 \implies f$ is constant almost surely.
Suppose that $Lf = 0$. Then, we compute $\langle f, -Lf\rangle_\mu$:

$$\begin{aligned} 0 = \langle f, -Lf\rangle_\mu &= \frac{1}{M}\int (2f^2(w) - f(\tau_1 w)f(w) - f(\tau_{-1} w))\,d\nu_0 \\ &= \frac{1}{M}\int (f^2(w) + f^2(\tau_1 w) - 2f(\tau_1 w)f(w))\,d\nu_0 = \frac{1}{M}\int (f(w) - f(\tau_1 w))^2\,dw. \end{aligned}$$

again by shift invariance of ν_0. So, we have that all terms in the integrand are non-negative. We conclude that $f(w) = f(\tau_1 w)$, thus f is constant ν_0-almost surely by the assumed ergodicity of ν_0 under shifts.
By proposition 23, we only need to check that $\mathbb{E}_{\mathbb{P}_\mu}(\frac{1}{t}[x,x]_t) < \infty$ to guarantee that $x_t - x_0$ satisfies the martingale CLT. Note that we can couple the positional process and the environment process by introducing the notation $w(x_s, i) = w_s(0,i)$. Also note that we can exchange the integrals by the Fubini-Tonelli theorem:

$$\mathbb{E}_{\mathbb{P}_\mu}\left(\frac{1}{t}[x,x]_t\right) = \int_0^t \frac{1}{t}\int \frac{2}{Mw_s(0)}\,w_s(0)\,d\nu_0\,ds = \frac{2}{M} = \sigma^2 < \infty. \qquad \square$$

We can now generalise this theorem to include multiple layers, with layer-dependent active and diffusive components as well. This provides an analogous theorem to 26 in random environments. However, as soon as we introduce layer-dependence in any fashion into the model, the calculation of D becomes irrecoverable. Therefore, we do not compute D and we will have to be content with a purely abstract invariance principle instead.

Theorem 33. *Central limit theorem for random waiting times:*
Fix a random environment of waiting times w. Let $(\nu_0, \nu_1, \dots, \nu_n)$ be the layer-dependent distributions of $w(\cdot, i)$. Let $X_t = (x_t, \sigma_t)$ be a random walk in $n+1$ layers.
The diffusive motion with random waiting times is based on w in each layer. Note that the distribution of w is i.i.d. inside each layer, but may be different between layers. Also, suppose that each layer has an active component with velocity v_i at rate 1.
Then, the process $\frac{x_{kt}-ckt}{\sqrt{k}}$ converges in probability to Brownian motion.

Proof. This time, note that $L_w(x) = \sum_{i=0}^{n} \mathbb{1}_{\sigma_s=i} v_i$.
Therefore, $x_t - x_0 - ct = M_t + \int_0^t \sum_{i=0}^{n} \mathbb{1}_{\sigma_s=i} v_i - c\,ds$ for some martingale M_t.
By lemma 25, we have that $\int_0^t \sum_{i=0}^{n} \mathbb{1}_{\sigma_s=i} v_i - c\,ds$ satisfies a central limit theorem. We can deal with M_t in the same way as in theorem 26. In this case, the generator is given by:

$$L_w f(x,\sigma) = \frac{1}{w(x,\sigma)}(f(x+1,\sigma)+f(x-1,\sigma)-2f(x,\sigma))+(f(x+v_\sigma,\sigma)-f(x,\sigma)+\sum_{i\neq\sigma}(f(x,i)-f(x,\sigma)).$$

Define $\frac{1}{m_i} := \int \frac{1}{w(0,i)}\,d\nu_i$. For the quadratic variation, we have:

$$\frac{1}{t}[M,M]_t = \frac{1}{t}\int_0^t \sum_{i=0}^{n} \mathbb{1}_{\sigma_s=i}(v_i^2 + \frac{2}{w(x_s,i)}) \leq \left(\sum_{i=0}^{n} v_i^2 + \frac{2}{m_i}\right) < \infty.$$

Again, $\mathbb{E}_{\mathbb{P}_\mu}(\frac{1}{t}[M,M]_t) < \infty$ now, since the expectation is taken over a probability measure.
So, by the Martingale CLT, the rescaled and centered process converges to Brownian motion for some diffusion constant D. To actually find D however, one would need to compute $\mathbb{E}_{\mathbb{P}_\mu}(\frac{1}{t}[x,x]_t)$ and find the stationary ergodic measure μ. □

Note that without the need for computing D, we could refrain from using the environment process and only needed to find a bound for the quadratic variation. This allowed for a vastly more general result than given in theorem 32.
Also note that the above theorem still holds when some layers do not have a diffusive component. We already allow each layer to have a different distribution on $[k,K]$, however it is still equally valid when some layers have "$w(\cdot,\sigma) = \infty$" everywhere, as long as any position can still be reached from any other position. Also note that the theorem is still equally valid when all layers do not only have i.i.d. waiting times, but have exactly equal waiting times. By this we mean, $w(x,\sigma) = w(x,0)$ always. In this case, the waiting times only need to be drawn for one specific layer and can be copied onto all others. This is a more realistic model in applications in which diffusivity is caused by the surroundings of the particle and therefore independent of the internal state of the particle.

DISCUSSION

The Bouchaud trap model has two main benefits from which the theorems given in the previous section follow. These benefits make it a simpler model to work with than the random conductance model. Firstly, we already discussed that sites have no preferential direction. As a result, the limit theorems hold for any possible Bouchaud environment. In the conductance model, environments which have a global preferential direction are technically possible. Even though the chance of drawing such an environment is 0, we certainly cannot prove limit theorems for these specific environments. Therefore, in the random conductance model, any convergence theorem must be in ν-probability, where ν is the joint distribution of the random environment.
The second benefit of the trap model is the additive functional only depending on the layer process. For example, in theorem 32, the term was zero. In theorem 33, the term was equal to the layer-dependent drift, which enabled use of the drift functional lemma.

3.4.3. THE RANDOM CONDUCTANCE MODEL

THE FRAMEWORK

In the previous section, we introduced the Bouchaud model, in which transition rates were picked for each site. In the random conductance model, the transition rates are drawn for each connection separately. So, let $\xi(x,y,i)$ denote the transition rate from $x \to y$ in layer i; i.e. the transition $(x,i) \to (y,i)$. Then, we draw $\xi \in \zeta \subset [k,K]^{(\mathbb{Z}\times I)^2}$.
Here, ζ is the set of environments that satisfy the following two equations:
1) $\xi(x,y,i) = \xi(y,x,i)$ 2) $\xi(x,y,i) \neq 0 \implies |x-y| = 1$
The transition operators τ_e are functions from $\zeta \to \zeta$ with $(\tau_e\xi)(x,y,i) = \xi(x+e,y+e,i)$.
For any random walk x_s in a random conductance environment, the environment process is taken as $\xi(x_s, x_s+e, i) = \tau_{x_s}\xi(0,e,i) := \xi_s(0,e,i)$.
We shall work with environments that are layer-independent. Therefore, we can drop the third variable and just write $\xi(x,y,i) = \xi(x,y)$. With this new formulation, we still leave the possibility open that some layers do not have a passive component at all, as these components can be left out of the generator. Similarly to the random waiting time model, we draw the environment from a joint distribution ν.

The conductance model is the first example where the additive functional also depends on x_s. Therefore, we will have to rely on results and techniques presented in (Landim et al., 2012) for dealing with this integral. Unfortunately, this greatly restricts the systems for which a CLT can be proven. In particular, any case where the drift c is nonzero is problematic. Therefore, we will not derive a comparable theorem to 33 in this section.
We shall briefly summarise the relevant available results from (Landim et al., 2012) here. We begin with a Markov process X_s that is ergodic and stationary with respect to measure π. (Landim et al., 2012) provides conditions that guarantee that $\int_0^t V(X_s)ds$ satisfies a central limit theorem. Here, $V(X_s)$ is a general function with $V \in L^2(\pi)$ and with $\mathbb{E}_\pi(V) = 0$. Notably, we have often not met the second requirement $\mathbb{E}_\pi(V)$ so far. Therefore, we need to additionally reduce the models that include drift to models with $\mathbb{E}_\pi(V) = 0$ so that we can use the results of (Landim et al., 2012).
On the next page, we give the simplest set of conditions that is sufficient for a functional central limit theorem.

Theorem 34. *Suppose that $V = Lf$ for some $f \in D(L)$ combined with $\sigma^2(V) < \infty$ where $\sigma^2(V)$ is a Green-Kubo integral for the variance of the limiting Brownian motion:*

$$\sigma^2(V) = 2\int_0^\infty \mathbb{E}_\pi(V(X_s)V(X_0))ds = 2\langle(-L)^{-1}V, V\rangle. \tag{3.27}$$

Then, the additive functional $\int_0^t V(X_s)ds$ satisfies a central limit type theorem.

For the remainder of the book, (Landim et al., 2012) considers a more general V than $V = Lf$ and tries to reduce it back to the original case, for example through approximating sequences. For these general functions, (Landim et al., 2012) provides several equivalent inequalities to $\sigma^2(V) < \infty$ that we can still employ when $V = Lf$ as well!
One of these equivalent inequalities is known as the sector condition:

Theorem 35. *Sector condition (Landim et al., 2012):*
Let $V \in L^2(\pi)$ such that $\mathbb{E}_\pi(V) = 0$. Suppose that
1) $V \in \mathscr{H}_{-1}$.
2) $\exists K \in \mathbb{R} : \forall f, g \in D(L) : \langle f, Ag\rangle_\pi^2 \leq K\langle f, (-Lf)\rangle_\pi \langle g, (-Lg)\rangle_\pi$
Here, A is the anti-symmetric part in the decomposition of L, i.e. $L = S + A$ for some self-adjoint S.
Then, the additive functional satisfies a central limit type theorem.

We shall not go into further detail on the meaning of $\mathscr{H}_{-1}$ here as it is irrelevant to our purposes. In any case, $V = Lf \implies V \in \mathscr{H}_{-1}$, so we may use the sector condition.

RESULTS

Equipped with the sector condition 35, we can begin proving some first CLTs in random conductance models. As soon as we introduce any drift into the model, we can no longer use the sector condition directly. Even indirectly, through reducing a process with drift to a drift-less one, we still run into trouble. We shall illustrate why the combination of random conductances and drift is problematic in this section. Furthermore, we will conclude with some illustrative examples of models with no drift for which CLTs do hold.

When we only have diffusive processes in each layer, the process automatically is drift-less. From a similar calculation as in theorem 32, we can conclude an analogous central limit theorem for diffusive processes.

Theorem 36. *Central limit theorem for random conductance models:*
Take a random conductance environment ξ based on the distribution $\nu = (\nu_0, \nu_1, \ldots, \nu_n)$. Let $X_t = (x_t, \sigma_t)$ be a random walk in $n+1 \leq 2$ layers, with diffusive motion in each layer based on the random conductances based of ξ. Assume that the layer process is irreducible.

Then, the rescaled and process $\frac{x_{kt}}{\sqrt{k}}$ converges in probability to Brownian motion with some diffusion constant D.

Proof. This time, $L_\xi(x) = \sum_{i=0}^n \mathbb{1}_{\sigma_s = i}\xi_s^i(0) - \xi_s^i(-1)$.
Therefore, $x_t - x_0 = M_t + \int_0^t \sum_{i=0}^n \mathbb{1}_{\sigma_s=i}\xi_s^i(0) - \xi_s^i(-1)ds$ for some Dynkin martingale M_t.

Note that the generator of the environment process is given by

$$Lf(\xi,\sigma) = \xi^{\sigma}(0)(f(\tau_1\xi,\sigma)-f(\xi,\sigma))+\xi^{\sigma}(-1)(f(\tau_{-1}\xi,\sigma)-f(\xi,\sigma))+\sum_{j=0,j\neq\sigma}^{n}(f(\xi,j)-f(\xi,\sigma))\lambda_{\sigma,j}.$$

We postulate that the stationary distribution of the environment process is $\mu = (\nu_0\pi_0,\ldots,\nu_n\pi_n)$ with π being the stationary distribution over the layers. For stationarity, we have:

$$\begin{aligned}
&\int\sum_{i=0}^{n}\left(\mathbb{1}_{\sigma=i}\xi^i(0)(f(\tau_1\xi,i)-f(\xi,i))+\xi^i(-1)(f(\tau_{-1}\xi,i)-f(\xi,i))+\sum_{j=0,j\neq i}^{n}(f(\xi,j)-f(\xi,i))\lambda_{ij}\right)d\mu\\
&=\int\sum_{i=0}^{n}\left(\xi^i(0)(f(\tau_1\xi,i)-f(\xi,i))+\xi^i(-1)(f(\tau_{-1}\xi,i)-f(\xi,i))+\sum_{j=0,j\neq i}^{n}(f(\xi,j)-f(\xi,i))\lambda_{ij}\right)\pi_i d\nu_i\\
&=\int\sum_{i=0}^{n}\left(\xi^i(0)(f(\tau_1\xi,i)-f(\xi,i))+\xi^i(0)(f(\xi,i)-f(\tau_1\xi,i))+\sum_{j=0,j\neq i}^{n}(f(\xi,j)-f(\xi,i))\lambda_{ij}\right)\pi_i d\nu_i\\
&=\int\sum_{i=0}^{n}\left(\sum_{j=0,j\neq i}^{n}(f(\xi,j)-f(\xi,i))\lambda_{ij}\right)\pi_i d\nu_i=\int Lf\,d\mu=0
\end{aligned}$$

Due to the last switching term being equal to $\pi Q\vec{f}$ for a layer-process generator Q with stationary distribution π, it is equal to zero.
We will only show the ergodicity for the 2-layer system. Assume $Lf=0$, then:

$$\begin{aligned}
\langle f,-Lf\rangle_\mu &= \int\sum_{i=0}^{1}\pi_i\left(\xi^i(0)(f^2(\xi,i)-f(\tau_1\xi,i)f(\xi,i))+\xi^i(-1)(f^2(\xi,i)-f(\tau_{-1}\xi,i)f(\xi,i))\right)d\nu_i\\
&+\left(\pi_0(f^2(\xi,0)-f(\xi,1)f(\xi,0))\lambda_{0,1}+\pi_1(f^2(\xi,1)-f(\xi,0)f(\xi,1)\lambda_{1,0})\right)d\nu\\
&=\int\sum_{i=0}^{1}\pi_i(\xi^i(0)(f(\xi,i)-f(\tau_1\xi,i))^2)d\nu_i+\frac{\lambda_{1,0}\lambda_{0,1}}{\lambda_{1,0}+\lambda_{0,1}}(f(\xi,0)-f(\xi,1))^2d\nu=0
\end{aligned}$$

All terms in the integrand are non-negative, but the integral is zero; therefore, all terms in the integrand must be zero, which confirms that f is ν-almost surely constant. The proof for $n+1=1$ is analogous and simpler.
Now that we know the stationary ergodic measure, we can verify the sector condition by showing that L is self-adjoint. Using the sector condition 35, we can conclude that the additive functional satisfies a CLT, as A must be equal to zero when L is self-adjoint.

$$\begin{aligned}
\langle f,Lg\rangle_\mu &= \int\sum_{i=0}^{1}\pi_i f(\xi,i)\left(\xi^i(0)(g(\xi,i)-g(\tau_1\xi,i))+\xi^i(-1)(g(\xi,i)-g(\tau_{-1}\xi,i))\right)d\nu_i\\
&+\frac{\lambda_{1,0}\lambda_{0,1}}{\lambda_{1,0}+\lambda_{0,1}}\left(f(\xi,0)(g(\xi,0)-g(\xi,1))+f(\xi,1)(g(\xi,1)-g(\xi,0))\right)d\nu\\
&=\int\sum_{i=0}^{1}\pi_i\left(\xi^i(0)(g(\xi,i)f(\xi,i)-g(\xi,i)f(\tau_1\xi,i))+\xi^i(-1)(g(\xi,i)f(\xi,i)-g(\xi,i)f(\tau_{-1}\xi,i))\right)d\nu_i\\
&+\frac{\lambda_{1,0}\lambda_{0,1}}{\lambda_{1,0}+\lambda_{0,1}}\left(g(\xi,0)(f(\xi,0)-f(\xi,1))+g(\xi,1)(f(\xi,1)-f(\xi,0))\right)d\nu=\langle Lf,g\rangle_\mu.
\end{aligned}$$

We can deal with M_t in the same way as in theorem 26. The generator is given by:

$$L_\xi f(x,\sigma) = \xi^\sigma(x,x+1)(f(x+1,\sigma)-f(x,\sigma))\xi^\sigma(x,x-1)(f(x-1,\sigma)-f(x,\sigma))+\sum_{i=0,i\neq\sigma}^{n}(f(x,i)-f(x,\sigma))\lambda_{\sigma,i}.$$

Thus resulting in:

$$\frac{1}{t}[M,M]_t = \frac{1}{t}\int_0^t \sum_{i=0}^{n} \mathbb{1}_{\sigma_s=i}(\xi_s^i(0)+\xi_s^i(-1)) \leq 2K < \infty$$

So, by the martingale CLT, the rescaled process converges to Brownian motion. □

We will now illustrate why attempting to bring an active component into the system fails. As an example, we consider a 2-layer model with only active motion in the bottom layer and only diffusive motion in the top layer.

$$L_\xi f(x,\sigma) = \mathbb{1}_{\sigma=1}\Big(\xi(x,x+1)(f(x+1,1)-f(x,1))+\xi(x,x-1)(f(x-1)-f(x))$$
$$+\lambda_{1,0}(f(x,0)-f(x,1))\Big)+\mathbb{1}_{\sigma=0}\Big(f(x+1,0)-f(x,0)+\lambda_{0,1}(f(x,1)-f(x,0))\Big).$$

The Dynkin martingale is: $x_t - ct = x_0 + \int_0^t \mathbb{1}_{\sigma_s=1}(\xi_s(0)-\xi_s(-1)) - c + \mathbb{1}_{\sigma_s=0}ds + M_t$.
As always, we can get rid of $\mathbb{1}_{\sigma_s=0} - c$ by using lemma 25. We then recognise that the remaining terms have the same form as in theorem 36. However, it would be incorrect to infer that the term $\mathbb{1}_{\sigma_s=1}(\xi_s(0)-\xi_s(-1))$ is the exact same as in the purely diffusive scenario. In fact, in our active model, the environment is not explored in a symmetric fashion. Consequently, $(\xi_s(0)-\xi_s(-1))$ is not the same process as it was in theorem 8. This asymmetry therefore forbids us from reducing to the previous theorem.
A second natural attempt is to incorporate the drift correction $-ct$ into the process such that we once again get an additive functional $\int V(X_s)ds$ with V having mean zero. So, we could view $x_t - ct = y_t$ as a new process with generator

$$L_\xi f(y,\sigma) = \lambda_{\sigma,1-\sigma}(f(y,\sigma)-f(y,\sigma))+\mathbb{1}_{\sigma=0}(f(y+1,0)-f(y,0)-c(f(y-1)-f(y)))$$
$$+\mathbb{1}_{\sigma=1}(\xi(y,y+1)(f(y+1,1)-f(y,1))+\xi(y,x-1)(f(y-1)-f(y))-c(f(y-1)-f(y))).$$

Unfortunately, it becomes impossible to show that the sector condition is upheld. Therefore, we can only apply the theory to specific cases in which either of these two methods of reduction actually do work.

Example 4. *The first natural example in which the second approach works is when the active component is the exact same in each layer. If this is the case, adding the drift correction into the generator combines in each layer to form a diffusive term. By theorem 36, we know that these processes satisfy a central limit theorem.*
The second natural examples occur when the total drift is zero and the process is reversible in nature. By reversibility, we mean that it satisfies time symmetry. For example, consider a 2-layer process in which we have the same diffusive behaviour in both layers and equal switching between layers. Additionally, we have an active component with rate 1 in each layer, but with jumps in opposite directions. By a theorem of (Kipnis and Varadhan, 1986), this process follows a central limit type theorem.

DISCUSSION

Despite the fact that the homogeneous environment central limit theorems of 29, 30 and 31 could be generalised without too much trouble for the Bouchaud trap model, the random conductance model turned out to be surprisingly resistant to analysis for environments with drift. The asymmetric exploration of the environment combined with the asymmetric rates at every site lead to processes that cannot be resolved with the available mathematical theory. Although much research has been done on both layered systems without any random conductances and multi-dimensional random-conductance models without any layer-dependence, there seems to be little literature connecting the two. By multi-dimensional random conductance models, we mean in particular the so-called ballistic random walks. Similarly to our jump processes, these are Markov processes in multi-dimensional grids where the particle drifts off in a certain direction. Many invariance principles have been established for these ballistic processes in environments with random conductances, such as in $\mathbb{Z}^d$ (Rassoul-Agha and Seppäläinen, 2009; Sznitman, 2002; Zeitouni, 2004).
However, it is important to note that all of these ballistic models assume that all states in the entire space follow the same distribution. Therefore, these results cannot be recovered easily when introducing layer-dependence into the model, which is the main purpose of using the layers in the first place!

Nevertheless, we were still able to provide CLTs for some specific random conductance systems that are useful in physical applications. In example 4, we presented a system of 2 layers with flipped active motion in the layers, which is comparable to the one-dimensional behaviour of the bacteria *M. xanthus* mentioned in section 2.3.2. We can conclude that the motion of this run and tumble particle is similar to a Brownian particle and that this holds in both homogeneous and inhomogeneous environments.

Theorem 33 is our main evidence for homogenisation being valid in a wide range of Bouchaud trap models and this indicates that we should expect homogenisation to apply to the random conductance model as well. In the upcoming chapter 5, we will also perform some numerical simulations to further investigate homogenisation in the random conductance model and these initial simulations support this hypothesis.

CHAPTER 4

NUMERICAL METHODS AND SIMULATING JUMP PROCESSES

In chapter 2, we described potential applications of the theory of jump processes. Recall that these layered jump processes are trajectories of particles with multiple internal states and exponentially distributed transitions. In the previous chapter 3, we conducted a mathematical analysis of single-particle jump processes. However, we are also interested in studying physical, multiple particle systems, which are more difficult to analyse mathematically. We shall now present a general numerical model for jump processes that can be used to simulate multi-particle systems. In the next chapter 5, we will apply this numerical model specifically to modelling RNA transcription.

4.1. GILLESPIE ALGORITHM

The Gillespie algorithm is widely used in biochemical applications to simulate systems with jump process transitions numerically (Kierzek, 2002). The algorithm is an exact method and completely based on the formulation given in proposition 13.
The algorithm operates as follows (Gillespie, 2007). Given a certain initial system, we sum the rates of all possible transitions into one total rate λ. As the probability of having two events occur simultaneously is 0, we may consider the transitions sequentially. We determine how long it takes for one transition to occur by sampling a time increment from an $\exp(\lambda)$ distribution. Afterwards, we use theorem 12 to pick the precise transition that occurred. For example, the transition from state i to j is picked with probability $\frac{\lambda_{i,j}}{\lambda}$. However, the transition is only performed if the exclusion interaction allows it.

4.2. IMPLEMENTING THE GILLESPIE ALGORITHM

Although the Gillespie algorithm itself is quite simple, there are various techniques that can be used to speed up the implementation of the algorithm. Optimisations and often even approximations methods are necessary, as the Gillespie algorithm is computationally intensive for systems with many particles (Rao and Arkin, 2003).
Instead of using an approximation method, we will be looking at 3 main aspects of our model that can be sped up significantly by optimising the naive implementation. We will outline these optimizations using an example to illustrate the ideas more clearly.

All optimizations focus on the part of the Gillespie algorithm where the occurred event needs to be picked. This event is picked from a list of all possible transitions at that moment, where each transition has a probability $\frac{\lambda_{i,j}}{\lambda}$ to be picked. A natural way to implement this picking procedure is through the use of **lists of cumulative sums**. We will illustrate what is meant by this through an example.

Example 5. *We consider a* 3*-state system containing a particle in both states* 0 *and* 1*. For illustration, we take* $\lambda_{0,1} = 1, \lambda_{0,2} = 0.5$ *and* $\lambda_{1,2} = 0.25$ *with* $\lambda_{i,j} = \lambda_{j,i}$*. The time increment is generated with rate* $1+0.5+1+0.25 = 2.75$*. We assign an ordering to construct the list of cumulative sums. In this case, we can take the order in which the transitions rates where given, resulting in the cumulative list* $[1, 1.5, 2.5, 2.75]$*. We then generate a* $U[0, 2.75]$*-distributed random variable* X *to determine which transition occurs. To do so, we use a built-in bisection method to find the position in which* X *can be inserted in the list while keeping the ordering intact. For example, if* $X \in (2.5, 2.75)$*, we can insert it in the fourth position of the list. Consequently, we pick the fourth transition to take place, which is* $(p_1, p_2) = (0,1) \to (0,2)$*, where* p_i *is the coordinate of particle* i*. Note that the probability that this transition took place is exactly* $\frac{2.75-2.5}{2.75} = \frac{\lambda_{2,3}}{\lambda}$*. Due to the use of cumulative sums and a uniformly distributed* X*, we always attain a probability* $\frac{\lambda_{i,j}}{\lambda}$ *for transition* $i \to j$*. Had* X *been lower than* 1*, we see that the transition* $(0,1) \to (1,1)$ *is attempted. However, this transition is forbidden by the exclusion interaction and hence blocked. In this case, the time is still increased by the increment, but the state of the system remains unchanged.*

A naive implementation would be to fully rebuild the cumulative list every time step. However, previous lists can also be used to build new ones and this greatly reduces the amount of computations needed. Note that there are three categories of transitions: particles can enter or leave the system, or transition between certain states in the system. For each of category, we will present a method to update the cumulative list.

Firstly, the addition of new particles into the system. When this happens, we can account for the extra rates of a new particle by appending them to the previous list. In example 5, adding a particle to state 3 would change the list to $[1, 1.5, 2.5, 2.75, 3.25, 3.5]$. Of course, this example is a bit redundant as the entire system is now clogged up due to exclusion. However, in other systems with more free space, the technique of appending the new rates works similarly. By working in this manner, we ensure that the rates for all other particles do not need to be altered, which is where most of the time-save is gained.

Secondly, the removal of particles from the system. Since we are looking at systems in which particles cannot overtake each other, we know that the first particle to enter the system is also the first to exit the system. Therefore, when a particle leaves, we would have to remove the first element of the cumulative list and update all the others. However, we can instead also change the way we draw the random variable X by using offsets. Suppose that the first particle leaves and used to contribute a total transition rate of λ_1 to the total rate λ. We remove all entries of the list associated with the first particle, but leave all others unchanged. So, in example 5 this would change the list from $[1, 1.5, 2.5, 2.75]$ to $[2.5, 2.75]$. We add λ_1 to our offset o and from now on draw X from $[o, \lambda]$ instead of $[0, \lambda - \lambda_1]$. Consequently, none of the transition probabilities change, but we also do not need to subtract λ_1 from all other entries, thus saving a lot of computations.

The final optimisation technique deals with transitions of existing particles. Suppose that the i-th particle transitions. Note that only the sums at indices of at least i will change due to this transition. On average, when the rates at each position in the system are comparable to each other, only half of the list needs to be updated after a transition.

Combining these three improvements already significantly improves the speed at which the model computes simulations. In fact, most of our simulations on RNA transcription could be completed within a couple of hours, which was sufficiently fast for gathering the needed data for our purposes. For more complicated systems, such as ones with thousands of particles, further optimizations or even approximations can be introduced.

4.3. THE NUMERICAL MODEL OF TRANSCRIPTION

In figure 4.1, we present a schematic of a general model that can simulate any jump process and which employs all of the outlined techniques of section 4.2. This schematic was implemented in Python code and used for simulations on RNA transcription. This Python code can be found in (van der Post, 2024).

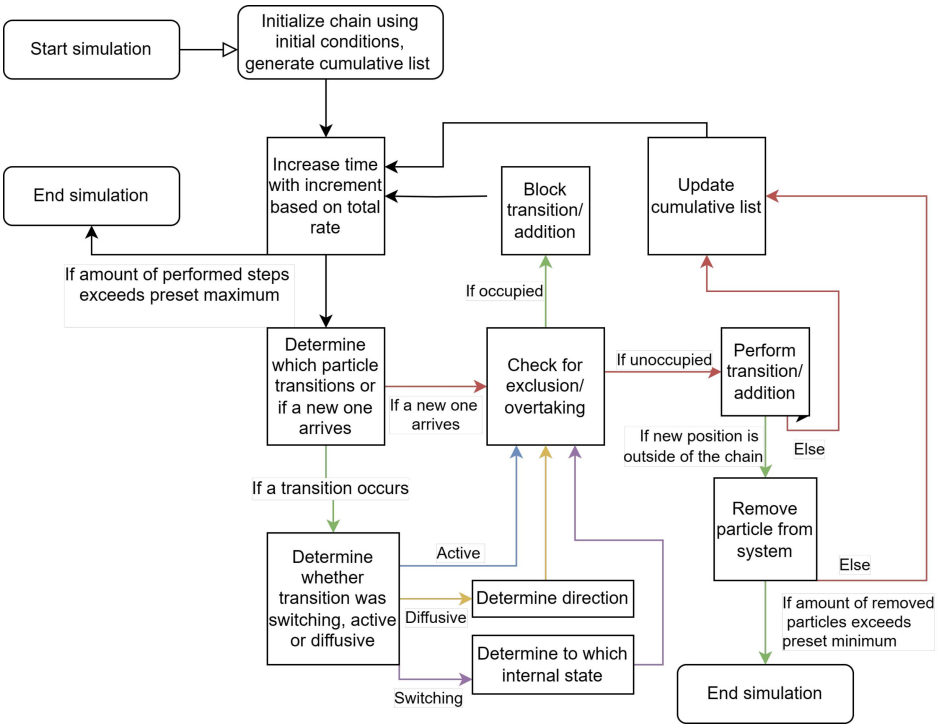


Figure 4.1: A comprehensive flowchart detailing the general model for simulating jump processes. This model will be applied to simulating RNA transcription in the next chapter 5. The boxes represent tasks that need to be performed for the simulation. The arrows indicate chronological transitions between tasks. The different colours of the arrows are used to distinguish between different options exiting the same box. Note that the schematic is divided into three main parts. Firstly, the upper two boxes represent the initiation or start of the simulation. Secondly, the leftmost and bottom-right boxes indicate the termination or end of the simulation. Everything else is a loop that is continuously repeated to simulate space for a prolonged period of time.

The schematic contains several components that require some more clarification, which we will now present in more depth.

EXCLUSION / OVERTAKING

In the flowchart we emphasise that both exclusion and overtaking are checked. In principle, the distinction between overtaking and exclusion is not necessary for our transcription model. One of our key assumptions, outlined in section 2.3.3, is that the RNAps do not overtake each other. The mentioning of overtaking is therefore exclusively for other potential applications, such as the ones given in section 2.3.

INITIALIZATION

The first stage of the simulation is initialising the DNA chain. In essence, this means actually imposing the initial conditions of the system. In most simulations, we will begin with an empty chain. Since we want to conduct steady-state simulations, it may seem counterproductive to start with an empty initial configuration. However, we shall see in section 5.2 that the system quickly converges to steady state in our model. Hence, the small period in which it is not steady state does not affect the eventual results.
In a handful of simulations, we want to be able to force the chain to attain a certain polymerase density to analyse congestion effects. In these scenarios, we need to begin with a chain that already has a specific RNAp density.

TERMINATION

There are two conditions that automatically end the simulation once they are fulfilled. The first condition is based on the amount of polymerases that reached the end of the chain. After a preset amount N_{finished} have finished transcription, we stop the simulation since enough data has been gathered. The second condition is based on the amount of computation steps and is intended as a failsafe. We place a restriction on the maximum amount of steps that can be performed before the simulation ends. In case the acquisition of data takes too long, or something else goes wrong, this condition ensures that the simulation will abort eventually. Under normal conditions, N_{finished} is obviously chosen small enough such that it is reached before the maximum amount of steps.

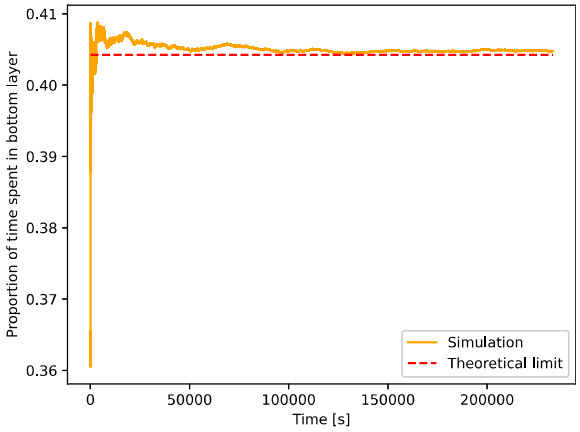
DISTINCTION IN DETERMINING EVENTS

In section 4.2, we picked the occurring transition based on one list containing all possible transitions. In practice, we actually divide the picking into multiple stages. First, we determine which RNAp makes a transition, based on the total rate of particle: λ_i. This also includes the possibility that a new particle enters the model, which happens with rate α. Only after picking the particle, we determine what happens to that RNAp. This can be an active jump, a diffusive jump or a switching of layer. Once the type of jump is determined, the exact location to which the RNAp moves is picked. Note that by theorem 12, dividing the picking procedure into stages does not affect the actual outcome.

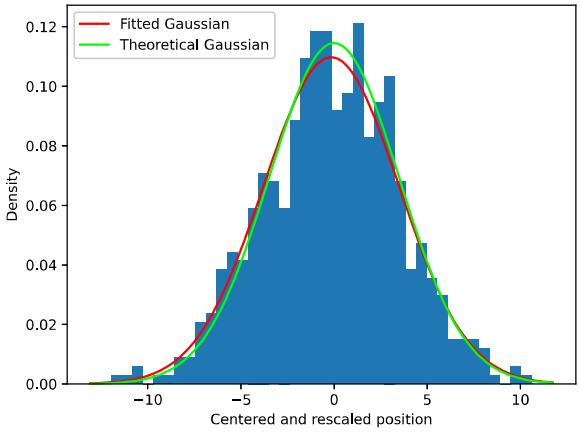
4.3.1. VERIFYING THE VALIDITY OF THE NUMERICAL MODEL

Since the numerical model is already quite involved, it is good practice to perform some sanity-checks on it. For instance, we established in theorem 18 and corollary 19 for finite state spaces, what fraction of time a particle should be in each internal state. Since the layer-process of the particles is independent of all other processes, we can consider this finite state space process separately from the rest and apply this corolary. We shall now verify that results generated from this model back up the conclusion obtained in corollary 19. For this purpose, we took a sample 3-layer system. We conduct 600 single-particle simulations of $4 \cdot 10^6$ transition steps each. The size of the system is larger than $4 \cdot 10^6$, so that the particles will not reach the boundaries during the simulations and thus this is comparable to simulating in infinite space.
The transition rates are picked as somewhat arbitrary coprime integers:
$\lambda_{0,1} = 1, \lambda_{0,2} = 5, \lambda_{1,0} = 11, \lambda_{1,2} = 7, \lambda_{2,0} = 3, \lambda_{2,1} = 2$.

Layer 0 only has active motion at rate 1 and size $v_0 = 2$, layer 1 is the pause layer and layer 2 only has diffusive motion at rate 10. The outward transition rates are picked highest for layer 1, so that the particle spends most of the time in the more interesting layers.



(a) A realization of the layer process of one of the 600 runs of $4 \cdot 10^6$ time steps. By theorem 18, we expect the layer presence to converge to the theoretical limit of 0.40426 and the simulated data did in fact closely approach this limit. In the figure, we present the lowest layer only; the trajectories looked analogous for the other layers. Obviously, the trajectory is quite volatile in the first couple of time steps due to the initial condition. Therefore, the first 1000 steps are left out of this data set, so that we can view the convergence on a closer scale for the remaining time steps.



(b) The simulated distribution of the rescaled process after $4 \cdot 10^6$ time steps for 600 runs. By theorem 31, we expect the distribution to be Gaussian with mean 0 and standard deviation 3.48. The theoretical Gaussian is plotted in green and had a Kolmogorov-Smirnov (KS) p-value of 0.288. We also add a Gaussian fit to the data, which resulted in fitted parameters $-0.153, 3.634$. The fitted Gaussian had a KS p-value of 0.929. These results and p-values increase our confidence in the validity of the implementation of the model, as they correspond reasonably well to the theoretical predictions.

By theorem 18, the proportion of time spent in each layer i should converge to π_i, where π is the stationary distribution. By equation 1.18, the stationary distribution is given by $(\pi_0, \pi_1, \pi_2) \approx (0.40426, 0.07979, 0.51596)$. In figure 4.2a, we see a realization of the layer presence of one of the particles. Clearly, this realization closely approaches the theoretically predicted limit. In fact, averaging over all 600 runs gave $(0.40428, 0.07977, 0.51594)$ as average proportions, with respective standard deviations of $(5 \cdot 10^{-4}, 2 \cdot 10^{-4}, 5 \cdot 10^{-4})$. The correspondence between theoretical predictions and simulation results confirms that the switching processes in the simulation are operating as expected.

Secondly, we test the result of the drift formula 24 in a different set of simulations. The theoretical value for the drift is $c = \frac{38}{47} \approx 0.8085$. We perform 100 distinct simulations to estimate the average of $\frac{x_t}{t}$ after $2 \cdot 10^6$ time steps. The obtained data set has a mean of 0.80846 and standard deviation 0.0112, which again aligns nicely with the prediction.

We also check the result of central limit theorem 31 against the simulation. We use the original 600 runs to estimate the distribution of $(x_t - ct)t^{-0.5}$ after $4 \cdot 10^6$ steps. In figure 4.2b, the resulting distribution is compared against the theoretical Gaussian and a fitted Gaussian. The predicted and simulated means and standard deviations are $0, -0.15$ and $3.48, 3.63$ respectively. The simulations and predictions are close, but not as convincingly as for example for the layer process. This suggests that the convergence to the Gaussian goes at a slower pace and more steps are needed to get a closer approximation.

All in all, the numerical simulations correspond closely to the theoretical predictions, which implies that the implementation is functioning as intended.

CHAPTER 5

MODELLING RNA TRANSCRIPTION

In the previous chapter 4, we introduced the computational model of layered jump processes in a general setting. Now, we can finally apply the model to perform multi-particle simulations on RNA transcription. In particular, we are interested in the rate of transcription. From the biological perspective, this rate is of utmost importance as the amount of transcription that is performed on a particular gene greatly influences how fast and how many specific proteins can be synthesised. Therefore, the transcription rate is the main feature the cell will aim to regulate to accommodate for the biological needs at a certain moment. We will see that the way in which different RNAps traverse the DNA and interact with each other naturally leads to a maximal flow of RNAps that can pass through the system and thus that the transcription rate is limited by this optimal flow.

Recall from section 2.3.3, that the cell has a high degree of control over the binding rate on the gene, via the use of transcription factors. We will therefore mainly be investigating the effect different binding rates have on the total transcription rate and the different phenomena that take place on the DNA to optimise this transcription rate. For this purpose, we will be closely following the analysis performed by (Klumpp, 2011) and a similar model as in (Klumpp, 2011 and Klumpp and Hwa, 2008) and aim to reproduce their results. These studies identified in particular a phenomenon through which different interacting RNAps push each other forward on the DNA, thus leading to a higher RNAp velocity and a higher transcription rate. We will investigate how this phenomenon arises and reproduce this "RNAp cooperation" effect in our simulations as well.

Additionally, we will compare our numerical results to the mathematical results. We especially expect that the two yield similar results in the case of sparse polymerase density on the DNA chain. Nevertheless, the mathematical analysis is still useful for high RNAp densities as well. In particular, it helps us identify the main aspects that are different or that change when comparing dense environments to sparse environments. These considerations in turn aid us in explaining the new phenomena that arise in dense environments, such as the cooperation effect.

5.1. THE MODEL AND PARAMETERS

We have already extensively discussed general layered jump process models in chapters 1 and 3 and presented how to apply these models to the particular case of transcription in sections 1.2.3 and 2.3.3. We will now explain the specific layered jump process model of transcription in more detail. Additionally, we introduce the actual used values of the relevant physical parameters, which are taken from (Klumpp, 2011).

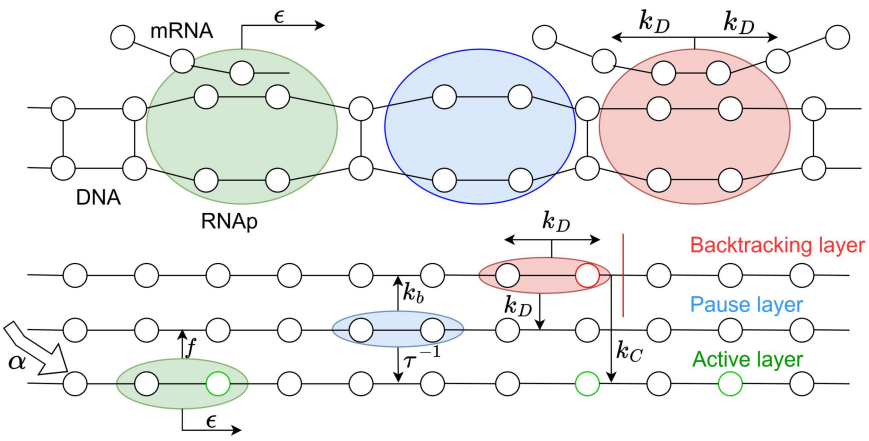


Figure 5.1: The basic jump process model of transcription. Note that the string of beads represent the DNA chain, while the coloured particles depict RNAps in different internal states. There are three internal states: backtracking (red), pausing (blue) and active transcribing (green). In the upper part of the figure, we sketch the scenario on the actual DNA chain, with zipped open DNA parts, transcribed mRNA and moving RNAps. In the lower part, we depict how we model this process by using different internal layers.
At the left, new polymerases attempt to bind to the DNA with rate α. However, they can only enter if enough space has been vacated, as RNAps interact via the exclusion interaction. RNAps exit the chain at the end of the gene after finishing transcription. In the active state, the RNAp moves forward with stepping rate ϵ and switches to the pause state with rate f. From the pause state, the RNAp either returns to the active state with rate $\frac{1}{\tau}$ or enters a backtrack. In a backtrack, the RNAp moves diffusively with rate k_D, but cannot pass the initiation site. It can either re-enter the pause state at the initiation site with rate k_D, or cleave to the active state with rate k_c. We will consider both a scenario in which particles can switch states at every site and one in which RNAps can only do so at specific "pause sites". The latter case is depicted by the coloured pause sites.

We take RNAp width of $w = 50$ nucleotides (nt), gene length of $L = 3000$ nt and a variable initiation attempt rate α. The stepping rate is $\epsilon = 100\text{s}^{-1}$, while the diffusive rate is $k_D = 10\text{s}^{-1}$. For switching rates we let $f = 10\text{s}^{-1}, \tau = 1\text{s}, k_b = k_D, k_c = 0.05\text{s}^{-1}$. It is however often assumed that the RNAp can only enter the pause state at specific pause sites on the DNA. As was explained in section 2.3.3; pause sites form at certain positions on the DNA due to the presence of histones, transcription factors and recognisable nucleotide sequences. The switching rate of $f = 10\text{s}^{-1}$ is therefore only applicable at these pause sites. Usually, we shall assume that there are 50 pause sites, which is slightly higher than the 1% pause site density taken in (Klumpp, 2011).

Recall that the model presented in figure 5.1 is based on the general jump process systems, but needs some additional components to simulate RNA transcription more realistically. In particular, we require the specific pause sites and the backtracking barrier preventing the polymerase from moving past the initiation site. Additionally, the cleaving process of the RNAp is site-dependent. Although the RNAp can cleave the transcript to return to the active state after a backtrack anywhere along the DNA, re-entering the pause layer from the backtracking layer can only be done at the site where the backtrack was initiated. We will perform analyses with and without these extra features to compare and contrast between the different results. Nevertheless, the simplest possible model, with no pause sites, no barriers and switching rate k_D and k_c at any position in the third layer already exhibits some interesting behaviour, such as RNAp clustering and optimal RNAp flow, that will continue to occur in the more complex models as well.

Therefore, we will begin with an analysis of this simple model. Since we now assume that every position in the first layer operates as a pause site, we need to adjust the rate f to accommodate for this large increase in the amount of pause sites. Hence, we adjust f in such a manner that the proportion of time a particle spends in a backtrack is roughly the same for the model with 50 pause sites and the simple model. We know that 50 pause sites of rate 10s^{-1} and 3000 pause sites of rate $10 \cdot \frac{50}{3000}$ will produce the same results in this regard, and will therefore employ the corrected rate of $\bar{f} = \frac{1}{6}\text{s}^{-1}$ in these simulations.

We will use steady-state simulations to measure the transcription rate and a couple of related quantities. Some initial trial simulations revealed that steady state is without a doubt reached in all models after 200 RNAps have finished transcription. Steady state can for example be recognised by the computation time increasing linearly with the amount of computation steps and by the rate at which RNAps finish the chain stabilising. The quantities that we will investigate in steady state are the average velocity, the completion rate, the steady state density and the transcription rate. The average velocity is defined as the velocity an RNAp particle has on average, throughout its journey on the DNA track. So, this velocity is akin to the drift velocity and calculated by the length of the gene divided by the time it takes for the particle to complete transcription. The completion rate is defined as the amount of RNAp that finishes every second. The polymerase density is defined as the amount of polymerase per nucleotide that is on average on the chain. Therefore, this density is at most $\frac{1}{w} = 0.02$ when the chain is entirely filled. Lastly, the transcription rate quantifies the amount of mRNA that is transcribed on the gene each second, by measuring the flow of RNAps on the DNA. It can be found via two equivalent calculations. Firstly, the transcription rate is the completion rate multiplied by the gene length of 3000. Equivalently, the transcription rate can also be found by multiplying the RNAp density with the RNAp velocity and the length of the gene.

Hence, the transcription rate is the result of a fine balance between the amount of RNAps on the chain and the speed at which these RNAps can move. As the amount of particles increases, we expect the maximal speed to drop. Picking a density that is too high will be outweighed by a drastically lower velocity and vice versa. As a result, there will be a certain maximal transcription rate characterising the maximal flow through the network. We will refer to this maximal rate and the corresponding RNAp density as optimal. In the upcoming section, we will show that the system automatically approaches the optimal conditions if the binding rate is chosen large enough. Additionally, if the DNA chain temporarily attains a higher RNAp density, it will eventually revert back to the optimal scenario. In a sense, the optimal density can therefore be seen as a stable equilibrium.

5.2. DYNAMICS IN THE BASIC TRANSCRIPTION MODEL

In the basic jump process model, we do not include the pause sites, barriers and site-dependent cleaving and use pause rate $\bar{f}$ instead of f. We will now study the dependence of transcription on the initiation attempt rate α and the consequent RNAp density ρ on the chain. We will see that the results are quite similar to one would be expect for traffic flow through a network. As such, we will also interpret the results from this perspective.

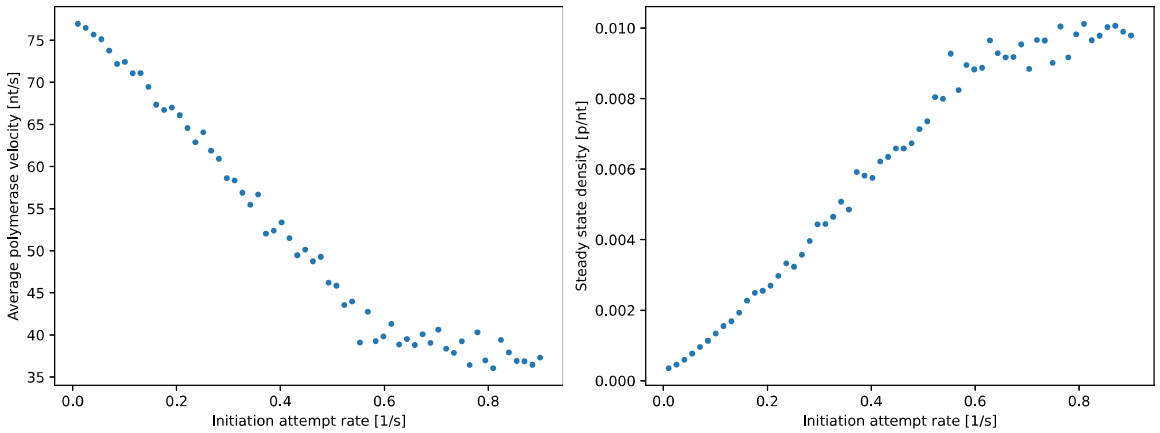


Figure 5.2: Two figures depicting the RNAp dynamics on the DNA chain for 50 different values of the initiation attempt rate α. Each data point is based on a simulation with 1500 finished RNAps.
In the left figure, we depict the average velocity of the 1500 RNAps. In the right figure, we present the accompanying average steady state density. In sparse conditions (low α and thus low density), we see that the drift velocity is slightly higher than the theoretical prediction of 75.9nt/s; which can be found from theorem 24 using the specific model parameters. We attribute this small difference to the fact that all RNAps are initialised in the active state and hence have a small bias for the first layer that is noticeable on a scale of 3000 nt. As α increases, the chain becomes more crowded and the RNAp density increases. Dense traffic begins to form on the DNA causing a decrease in overall speed. The decrease is first linear, but flattens out for larger α and settles at around half of the maximal velocity. As the velocity converges, the density also approaches a limit of about 0.01 at large α. Note that 0.01 RNAp/nt corresponds to 50% occupation, as the RNAps have a width of 50 nt and thus 100% occupation would mean 1 RNAp for every 50 nucleotides. The density is not completely stable in the high α region, which indicates that the different simulations exhibit slight differences in long term behaviour, possibly due to long-term congestion forming in some chains but not in others.

In figure 5.2, we present the RNAp dynamics of this initial, simplified model. As expected, an increasing attempt rate α leads to a rising density and a decreasing velocity. The density seems to be bounded by an occupation of circa 50%, no matter the rate α. This suggests that the optimal flow of the system, as discussed in section 5.1, is obtained at this combination of density and velocity. The reasoning behind this conclusion is as follows. By mass conservation, this maximal flow is a balance between the inflow and the outflow. Since the outflow is bounded, it takes time for particles to exit the system. As a result, the particle density increases as more and more particles enter with increasing α. The inflow is in turn dependent not only on α, but also on the availability of free space on the chain. At higher densities, this availability is diminished, making the arrival of extra particles unlikely, even despite the high α. At some point, we reach the optimal density which maximises the outflow. The inflow is restricted by this maximal outflow; so as we keep increasing α, the inflow does not increase further due to a lack of available space. Hence, we expect the system to steer towards a local maximum in flow with increasing α.

To show that the above intuitive reasoning holds true, we need to verify that the RNAp flow is suboptimal at densities higher than 50%. We were able to confirm this result through two different methods. Firstly, we conducted steady-state simulations of systems initialised at high α and high densities such as 100%. We looked at the steady-state

flow through the network after 200 particles finished transcription. These simulations revealed that it is actually impossible to sustain a high density in steady state. No matter the initial condition, the system converges to a steady state in the condition of optimal flow at 50% occupation. From this, we can conclude that the **state of optimal flow is stable** and systems at high densities temporarily have an imbalance between inflow and outflow automatically returning them to the state of optimal flow.
To quantify this in- and outflow imbalance, we also performed measurements to determine the temporary outflow at high densities. Each measurement is an average over 1000 simulations of systems at the same high density. Each system is initialised with evenly spread RNAps that together constitute the total density. For each simulation, we estimate the transcription rate by measuring the amount of mRNA actively transcribed by all RNAps before the first particle finishes and divide by the elapsed time. The results of these simulations can be seen in figure 5.3.

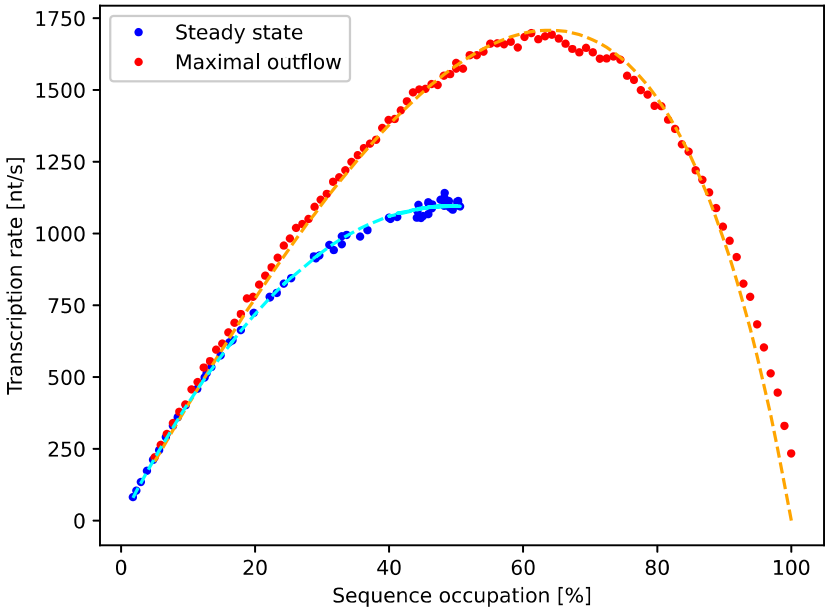


Figure 5.3: Simulations comparing the steady state flow to the temporary outflow immediately after initialization at a specific density ρ. The transcription rates coincide for low values of ρ, but even before the optimal flow has been reached, the temporary outflow begins to exceed the optimal steady state flow. Nevertheless, the graph of maximal outflow confirms our intuition that there is a fine margin between having a too large density, causing the decrease in velocity to outweigh the gain in amount of polymerase, and having a too small density, with the opposite effect. As a result, the graph follows a standard shape of traffic flows (Klumpp, 2011), which has been added as fit based on the relationship $c\frac{p(1-p)}{k-kp+p}$, where c and k are constants. The fits closely match the observed data and have R^2 values of 0.998 and 0.986 for steady state and outflow respectively. In the same order, the fit constants were $(c,k) = (4.21\cdot10^3, 0.921), (1.31\cdot10^4, 3.15)$.

What immediately stands out in figure 5.3, is the maximal outflow already exceeding the optimal flow at low densities. This can be explained by the fact that the chain was initialised with evenly spread out RNAps. From the discrepancy, we see that in steady state something different happens. We can conclude that in reality, the RNAps somewhat clump together in the density range $20-50\%$, thereby lowering the flow of the system.

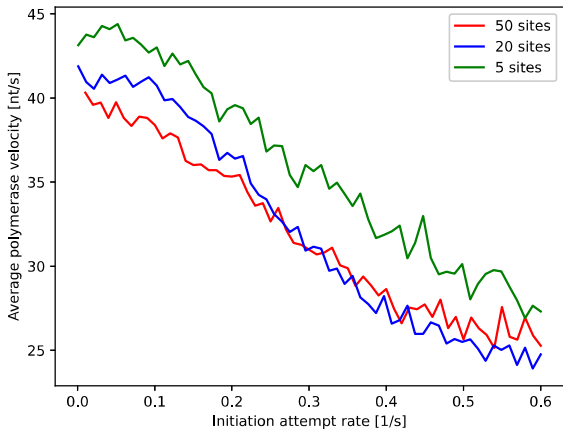
The clumping of RNAps is therefore an inherent property of our stochastic model. Intuitively, clustering makes sense; we expect that particles will catch up to their predecessors if these predecessors are temporarily in a (backtracking) pause. **As such, the RNAps will automatically come closer together and form groups or clusters**. This result confirms previous research as clustering was extensively studied by (van den Berg, 2017).
To verify the clustering with some more direct evidence, we took 100 samples of steady-state simulations at 50% density to measure the nearest-neighbour distances. The average mean of these distances was 65.2 with a standard deviation of 25.4, while the average median was 12.3 with a standard deviation of 20.9. Note that the mean is not a great estimator for clustering, as it also measures the distance between clusters, which may be quite large. However, the median provides a strong indication of clustering occurring, since, on average, half of the nearest neighbour pairs are at most 12.3 nt separated. On the other hand, the high standard deviations imply large variations between the different simulations. In fact, 40% of the simulations had medians of 3 or lower, while there were a couple of outliers with medians larger than 50. We can conclude that congestion and clustering is likely to occur in steady state, but not guaranteed. In figure 5.2, we saw a similar result where the density at high α was not completely stable around the limit.

Recall that one of our main goals is to reproduce the phenomenon of RNAp cooperation obtained by (Klumpp, 2011). An important requirement for this is the clustering of RNAps, so that they are close enough to cooperate with each other to begin with! As such, the simulations of the simple model already reveal the significant phenomenon of clustering. Once we introduce the diffusion barriers, clustering causes a delicate interplay between RNAp interaction and barrier interaction, resulting in the RNAp cooperation.

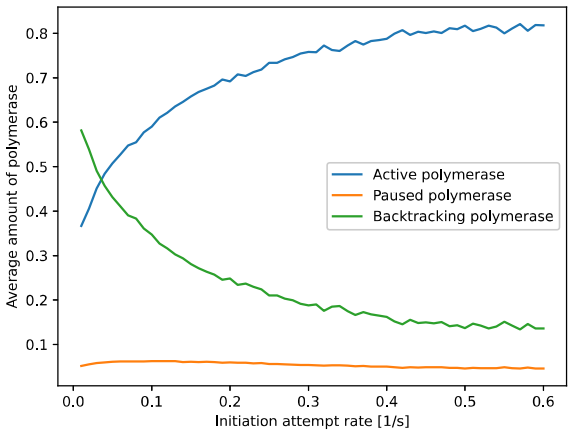
5.3. TRANSCRIPTION DYNAMICS WITH DIFFUSION BARRIER, PAUSE SITES AND CLEAVING

In the simple model, we were already able to establish two important conclusions. Firstly, we saw that the DNA chain has a certain optimal flow, which is also stable. Secondly, we concluded that the RNAps will form clusters. Especially this second point is crucial to be able to observe cooperative interactions between the RNAps. The cooperation effect is one of the main phenomena (Klumpp, 2011) derived; a slight increase in the velocity of RNAps for low densities. This bump stems from the fact that RNAps at low densities were able to push each other forward, by shortening the length of a backtrack, but simultaneously without hindering each other too much in their forward motion.
This bump only occurs with the addition of two components to the basic model. Firstly, recall that the RNAp is precluded from moving past the initiation site, which can be visualised as a barrier at this site. Furthermore, the RNAp has two mechanisms of exiting a backtrack. It can re-enter the pause state at the initiation site, or cleave the transcript and reactivate at any other site. Re-entering the pause state occurs at rate 10s^{-1}, while cleaving occurs at rate 0.05s^{-1}. Hence, the RNAp will exit a backtrack much quicker when it is placed on the initiation site. If we have several RNAps in close proximity, exclusion interaction can force a backtracking RNAp forwards to return to the initiation site. As a result, RNAps can cooperate and push each other out of a long backtrack! This phe-

nomenon is consequently observed as an uptick in the RNAp velocity at low densities. We performed several simulations to reproduce this result obtained by (Klumpp, 2011). We took 60 simulations in the range $\alpha \in [0, 0.6]$ with 3000 RNAp finishes in each.



(a) We compare the steady state velocity in three different systems, each having a different amount of pause sites. The switching rates are corrected for the pause sites, which should lead to equivalent layer presence. The velocity uptick becomes more apparent with fewer pause sites. We attribute this to fewer pause sites being more spread out. As a result, the clusters of RNAps will consist of many active particles stuck behind a backtracking particle, which leads to more pushing. Surprisingly, the velocities are a bit separated between the three systems. We will go into more detail on this result when discussing the transcription rate itself.



(b) This figure depicts the layer presence of the system with 50 pause sites at each value of α. Note that this data has been normalised at each different α. In reality, the amount of particles in the system is also increasing with α.
Despite the fact that we do not observe a velocity bump for the 50-sites system, we still see the phenomenon of shortened backtracks in full effect here. As the density on the chain is increased, the amount of RNAp interactions also increases. We see that as a result of this, the proportion of backtracking RNAps decreases, which is exactly the anticipated cooperation phenomenon.

Figure 5.4: On the RNAp cooperation effect and its dependence on the amount of pause sites.

We came to the conclusion that **the cooperation effect is highly dependent on the distribution of pause sites** on the DNA sequence. For example, in the basic model with every nucleotide functioning as a backtracking site, we found no noticeable uptick in the velocity of the RNAps. In fact, even for a model with 50 sites of rate 10, there is no real bump, although the slope at which the velocity decreases is lower for low α. Therefore, we also moved to models with 20 or even 5 sites, where the switching rates were compensated for the lower amount of pause sites, so that the particle distribution over the layers remained similar. With fewer pause sites, we do observe the hypothesised uptick in velocities. We propose that this phenomenon is therefore highly dependent on the spread between different pausing sites. The farther these sites are removed, the more likely it is that all particles closely behind a backtracking particle are actively transcribing. As a result, these particles are more likely to push the backtracking particle forward. This reasoning is in line with findings of other studies. (Nudler, 2012) remarks: "The probability of backtracking varies dramatically even at adjacent nucleotides, implying that, when leading EC backtracks, trailing EC would most likely be in the active mode, 'pushing' leading EC forward." (EC stands for *E.coli* here). Similarly, (Klumpp, 2011) found that: "Indeed if there are additional backtracking sites upstream of a backtracked RNAP, the trailing RNAP may be stuck there for some time and will therefore need a longer time to catch up with the leading RNAP."

In any case, the velocity uptick was remarkably stronger in the results of (Klumpp, 2011) than in ours. We attribute this disparity to the fact that (Klumpp, 2011) actually differentiates between pause sites and backtracking sites. In our model, we did not consider this distinction, which allows for a smaller amount of backtracking sites while not affecting the total amount of particles in pause state versus in active state. As such, (Klumpp, 2011) also did not introduce rate correction factors.

A second feature that stands out in figure 5.4a, is the fact that the velocities are drastically lower than in figure 5.2. This can also be attributed to the fact that in the basic model, the RNAp could re-enter the pause state with rate $10s^{-1}$ at any site. As a result, the RNAp spends more time in the backtracking layer as compared to the basic model.

Lastly, another aspect of the results that is highly dependent on the pause sites is the actual transcription itself. In figure 5.4a, we noticed that the velocity for the systems with fewer pause sites are consistently higher. In figure 5.5, we see that this is caused by a difference in polymerase densities. The combination of this difference in densities and velocities also causes varying transcription rates, as can be seen in figure 5.5.

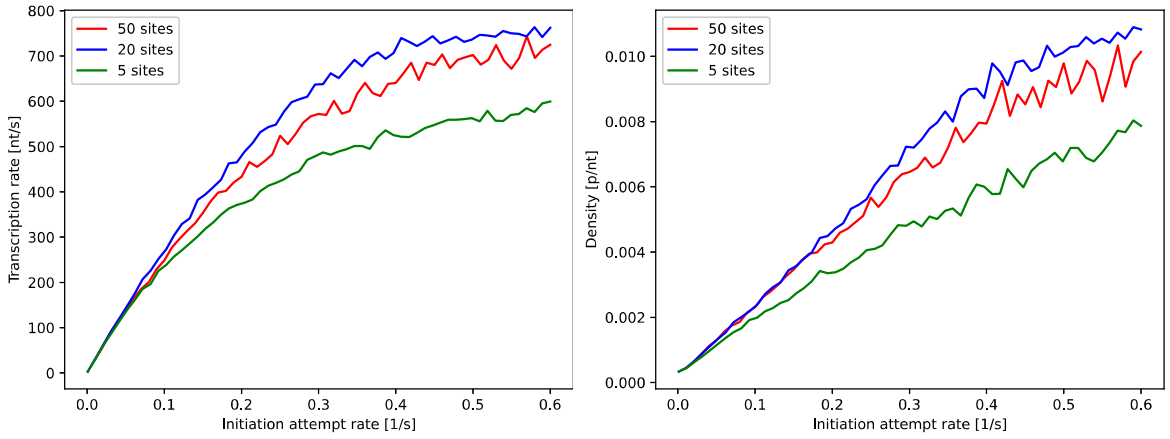


Figure 5.5: Overview of the steady state transcription rates and density for the three systems with different pause sites. These results were obtained with 60 simulations at different α and 3000 RNAp finishes for each simulation. Surprisingly, the density seems to depend non-monotonically on the amount of pause sites. We did not find any plausible explanations for this observation.

We were not able to identify a reason for this behaviour. It is particularly surprising that the inflow for 20 sites is highest, while the inflow for 5 sites is lowest. Therefore, the density does not depend monotonically on the amount of pause sites. Nevertheless, it is already important to establish and to realise that the transcription rate is highly dependent on the amount of pause sites. Additionally, we expect that not only the amount of pause sites, but also the distribution of these pause sites has a large impact on transcription. In our analysis, we only considered evenly spread pause sites, however results may change drastically when considering more concentrated pause site distributions. For instance, if all pause sites are placed at the start of the chain, this greatly affects the local

behaviour of the RNAps in this region. Consequently, the start of the chain will form a "bottleneck" where RNAps are moving forward slowly, with the added consequence that less RNAps can bind to the DNA as well. We therefore expect the actual distribution and locations of the pause sites to have a large influence on the actual transcription rate.

5.4. RELATING THE RESULTS TO THE MATHEMATICAL ANALYSIS AND EXPANDING TO INHOMOGENEOUS ENVIRONMENT

We have just concluded our numerical analysis of the basic and advanced transcription models. Both models are extensions of the general framework of layered Markov processes and exhibit behaviour typical to RNA transcription.
Especially the basic model is closely related to the systems that were analysed mathematically in chapter 3. The only differences being, that the basic model is in finite space and allows for the presence of multiple particles simultaneously. Nevertheless, the mathematical results are still applicable when the system is at a low RNAp density. Specifically the results on the particle distribution over the layers (corollary 19) and the drift formula (theorem 24) are useful theoretical equations for this particular application. The former of the two allows us to estimate the proportion of time the RNAp spends in backtracks. This quantity is biologically important as it relates to the amount of self-correction that is performed by RNAps on the gene. On the other hand, the drift formula provides an estimate for the velocity of RNAps and thus also for the transcription rate at low densities, as can be seen in figure 5.2. At higher densities, the drift formula serves as an upper bound to the actual velocity of RNAps; since the emergence of congestion causes the RNAp velocity to decrease when more particles enter the system. The drift formula can therefore function as a comparison standard for the velocity of particles in a multiple-particle system and is helpful in identifying the causes for deviations from this standard. We were able to identify the density regimes where the RNAps are significantly lower due to congestion, but also regimes in the advanced model where the RNAp velocity was actually higher due to RNAp cooperation. We found out that RNAp cooperation is the result of a delicate interplay between the spread of pause sites, the diffusion barrier and cleaving.

Aside from the drift formula, we also derived several invariance principles for the single-particle systems. These central limit theorems provide estimates for the variation of the particle's motion around the drift velocity. Unfortunately, the many additional complex features and the RNAp interactions are more dominant effects than the diffusive nature of the model. As a result, the RNAps cannot really be compared to Brownian particles, especially at high RNAp densities. Nevertheless, the invariance principles are still highly applicable for other physical applications, such as the run and tumble particles. Additionally, the CLTs and drift formulas and their generalisations to inhomogeneous environments serve as fundamental rigorous evidence that homogenisation takes place. These random environments provide a more realistic depiction of the actual DNA, where the RNAp's motion is position-dependent due to histones, transcription factors and the DNA sequence itself. Since the DNA chain is highly inhomogeneous, verifying whether or not homogenisation occurs is an important question with practical ramifications on the validity of homogeneous models of transcription.

Note that the inhomogeneity of the DNA strand comes in two different forms. Firstly, there is inhomogeneity in the roles of different DNA sites, since only a fraction of the sites act as pause sites. It turned out that this inhomogeneous feature is a particular example of something for which homogenisation does not take place. The spread in pause sites leads to the local effect of RNAp cooperation, which has an influence on the global outcome of the total transcription rate.

The second inhomogeneous aspect of the DNA is the fact that the transition rates between sites vary along the DNA chain. Note that these transition rates can either be for diffusive or for active motion. In our mathematical analysis, we have mostly focused on proving homogenisation for systems with randomly changing diffusive rates. We presented both the Bouchaud trap model and the random conductance model as potential candidates for modelling the RNAp's sequence dependence. Of these two, the random conductance model seemed the most realistic as this allows particular DNA sites to have a certain polarity. Nevertheless, most of our rigorous homogenisation results were found in the Bouchaud trap model. In the trap model, we found that the drift formula and the CLTs fully carried over, thus implying that homogenisation takes place for varying diffusion rates. Since we also expect the active rates of the RNAp to vary with the DNA sequence, a potential extension to our mathematical model would be to consider random active rates and to prove CLTs for these systems as well.

To confirm the expectation that homogenisation also takes place for the random conductance model, we performed several simulations with random environments of diffusive rates. For the random environments, we took the diffusive rates to be uniformly distributed on $[5, 15]$, instead of taking all rates equal to 10s^{-1}. Given a certain random environment, we performed a similar analysis to the one given for the homogeneous environments. However, it turned out that the consequent results on global phenomena such as transcription rates, average velocities and RNAp densities were all almost completely the same as in the case of homogeneous environment. Although the homogeneous and inhomogeneous models are quite different on a local scale, the striking similarity in global results clearly indicates that homogenisation is valid for this specific random conductance model and thus justifies the use of uniform diffusive rates in transcription models.

Chapter 6

Conclusion

In this thesis, we set out to investigate the motion of RNA polymerase (RNAp) during the elongation stage of RNA transcription by using a stochastic model based on Markov processes. In particular, our first goals were to identify and categorise the long-term movements of a single RNAp particle, to determine and to compare the effects of RNAp interactions when considering multiple particles and to relate our findings to the overall rate of transcription under different conditions of the DNA. We conducted both a mathematical and a numerical analysis of the model to derive practical results on single-particle and multi-particle transcription systems respectively.

Additionally, we identified the DNA on which transcription takes place as a highly inhomogeneous medium. Nevertheless, most standard transcription models assume that several aspects of the environment are completely uniform. Intuitively, it is reasonable to expect that small local variations in an inhomogeneous medium eventually average out, thus yielding similar global behaviour to particles in homogeneous environments. This process of "averaging out" is known as homogenisation of the environment. Therefore, our last goal was to validate the standard homogeneity assumptions by verifying that results obtained in homogeneous environments carry over to inhomogeneous environments with both rigorous evidence and evidence gathered from numerical simulations.

For the mathematical analysis, we embedded the specific transcription model into a general class of Markov processes that we call layered jump processes. Aside from transcription, this class has a wide array of applications in several scientific fields, such as molecular motors, electron spin dynamics and even the macroscopic study of the movements of groups of animals and humans.
For these general layered jump processes, we proved that the limiting behaviour of single particles is a linear drift determined by the drift velocity given in theorem 24. This drift equation is useful in the physical applications, as it estimates the speed of particles in these systems. Additionally, we found that the variation around this drift is comparable to Brownian motion and derived closed-form expressions for the diffusion constants in theorems 26, 30 and 31; these results are known as an invariance principles in probability theory. We concluded that single particles in layered jump processes can be seen as propelled Brownian particles that drift in one direction on average.
Furthermore, we were able to extend the invariance principles to the particular case of Bouchaud inhomogeneous environments in theorem 33. This provides initial rigorous evidence of homogenisation. Nevertheless, we did not fully succeed in confirming homogenisation for random conductance environments, bar a couple of specific cases. In particular, it proved to be difficult to combine random conductances with non-zero drift components. However, we did come to the conclusion that the random conductance models portray the inhomogeneity of DNA more accurately than the Bouchaud models. Therefore, these random conductance models may serve as interesting systems for

follow-up research to find whether the lack of homogenisation is actually an inherent property of the system or a symptom of a gap in available mathematical theory.

We complemented the single-particle mathematical analysis by conducting various numerical simulations of multi-particle systems. For these simulations, we implemented the general framework of layered jump processes and applied this specifically to the case of RNA transcription. In these simulations, we found that the Markov processes exhibit several intriguing phenomena which are typical to RNA polymerase and transcription.

First of all, we noticed that the DNA has an optimal RNAp density at which transcription is maximised. Cells can regulate the rate at which RNAp binds to the chain to achieve this optimal density. Therefore, it is particularly useful that our model indicated that the optimal condition is stable. Either having a high enough binding rate or starting from a high RNAp density always resulted in a steady state configuration at optimal RNAp flow.

Secondly, we observed that RNAps start to cluster together at higher RNAp densities. Once the DNA got busier, congestion started to form, thus slowing down the RNAp velocity. Clustering is also a necessary phenomenon for the occurrence of the "cooperation effect" described in the following paragraph.

Previous studies, such as (Klumpp, 2011; Nudler, 2012), observed that RNAps in close proximity to one another may cooperate to increase the rate of transcription. This cooperation causes back-tracking pauses to endure for shorter periods of time, leading to more active transcribing. Using a simple jump process model, we were able to reproduce this observation and we could also explain how it arises. It turned out that this cooperation effect is the result of a delicate interplay between the distribution in pause sites, the interactions between RNAps and the diffusion barrier.
Consequently, the cooperation effect is highly dependent on the configuration of pause sites on the DNA sequence and we inferred that the prevalence of the effect may vary between actual genes. A future research direction could therefore certainly be focused on the actual distribution and location of pause sites on specific genes.

Lastly, since we were unable to derive mathematical invariance principles for the random conductance model, we performed transcription simulations with an inhomogeneous medium of random conductances as well. Our numerical simulations yielded the same global results in both homogeneous and inhomogeneous environments, thus indicating that homogenisation is valid in our transcription model.

To summarise, our stochastic transcription model reproduces many aspects of the movements of RNAps that have been observed experimentally. Hence, we can conclude that the model is realistic and warrants further theoretical analysis, such as a mathematical analysis of multi-particle systems. The general class of layered jump processes has proven itself to be applicable in a wide array of physical systems, such as RNA transcription, and to be suitable for a deep mathematical analysis with striking fundamental results, such as invariance principles and homogenisation.

BIBLIOGRAPHY

Akhmanova, A., & Kapitein, L. C. (2022). Mechanisms of microtubule organization in differentiated animal cells. *Nature reviews. Molecular cell biology*, *23*(8), 541–558. https://doi.org/10.1038/s41580-022-00473-y

Bechinger, C., Di Leonardo, R., Löwen, H., Reichhardt, C., Volpe, G., & Volpe, G. (2016). Active Particles in Complex and Crowded Environments. *Reviews of modern physics*, *88*(4). https://doi.org/10.1103/revmodphys.88.045006

Bhattacharya, R., & Waymire, E. (2023, January). *Semigroup Theory and Markov Processes*. https://doi.org/10.1007/978-3-031-33296-8_2

Biskup, M. (2011). Recent progress on the Random Conductance Model. *Probability surveys*, *8*. https://doi.org/10.1214/11-ps190

Bouchaud, J.-P., & Georges, A. (1990). Anomalous diffusion in disordered media: Statistical mechanisms, models and physical applications. *Physics reports*, *195*(4-5), 127–293. https://doi.org/10.1016/0370-1573(90)90099-n

Chowdhury, D., Santen, L., & Schadschneider, A. (2000). Statistical physics of vehicular traffic and some related systems. *Physics reports*, *329*(4-6), 199–329. https://doi.org/10.1016/s0370-1573(99)00117-9

Demaerel, T., & Maes, C. (2018). Active processes in one dimension. *Physical review. E*, *97*(3). https://doi.org/10.1103/physreve.97.032604

Doyle, P., & Snell, J. (1984, January). *Random Walks and Electric Networks*. https://doi.org/10.5948/upo9781614440222

Einstein, A. (1905). Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik*, *322*(8), 549–560. https://doi.org/10.1002/andp.19053220806

Fodor, É., & Marchetti, M. C. (2018). The statistical physics of active matter: From self-catalytic colloids to living cells. *Physica. A*, *504*, 106–120. https://doi.org/10.1016/j.physa.2017.12.137

Fraleigh, J. B., & Beauregard, R. A. (1995, January). *Linear algebra*. Addison Wesley.

Gillespie, D. T. (2007). Stochastic simulation of chemical kinetics. *Annu Rev Phys Chem*, *58*, 35–55. https://doi.org/10.1146/annurev.physchem.58.032806.104637

Griffiths, D. J., & Schroeter, D. F. (2018, August). *Introduction to quantum mechanics* (Third). Cambridge University Press. https://doi.org/10.1017/9781316995433

Grimmett, G., & Welsh, D. (2014, January). *Probability, an Introduction* (Second). Oxford University Press.

Grimmett, G. R., & Stirzaker, D. R. (2001, May). *Probability and Random Processes* (Third). https://doi.org/10.1093/oso/9780198572237.001.0001

Großmann, R., Peruani, F., & Bär, M. (2016). Diffusion properties of active particles with directional reversal. *New journal of physics*, *18*(4), 043009. https://doi.org/10.1088/1367-2630/18/4/043009

Hewitt, E., & Stromberg, K. R. (1965, January). *Real and Abstract Analysis*. Springer-Verlag. https://doi.org/10.1007/978-3-662-29794-0

Jacod, J., & Protter, P. (2004, January). *Probability essentials* (Second). Springer-Verlag Berlin Heidelberg GmbH. https://doi.org/10.1007/978-3-642-55682-1

Jonkers, I., & Lis, J. T. (2015). Getting up to speed with transcription elongation by RNA polymerase II. *Nature reviews. Molecular cell biology*, *16*(3), 167–177. https://doi.org/10.1038/nrm3953

Jülicher, F., & Bruinsma, R. (1998). Motion of RNA Polymerase along DNA: A Stochastic Model. *Biophysical journal*, *74*(3), 1169–1185. https://doi.org/10.1016/s0006-3495(98)77833-6

Kierzek, A. M. (2002). STOCKS: STOChastic Kinetic Simulations of biochemical systems with Gillespie algorithm. *Bioinformatics*, *18*(3), 470–481. https://doi.org/10.1093/bioinformatics/18.3.470

Kipnis, C., & Varadhan, S. R. S. (1986). Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Communications in Mathematical Physics*, *104*(1), 1–19. https://doi.org/10.1007/bf01210789

Klumpp, S. (2011). Pausing and Backtracking in Transcription Under Dense Traffic Conditions. *Journal of statistical physics*, *142*(6), 1252–1267. https://doi.org/10.1007/s10955-011-0120-3

Klumpp, S., & Hwa, T. (2008). Stochasticity and traffic jams in the transcription of ribosomal RNA: Intriguing role of termination and antitermination. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(47), 18159–18164. https://doi.org/10.1073/pnas.0806084105

Kolasinska-Zwierz, P., Down, T., Latorre, I., Liu, T., Liu, X. S., & Ahringer, J. (2009). Differential chromatin marking of introns and expressed exons by H3K36me3. *Nature genetics*, *41*(3), 376–381. https://doi.org/10.1038/ng.322

Krámli, A., Simányi, N., & Szász, D. (1986). Random walks with internal degrees of freedom. *Probability theory and related fields*, *72*(4), 603–617. https://doi.org/10.1007/bf00344723

Landim, C., Olla, S., & Komorowski, T. (2012). Fluctuations in Markov processes. *Grundlehren der mathematischen Wissenschaften*. https://doi.org/10.1007/978-3-642-29880-6

Li, B., Carey, M., & Workman, J. L. (2007). The Role of Chromatin during Transcription. *Cell*, *128*(4), 707–719. https://doi.org/10.1016/j.cell.2007.01.015

Lowther, G. (2009). *The pathological properties of brownian motion*. Almost Sure Math blog.

Merkenschlager, M., & Odom, D. T. (2013). CTCF and Cohesin: Linking Gene Regulatory Elements with Their Targets. *Cell*, *152*(6), 1285–1297. https://doi.org/10.1016/j.cell.2013.02.029

Mignot, T. (2007). The elusive engine in Myxococcus xanthus gliding motility. *Cellular and molecular life sciences*, *64*(21), 2733–2745. https://doi.org/10.1007/s00018-007-7176-x

Nakata, Y., Ohta, Y., & Wada, Y. (2021, January). *Transcription Dynamics: Cellular Automaton Model of Polymerase Dynamics for Eukaryotes*. https://doi.org/10.1007/978-981-16-7132-6_1

Nudler, E. (2012). RNA Polymerase Backtracking in Gene Regulation and Genome Instability. *Cell*, *149*(7), 1438–1445. https://doi.org/10.1016/j.cell.2012.06.003

Rao, C. V., & Arkin, A. P. (2003). Stochastic chemical kinetics and the quasi-steady-state assumption: Application to the Gillespie algorithm. *The Journal of Chemical Physics*, *118*(11), 4999–5010. https://doi.org/10.1063/1.1545446

Rassoul-Agha, F., & Seppäläinen, T. (2009). Almost sure functional central limit theorem for ballistic random walk in random environment. *Annales de l'Institut Henri Poincaré. B, Probabilités et statistiques/Annales de l'I.H.P. Probabilités et statistiques*, *45*(2). https://doi.org/10.1214/08-aihp167

Saeki, S. (1996). A proof of the existence of infinite product probability measures. *The American Mathematical Monthly*, *103*(8), 2. https://doi.org/10.1080/00029890.1996.12004804

Schwarz, W. (2022, January). *Random Walk and Diffusion Models: An Introduction for Life and Behavioral Scientists*. Springer Nature. https://doi.org/10.1007/978-3-031-12100-5

Sigman, K. (2022). *Continuous-time markov chains*. Columbia University.

Simon, S. H. (2019). *The Oxford Solid State Basics* (First). Oxford University Press.

Sims, R. J., Belotserkovskaya, R., & Reinberg, D. (2004). Elongation by RNA polymerase II: the short and long of it. *Genes and development*, *18*(20), 2437–2468. https://doi.org/10.1101/gad.1235904

Sneppen, K., Dodd, I. B., Shearwin, K. E., Palmer, A. C., Schubert, R. A., Callen, B. P., & Egan, J. B. (2005). A Mathematical Model for Transcriptional Interference by RNA Polymerase Traffic in Escherichia coli. *Journal of Molecular Biology/Journal of molecular biology*, *346*(2), 399–409. https://doi.org/10.1016/j.jmb.2004.11.075

Sznitman, A.-S. (2002). An effective criterion for ballistic behavior of random walks in random environment. *Probability theory and related fields*, *122*(4), 509–544. https://doi.org/10.1007/s004400100177

Tripathi, T., & Chowdhury, D. (2008). Interacting RNA polymerase motors on a DNA track: Effects of traffic congestion and intrinsic noise on RNA synthesis. *Physical review. E, Statistical, nonlinear and soft matter physics*, *77*(1). https://doi.org/10.1103/physreve.77.011921

van den Berg, A. A. (2017). The interplay between polymerase organization and nucleosome occupancy along dna how dynamic roadblocks on the dna induce the formation of rna polymerase pelotons. https://doi.org/10.4233/uuid:993e98ca-3c91-4591-9fbf-26bd6eea2354

van der Post, M. (2024). *Python code for rna transcription random walk model*. Online repository of the Idema group. https://gitlab.tudelft.nl/ig-projects/rna-transcription-random-walk-model

Van Strien, M. (2021). Was physics ever deterministic? The historical basis of determinism and the image of classical physics. *The European physical journal. H*, *46*(1). https://doi.org/10.1140/epjh/s13129-021-00012-x

van Ginkel, B., van Gisbergen, B., & Redig, F. (2021). Run-and-tumble motion: The role of reversibility. *Journal of Statistical Physics*, *183*(3). https://doi.org/10.1007/s10955-021-02787-1

van Kampen, N. (2007, January). *Stochastic Processes in Physics and Chemistry* (Third). Elsevier. https://doi.org/10.1016/b978-0-444-52965-7.x5000-4

Zeitouni, O. (2004, January). *Part II: Random Walks in Random environment*. https://doi.org/10.1007/978-3-540-39874-5_2