2015-07-31

Scaling Limit of a Generalized Pólya Urn Model

Zhe Zhang
University of Miami, zhezhangUM@yahoo.com

UNIVERSITY OF MIAMI

SCALING LIMIT OF A GENERALIZED PÓLYA URN MODEL

By

Zhe Zhang

A DISSERTATION

Submitted to the Faculty
of the University of Miami
in partial fulfillment of the requirements for
the degree of Doctor of Philosophy

Coral Gables, Florida

August 2015

UNIVERSITY OF MIAMI

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

SCALING LIMIT OF A GENERALIZED PÓLYA URN MODEL

Zhe Zhang

Approved:

Ilie Grigorescu, Ph.D. Associate Professor of Mathematics	George C. Cosner, Ph.D. Professor of Mathematics
Victor C. Pestien, Ph.D. Associate Professor of Mathematics	Dean of the Graduate School
Burton J. Rosenberg, Ph.D. Associate Professor of Computer Science	

ZHANG, ZHE (Ph.D., Mathematics)
Scaling Limit of A Generalized (August 2015)
Pólya Urn Model

Abstract of a dissertation at the University of Miami.

Dissertation supervised by Associate Professor Ilie Grigorescu.
No. of pages of text. (120)

We derive a fluid limit for a multi-type urn model, also known as a hydrodynamic limit, in the sense that random trajectories of the microscopic process are shown to follow, as the scaling factor L, proportional with the initial population, approaches infinity, a unique trajectory characterized as the strong global solution of a specific dynamical system (the macroscopic equation). The result is a weak Law of Large Numbers of random variables with values in the Skorokhod space of right continuous with left limits functions with values in $\mathbb{R}^k$, where k is the number of types. We obtain that while the macroscopic process does not vanish in finite time, the microscopic process has a probability of extinction no larger than $O(L^{-1})$. A similar limit is proven for the normalized vector of population proportions, together with qualitative results on its asymptotic behavior. Both limits are in probability, uniformly in time for any fixed time interval. The model and the scaling studied is inspired by earlier work by Schreiber, Benaïm et al. [35, 3] and generalizes the well known *Replicator model* with applications in mathematical ecology and genome population dynamics.

Dedication

For my family,

who always love;

for my friends,

who are supportive;

for everyone,

who understands.

Acknowledgements

I would like to express my respect and gratitude to my advisor, Dr. Ilie Grigorescu, for his guidance and instructions, not only in academic but in life.

I am indebted to him as he is always patient, supportive and encouraging. He devoted enormous time teaching and discussing with me. He is willing to help and gives valuable suggestions. He is a mentor and a friend. I learn from him and I want to say thank you.

I also give my thanks to Dr. Shigui Ruan and Dr. George Cosner, who help me a lot in my graduate studies and whose teachings lead me into a new field that opens eyes and is inspiring. They provide me with feedbacks that encourage and bring strength.

I want to thank Dr. Victor Pestien and Dr. Leticia Oropesa. They are so kind and always there to help. They are caring and give me precious support.

I thank so many people in the Department of Mathematics in University of Miami. I can have such a wonderful place to learn and grow because of you.

Of course I thank my family, I love you too.

Contents

Chapter 1

Introduction and Results

1.1 Historical background

Urn models consist of a rich family of mathematical models which could be used to create mental experiments and explain complex real world phenomena. By putting objects of interest in the context of urns and colored balls, the urn scheme provides a framework to study in a systematic way mathematically common situations in different areas and is proved to successful. It has application in a wide range of fields like natural science, engineering and more, see for example [18] for an introduction to many classical urn models and with applications.

In a wide sense, an urn problem can refer to any mathematical problem that is presented, or can be presented, as an urn scheme and the word urn model may refer to the urn problem, the urn scheme or the mathematical mechanism, notably probability methods, involved to solve the problem. The most recognized application is to random redistribution models of a population of discrete types. Most commonly, an urn model is constructed by considering colored balls in urns together with a set of rules for drawing and the permissible actions to perform, depending on drawing results. Balls can be any objects of interests such as people, atoms, books etc. and different colors represent some characteristic of the objects that one intends to differentiate. Urns can be concrete containers such as boxes, bags which hold the colored balls or serve as abstract concepts like sets, groups or categories. A most fertile example is modeling of genomic types.

In most cases, balls are assumed to be indistinguishable other than their colors (types) and each ball in the same urn is equally likely to be drawn. A set of rules must be specified such as the number of draws to be performed or when to stop, what actions to be performed for different drawing results etc. Some phenomena concerning the evolution of urn models that are of general interest include the (random) distribution of the colored balls starting with a certain state after a number of draws, limiting distribution if it exists (ergodic properties), speed of convergence (exponential, polynomial, and other stability questions) to the equilibrium configuration. Another issue is *non-extinction* of types, as will be seen in the present work in Chapter 8.

The first urn model in history is hard to decide but its appearance in literature can be traced back to as early as the 17th century. For example, the famous gambler's ruin problem, whose name was adapted much later, was considered in correspondences between Pascal and Fermat in 1656 and restated in Huygens' work in latin (1657), see [17] for a mordern version. The original problem was phrased in the language of dice and scores but it is rather easy to be converted to an urn scheme. It was presented as a classical urn problem, for example, in [27]. This is also one of many problems that motivated the early development of probability theory.

In combinatorics, probability and statistics, many basic ideas can be readily understood with the aid of urn models. Urn models can be constructed to derive many important discrete distribution such as (negative) binomial, (negative) hyper-geometric. Moreover, by considering limiting cases, even more distributions, like Poisson, Gaussian, Gamma, Beta distributions can also be obtained. Due to its simple yet extremely flexible nature, urn models gain their popularity and spring up in biology, chemistry, physics, economics, etc.

For example, the classic Ehrenfest model, an urn model named after Paul Ehrenfest, was proposed in early 1900s in an attempt to explain the second law of thermodynamics, with a simple statistical setup. However, it was reported that D. Bernoulli was the first who came up with this model, see [36] .

In 1923, a paper [10] by Eggenberger and Pólya considered an urn scheme modeling the spread of contagious disease which is later known as Pólya-Eggenberger Urn or simply Pólya's Urn. It is constructed as follows. Starting with a single urn containing w white balls and b black balls, a ball is drawn from the urn at random, with equal probability (uniformly). The color of the ball drawn is observed and the ball is returned to the urn, together with s balls of the same color. An urn scheme like this can be conveniently represented in a matrix form

$$A_{2\times 2} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

Here the matrix A is called the replacement matrix and its row index corresponds to the color of the ball being drawn and its column indexes represent the color of the balls to be added. More specifically, if we sampled a ball of color i, we shall check i-th row and add a_{ij} balls of the j-th color to the urn for all j. In this way we may represent the original Pólya's urn scheme as

$$\begin{bmatrix} s & 0 \\ 0 & s \end{bmatrix}. \tag{Polyá's urn}$$

In a mental experiment our urn can have unlimited space and the number of additional balls available is also unlimited. If $s > 0$ this process can be continued forever and it is known, see [9], that the fraction of the white balls in the urn converges almost surely to a random variable which has Beta distribution with parameters $\frac{w}{s}$ and $\frac{b}{s}$, i.e., $Beta(\frac{w}{s}, \frac{b}{s})$. Note that this model had been studied before 1923. For example, Markov,A.A. investigated the case for $s = 1$ in 1906 [29] and in general for $s > 0$ in 1917 [28].

There are many ways to generalize the Pólya's urn scheme. For instance, depending on the draw, s' balls of the opposite color can be added to the urn besides s balls of the same color, which corresponds to the scheme

$$\begin{bmatrix} s & s' \\ s' & s \end{bmatrix}.$$

In fact, this generalization was studied by Bernard Friedman in [14] and was named after him. Also s, s' can be negative integers as well in which case adding s or s' balls shall be interpreted as removing $|s|$ or $|s'|$ balls. We also mention that the Ehrenfest model can be interpreted as an urn of fixed total N with replacement matrix

$$\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}.$$

1.2 Generalizations of the urn model and recent work

The study of the Pólya urn scheme is evolving as more generalized schemes and their respective properties are considered. First note that the entries of the replacement matrix can be pretty arbitrary at least in the nonnegative cases [23]. Also, the 2×2 replacement matrix can be entended to any $k \times k$ matrix, where $k > 0$ is an integer, for example see [2]. More than that, it is possible to extend the Pólya's urn scheme to an infinite scheme, for example in the classical work [6] by Blackwell and MacQueen (1973). Once an urn scheme is formulated, it is very natural to consider it in the context of stochastic processes and there is no reason to restrict us to a deterministic replacement matrix. More complex sampling rules can be considered. For example, a total of k', $2 \leq k' \leq k$ balls may be sampled in each step either with (or without) replacement. If a multi-index α of k' colors has been chosen, the matrix of the urn scheme can be represented by (possibly random) k-dimensional vectors R_α, with components a $R_\alpha^{(i)}$ being the number of individuals of type i to be added.

The *Replicator model* discussed in Section 1.3 is such an example with random matrix and $k' = 2$. Also, even more complicated structure like random trees which are connected to Pólya urn models are studied, see for example [26]. Other multi-type population dynamics are briefly discussed in subsections 1.5.2 and 1.5.3 of Section 1.5 on branching processes.

1.3 The replicator equations and a replicator process

The replicator model is a classical model in the field of evolutionary game theory, for instance in [16], used to study the evolutionary dynamics of a population under the influence of natural selection. It can be set up as follows. Consider a population with individuals adopting various (but finitely many) strategies. For example, on an island there are different species which may compete with each other for resources and individuals of the same type (species) are assumed to take identical strategies. In that sense, types and strategies are, for all practical purposes, identical and we shall call an individual is of type i if it takes strategy i, $1 \leq i \leq k$. When two individuals meet, according to their types, they will replicate (without mutation) or die, which will change the configuration of the population. Now we normalize the population, or equivalently we consider the frequencies of individuals of different types, each of which is programmed to take exactly one of those k different strategies. We will use a column vector $\vec{x} = (x_1, x_2, ..., x_k)^T$ to represent such a state of the population, where x_i is the frequency of individuals of type i, $1 \leq i \leq k$. Clearly, as frequencies, we shall have $x_i \geq 0, 1 \leq i \leq k$ and $\sum_{i=1}^{k} x_i = 1$.

In the following, we will use $|\cdot|$ to represent the L^1 norm for a vector unless stated otherwise. Note that we may also use $|\cdot|$ for a scalar which is just the absolute value and for a set in which case it would be the cardinality function. It should be understood in the appropriate sense depending on the nature of the input.

For a (nonnegative) state vector $\vec{x}$, we have $|\vec{x}| = \sum_{i=1}^{k} |x_i| = \sum_{i=1}^{k} x_i = 1$, it follows that the set of all possible states is Δ^{k-1}, which is the standard $(k-1)$ - dimensional simplex:

$$\Delta^{k-1} = \left\{ \vec{x} = (x_1, ..., x_k)^T \in \mathbb{R}_+^k \middle| |\vec{x}| = 1 \right\}. \tag{1.1}$$

Note that we take a vector to be a column vector by default.

In a general sense, according to the strategy one will take, each type $i \in \{1, 2, ..., k\}$ of individuals would be assigned a so-called fitness function $f_i : \Delta^{k-1} \to \mathbb{R}_+$. And the average fitness function $\bar{f}$ of the whole population would be $\bar{f}(\vec{x}) = \sum_{i=1}^{k} x_i f_i(\vec{x})$.

For a large population, we may consider the dynamical system where the frequencies are smooth enough to be approximated by the *replicator equations*, as presented in [16],

$$\frac{dx_i}{dt} \triangleq \dot{x}_i = x_i(f_i(\vec{x}) - \bar{f}(\vec{x})), \; i = 1, 2, ..., k. \tag{1.2}$$

Note that in the case $x_i = 0$, the population of individuals taking strategy i will be extinct (since we assumed no mutation) and negative x_i does not have biological meaning. If $x_i > 0$

the replicator can be rewritten as

$$\frac{\dot{x}_i}{x_i} = f_i(\vec{x}) - \bar{f}(\vec{x}), \quad i = 1, 2, ..., k, \tag{1.3}$$

and it can be interpreted that the relative growth rate $\frac{\dot{x}_i}{x_i}$ (logarithmic derivative) for the type i individuals is proportional to $f_i(\vec{x}) - \bar{f}(\vec{x})$, the difference between its fitness $f_i(\vec{x})$ and the average fitness of the whole population $\bar{f}(\vec{x})$. This models population growth under natural selection. We can expect that the frequency for species i with better fitness, $f_i(\vec{x}) - \bar{f}(\vec{x}) > 0$, will increase and those that don't fit well, with $f_i(\vec{x}) - \bar{f}(\vec{x}) < 0$, will decrease. Note that the replicator equations don't tell us how the total population, in absolute sense, will grow or diminish. Also note that this equation system is consistent, in the sense that, as frequencies, we need $\sum_{i=1}^{k} x_i = 1$ which leads to $\sum_{i=1}^{k} \dot{x}_i = 0$, and it can be verified by adding up equations (1.2) for $i = 1, ..., k$.

The case of most interest and extensively studied, see for example [16] and references therein, is when the fitness functions are assumed to be linear. More precisely, it is assumed that each fitness function f_i is of the form

$$f_i(\vec{x}) = \sum_{j=1}^{k} a_{ij} x_j, \; i = 1, ..., k, \tag{1.4}$$

or in a vector form

$$f(\vec{x}) = A\vec{x}, \tag{1.5}$$

where $\vec{x} = (x_1, ..., x_k)^T$, $f(\vec{x}) = (f_1(\vec{x}), ..., f_k(\vec{x}))^T$ and

$$A = A_{k\times k} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \dots & a_{kk} \end{bmatrix}.$$

The matrix entries (a_{ij}) can be interpreted as "payoffs": a_{ij} is the factor by which the fitness function of type i is affected when encountering a type j individual. With these notations the replicator equations in the linear case simplify to

$$\dot{x}_i = x_i((A\vec{x})_i - \vec{x}^T A\vec{x}),\ 1 \leq i \leq k. \tag{1.6}$$

Properties concerning equilibrium and stability for the replicator equation are extensively studied and many results are available [16]. But as one may have noticed, for the frequencies functions to be differentiable, we have to assume an infinite population or, in the approximately sense, a really large total population.

In contrast to this view, our interest lies in considering a finite population where randomness comes into play, and its limit phenomena under certain scaling.

The model we will investigate is a replicator process proposed by Schreiber in [35] and a following paper [3], which in a large part motivated our work. Schreiber et al. [35, 3] formulated an urn model which can be considered a discrete model for the replicator dynamics, in a generalized sense. Instead of a deterministic payoff matrix, random payoff, with potential for growth, was studied under a discrete scaling that essentially equates the number of updates per time unit with the size of the current population. Stability results of the resulting dynamical system and some other insightful results were given.

Our approach is different from the original authors though; we will introduce scaling by the standard construction of having independent unit rate exponential clocks for updates carried by each individual. In a natural way, it results that the holding times between updates of the system seen as a whole coincides with exactly the minimum of such an exponential time over all individuals, which is exactly, on the average $|Z|^{-1}$, when the population is $|Z|$. Since we deal with a large population, we shall assume $|Z| \sim L$, where L will be the scaling factor, typically large.

This makes the process under discussion a continuous time pure jump Markov process. We then investigate the scaling limit as $L \to \infty$, as described in more detail in (1.25) in Section 1.6.

The replicator model we will study is based on that of [35, 3] with slight modification. But we comment that although in [35, 3], more than one example was studied and the replicator process is just one of them, our results rely on the assumption they made but not on the process itself. Thus similar results can be obtained for other processes they mentioned. More than that, we found that the results we obtained will hold under an even relaxed assumption.

Note that the replicator model can be put in correspondence with an urn model by interpreting "balls" as individuals and "colors" as types. The discrete version of the replicator process we study can be formulated, in the context of a generalized urn model, as follows.

A single urn is composed of balls with up to $k \geq 2$ different colors and we label the colors with numbers $1 \leq i \leq k$. We say a ball is of color i, if the label for its color is i. We consider the discrete time Markov chain $(Z_n)_{n\geq 0}$, where $Z_n = (Z_n^1, ..., Z_n^k)$ with component Z_n^i, $i = 1, 2, ..., k$, representing the number of balls of color i after the n-th update. Starting with an initial state $Z_0 = a_0 = (a_0^1, ..., a_0^k) \in \mathbb{Z}_+^k$, each time two balls are sampled in order and with replacement. More precisely, we sample a ball first with replacement and then sample the second ball with replacement. Colors of the sample balls are observed and we record their labels in the order they were drawn with the ordered pair (i, j). For the n-th update, if $i = j$, then r_i balls of the same color are added to the urn. If $i \neq j$, then R_{ij} balls of color i and $\widetilde{R}_{ij}$ (contrast to $\widetilde{R}_{ji}$ in the original formulation [35, 3]) balls of color j are added to the urn. For $1 \leq i, j \leq k$, all $r_i, R_{ij}, \widetilde{R}_{ij}$ are independent random variables taking values in $\{-1, 0, 1, ..., m\}$, where $m \geq 2$ is an integer, according to arbitrary but predetermined distributions.

We point out that in this introduction we follow the presentation of that of Schreiber et al [35, 3]. As will be seen, the setup can be substantially generalized and our main results still hold.

For a predetermined integer $m > 1$, the set of all possible jumps w for the replicator model is denoted by J where

$$J = \left\{ w = (w_1, ..., w_k) \in \mathbb{Z}^k : \text{ at most two } w_i \text{ are nonzero}, \; -1 \leqslant w_i \leqslant m, \; i = 1, ..., k \right\} \tag{1.7}$$

and note that the number of elements in J is finite, which we shall denote by $|J|$, where $|J|$ is the cardinality of the set J not the L^1 norm for a vector.

Under our setup, the Markov chain $(Z_n)_{n\geq 0}$ is time homogeneous $t = n \geq 0$ with states $z, z \in \mathbb{Z}_+^k$ and its transition probability functions $\Pi(z, z + w)$ are given as follows.

(i) If $z = 0$, then $\Pi(z, z) = 1$, i.e. the null state is absorbing, the population remains extinct.

(ii) If $z \neq 0$, denote $x_i = \dfrac{z_i}{|z|}$. Then we have

$$\Pi(z, z+w) = x_i x_j P(R_{ij} = w_i) P(\widetilde{R}_{ij} = w_j) + x_j x_i P(R_{ji} = w_j) P(\widetilde{R}_{ji} = w_i), \tag{1.8}$$

$$\text{if } w = (0, \cdots, 0, w_i, 0, \cdots, 0, w_j, 0, \cdots, 0) \quad \text{with } i \neq j, w_i \neq 0 \text{ and } w_j \neq 0;$$

$$\Pi(z, z+w) = \sum_{j \neq i} x_i x_j P(R_{ij} = w_i) P(\widetilde{R}_{ij} = 0) + \sum_{j \neq i} x_j x_i P(R_{ji} = 0) P(\widetilde{R}_{ji} = w_i) + x_i^2 P(r_i = w_i)$$

$$\text{if } w = (0, \cdots, 0, w_i, 0, \cdots, 0) \quad \text{with } w_i \neq 0;$$

and

$$\Pi(z, z+w) = \sum_{j \neq i} x_i x_j P(R_{ij} = 0) P(\widetilde{R}_{ij} = 0) + \sum_{i=1}^{k} x_i^2 P(r_i = 0) \qquad \text{if} \quad w = \vec{0}.$$

1.4 Generalization of the replicator model

We notice that in all cases (1.8), the transition functions $\Pi(z, z+w)$ are quadratic functions in $x = \frac{z}{|z|}$, which we can call $p_w(x)$. It turns out that a more relaxed set of assumptions is sufficient. This is consistent with [3].

Assumption 1.1. *All the possible jumps are uniformly bounded by a positive integer $m' = m(J)$, where J denotes the set of all possible jumps, or precisely,*

$$(A1) \qquad \forall w \in J, \ |w| \leq m' \tag{1.9}$$

This means that we only add or remove a total of at most m' balls at each update.

Assumption 1.2. *There exist Lipschitz maps*

$$\left\{p_w : \Delta^{k-1} \to [0,1] : w \in \mathbb{Z}^k, |w| \leqslant m'\right\}, \tag{1.10}$$

and a real number $a > 0$ such that

$$(A2) \qquad |Z|\,|p_w(Z/|Z|) - \Pi(Z, Z+w)| \leqslant a\,, \qquad \textit{when} \quad Z \neq 0\,, \tag{1.11}$$

for any $w \in \mathbb{Z}^k$.

Remark 1.3. *In equation (1.11), we only required the condition for $Z \neq 0$. For this equation to make sense for $Z = 0$, we can formally take $\frac{Z}{|Z|} = 0$ and extend the domain of p_w to include 0 that:*

$$p_w(0) = 0 \quad \textit{for all} \quad w \in J\,. \tag{1.12}$$

However, $p_w(0) = 0$ is not essential and it can be taken to be any finite number. The point is we want to extend $|Z|p_w(\frac{Z}{|Z|})$ in a continuous way to the case $Z = 0$ (cf. Proposition 6.6).

Remark 1.4. *Note that as $|J|$ is finite, once a set of Lipschitz maps $(p_w)_{w\in J}$ is given, we can conclude there exists a constant C_p for all such functions that*

$$|p_w(X) - p_w(X')| \leqslant C_p|X - X'| \qquad \forall w \in J \quad \textit{and} \quad \forall X, X' \in \triangle^{k-1}. \tag{1.13}$$

For the replicator process model, the jump w_i for each component satisfies $|w_i| \leq m$ by construction and at most two components will change at one update. So $(A1)$ will hold for $m' = 2m$, where m is given in equation (1.7).

Remark 1.5. *The transition probabilities given by (1.8) are bilinear in x_i, x_j on a compact set, which would be Lipschitz and we can simply choose $p_w(Z/|Z|) = \Pi(Z, Z + w)$ in which case (1.11) must hold. Then* (A2) *is satisfied as well.*

Let $L > 0$ be a large number. Condition (A2) (1.11) can be re-written in terms of $x \in \Delta^{k-1}$ and L, with $Lx \in \mathbb{Z}_+^k$,

$$\forall x \in \Delta^{k-1}, \quad \forall w \in J, \quad \forall L > 0 \qquad |p_w(x) - \Pi(Lx, Lx + w)| \leqslant \frac{a}{L}. \tag{1.14}$$

The presence of the factor $L > 0$ suggests a scaling. The total population will be $O(L)$, as well as the number updated per time unit. It is reasonable to consider the time scale of $O(L)$ as well, acting heuristically as if each individual seeks an update with rate one. The minimum waiting time between updates becomes $O(L^{-1})$, as it is well known for the minimum of L independent exponential random variables with intensity one.

As outlined in Section 1.6, we derive a rigorous scaling limit of the k-type model at the full path level (known as *hydrodynamic limit* in the literature [22]) and prove that essentially the extinction is a negligible event.

To put our work in context, the next section gives a brief account of other scaling limits well known in the literature.

1.5 Scaling limits of some related stochastic processes

We are interested in studying some limit phenomena of a certain particle system which could be generated from a generalized Pólya scheme under an appropriate scaling. And it would be beneficial to discuss a few classic discrete time stochastic processes which can be constructed using an urn model, not of the type of Pólya though. In particularly, we discuss how do different scaling work in their cases for illustration and comparison purpose.

1.5.1 Symmetric random walk and Brownian motion

The first and possibly the most renowned example we'd like to mention is the fact that the limit of a symmetric random walk leads to a Brownian motion under a certain scaling. A standard symmetric random walk on the real line is constructed as follows.

Let $(\xi_n)_{n\geq 1}$be a sequence of independently and identically distributed (i.i.d.) random variables $(\Omega, \mathcal{F}, P) \to (\mathbb{R}, \mathbb{B}_{\mathbb{R}})$ such that $P(\xi_1 = 1) = P(\xi_1 = -1) = \dfrac{1}{2}$. Define recursively random variables $S_0 \equiv 0$ and $S_{n+1} = S_n + \xi_{n+1}$, the process $(S_n)_{n\geq 0}$ is called the symmetric (or simple) random walk on the real line. We comment that the random walk model can be easily constructed in the language of urn model if we are allowed to have negative number balls in an urn. Elementary calculations tell us

$$E[\xi_1] = 0, \quad Var[\xi_1] = 1, \quad E[S_0] = 0, \quad Var[S_n] = n.$$

A direct application of the Lindeberg-Lèvy central limit theorem yields:

$$\frac{S_n}{\sqrt{n}} \xrightarrow{\mathrm{D}} N(0, 1).$$

Next we will speed up the process and make smaller jump sizes and generate a new process. We will do it the following way: for a large integer $L > 0$, we define a new process

$$B_L(t) = \frac{S_{\lfloor tL \rfloor}}{\sqrt{L}}$$

where $\lfloor\cdot\rfloor$ is the floor function. Compared with $\frac{S_n}{\sqrt{n}}$, $B_L(t)$ is a continuous time process that, roughly speaking, runs L times as fast but making jumps $\frac{1}{\sqrt{L}}$ the size. Put it another way, let $t' = Lt, \xi'_n = \frac{\xi_n}{\sqrt{L}}$, and $S'_n = \sum_{i=0}^{n} \xi'_i$, we have $B_L(t) = S'_{\lfloor t' \rfloor}$. And as $L \to \infty$ we have

$$B_L(t) = \frac{S_{\lfloor Lt \rfloor}}{\sqrt{\lfloor Lt \rfloor}} \cdot \sqrt{\frac{\lfloor Lt \rfloor}{L}} \xrightarrow{D} N(0, t).$$

It is evident that $B_L(0) = 0$ and $B_L(t)$ has stationary and independent increments by construction. If we let $L \to \infty$ and denote $B(t)$ the weak limit of $B_L(t)$, it turns out that with probability 1 we can choose a continuous path for $B(t)$ which will make $B(t)$ the standard Brownian motion. We shall avoid the technicality of a strict proof here but rather explain why the scaling has to be chosen this way. That is, if we want to find a scaling of time and a scaling of the jumps (space) to make a possible limit process of simple random walk to be of interest as a stochastic process, we have to scale it in a way that an infinitesimal change must satisfy $\Delta t \sim \Delta x^2$ or equivalently $\dfrac{\Delta t}{\Delta x^2} \sim O(1)$.

Consider that we were given a simple random walk in a microscopic level. We determined that a change in the microscopic system would take too long to be noticed in a macroscopic level, so we consider large time scale, say L units at a time, i.e., L unit time in the microscopic level would be 1 time unit in the macroscopic level. Let's say the spatial scaling factor is a function $f(L)$. Now image what will we see in the macroscopic level: for one unit time in macro level, L jumps happen in the micro level and each of size $f(L)$ so its cumulative effect would be $f(L)S_L$. Its variance is $Var[f(L)S_L] = f(L)^2 Var[S_L] = Lf(L)^2$. Note that it is reasonable to require $\lim_{L\to\infty} Var[f(L)S_L] = c > 0$ for some constant c since we may want a possible limit process to be of finite variance in an unit time ($c < +\infty$), yet still possesses some randomness ($c \neq 0$). Otherwise either $c = 0$ which results in a deterministic limit or the case that the variance doesn't exist which means a limit process, if ever exits, would have no variance. This leads to $f(L) \sim O(\frac{1}{\sqrt{L}})$ and if we choose $f(L) = \frac{1}{\sqrt{L}}$ we get $B_L(t)$ as we constructed. Note that in the above argument, we demonstrated formally that if an infinitesimal time change satisfies $L \cdot \Delta t = 1$ or $\Delta t = \frac{1}{L}$, then an infinitesimal space change shall be $\Delta x \sim O(\frac{1}{\sqrt{L}})$ which means $\dfrac{\Delta t}{\Delta x^2} \sim O(1)$. For a discussion of non-symmetric random walk and its connection with diffusion processes, one may refer, for example, [13].

1.5.2 The Galton-Watson process and the Feller diffusion

The Galton-Watson process is one representative of a class of mathematical models used to study population dynamics mostly known as branching processes. The first model, or rather problem, was formulated by Francis Galton. He was unconvinced of the conjecture that aristocratic surnames are more likely to go extinct than common names and decided to study it in a mathematical way. Thus he proposed the "Problem 4001"[15] in the periodic "Educational Times" . However, he only received one solution and it was far from satisfactory to him [20].

At Galton's request, Henry W. Watson approached this problem and developed a useful mathematical device which is later known as (ordinary) generating functions, and which for a sequence $(p_i)_{i\in\mathbb{Z}_+}$ is of the form:

$$f(x) = \sum_{i=0}^{\infty} p_i x^i. \tag{1.15}$$

In probability we usually consider the case where p_i is the probability that a discrete random variable taking values in $\mathbb{Z}_+$ to be i in which case an ordinary generating function is referred as a probability generating function. Watson, with great ingenuity, realized that this problem can be reduced to iteration of functions. More precisely, if we define

$$f_1(x) = f(x), f_{n+1} = f_n(f(x)), \tag{1.16}$$

then the probability that the Galton-Watson process, say $(W_n)_{n\geq 0}$, takes value i is the corresponding coefficient of the x^i term in $f_n(x)$.

However, failed to recognize an extra but relevant root (fix point) in the supercritical case, Watson wrongly concluded that “All the surnames, therefore, tend to extinction in an indefinite time ”[37]. Nevertheless, this process is named after Galton and Watson due to their original work and great contribution. I refer the reader to [20] and [21] for more history remarks.

In modern language, the (single-type) Galton-Watson process can be formulated as a discrete time homogeneous Markov chain and is created as follows. Let $(\xi_{i,j})_{i\geq 0, j\geq 1}$ be a sequence of i.i.d. random variables taking values in $\mathbb{Z}_+ = \{0, 1, 2, \cdots\}$ with the identical probability generating function

$$f(x) = \sum_{i=0}^{\infty} p_i x^i.$$

Further define recursively $W_0 \equiv 1$ and $W_{n+1} = \sum_{i=1}^{W_n} \xi_{n,i}$. The sequence, or rather stochastic process, $(W_n)_{n\geq 0}$ is called the single type Galton-Watson process. It is called a single-type because it assumes that each individual reproduces according to the same distribution and a branching process because its transition probabilities $p(x, \cdot)$ satisfy the *branching property* (see, for example [11] [8])

$$p(x + y, \cdot) = p(x, \cdot) * p(y, \cdot) \tag{1.17}$$

where the operator “ $*$ ” represents convolution.

It's not hard to see that if a process has the branching property, then the law for the process to start at a state $x + y$ would be the same with the law of the sum of two such processes starting with x and y respectively.

We can interpret $\xi_{n,i}$ as the random number of offspring for the i-th individual in the n-th generation and W_n as the total number of population in the n-th generation. It is assumed that $E[\xi_{n,i}] < \infty$ which means that the expected progeny of a single individual is finite.

Note that the Galton-Watson process is possibly the most famous branching process. It was introduced as a discrete time model but it can be made a continuous time Markov process, for example, by using the construction method in Section 2.2. It describes the behavior of a population that individuals reproduce according to the same rule (distribution) and independent of each other. Now suppose W_1 has finite expectation ($\mu > 0$) and variance $\left(\sigma^2 > 0\right)$, i.e.,

$$E[W_1] = \mu, \quad Var[W_1] = \sigma^2,$$

simple calculation yields

$$E[W_n] = \mu^n,$$
$$Var[W_n] = \begin{cases} \dfrac{\sigma^2}{\mu(\mu - 1)}\mu^n(\mu^n - 1), & \text{if } \mu \neq 1; \\ n\sigma^2 \quad , & \text{if } \mu = 1. \end{cases}$$

The cases $0 < \mu < 1$, $\mu = 1$ and $\mu > 1$ are called sub-critical, critical and super-critical respectively. The sub-critical case isn't of much interest as $0 < \mu < 1$ implies $\lim_{n\to+\infty} E[W_n] = 0$ and this population will die out will probability 1. The critical case is somehow delicate and for the purpose of illustration of scaling, we shall only consider the case $\mu > 1$ as Feller did in [12].

In the supercritical case where $\mu > 1$, our population grows exponentially with probability one and for a large population we may approximate it by a continuous time process.

For the limit to be a diffusion process (if it can ever be), we shall try to scale the process in a way that as the time change $\Delta t \to 0$, we have

$$\begin{aligned} \frac{E[\Delta W | W_s = w]}{\Delta t} &\to \alpha(s, w), \\ \frac{E[(\Delta W)^2 | W_s = w]}{\Delta t} &\to 2\beta(s, w), \end{aligned} \tag{1.18}$$

for some parameter functions $\alpha(s, w)$ and $\beta(s, w)$, where $(W_s)_{s \geq 0}$ is the after scaling (continuous time) process. If this condition can be satisfied, for instance, for some (twice continuously differentiable) α and β, then the (after-scaling) limit process would be a diffusion proces and satisfies the Kolmogorov diffusion equations with *drift* $\alpha(s, w)$ and *diffusion coefficient* $\beta(s, w)$.

Now supppose we consider N time unit at a time and the scaling factor for population to be $f(N)$, we define the scaled process to be

$$W_t^{(N)} \triangleq \frac{W_{\lfloor Nt \rfloor}}{f(N)} \tag{1.19}$$

If we choose $f(N) = N$, we see that for an infinitesimal time change $\Delta t = \frac{1}{N}$, we have *infinitesimal mean displacement* and *infinitesimal variance* per particle, respectively

$$E[\Delta W] = \mu - 1 = \frac{\alpha}{N} \quad \text{and} \quad Var[(\Delta W)] = \frac{\sigma^2}{N}.$$

Since we have $\Delta t = \dfrac{1}{N}$, we obtain, again per particle,

$$\begin{aligned}\frac{E[\Delta W]}{\Delta t} &= \mu - 1 = \alpha(s, w),\\ \frac{Var[(\Delta W)]}{\Delta t} &= \sigma^2 = 2\beta(s, w).\end{aligned} \tag{1.20}$$

If the current population is W_t and the coefficients are constants, for simplification, then the mean change and variance from above are multiplied by a factor of W_t, giving

$$E[\Delta W] \simeq \alpha W_t \Delta t \qquad \text{and} \qquad Var[(\Delta W)] \simeq 2\beta W_t \Delta t\,, \tag{1.21}$$

which are exactly the infinitesimal form of (1.23).

These conditions indicate that the proper scaling to obtain a possible diffusion limit is such that $\dfrac{\Delta W}{\Delta t} \sim O(1)$, consistent with the presence of a drift α. In fact, the transition density function of the process $(W_t^{(N)})_{t\geq 0}$ for large N can be well-approximated by the solution $u_{(t,w)}$ to the diffusion equation

$$u_t(t, w) = \beta\{wu(t, w)\}_{ww} - \alpha\{wu(t, w)\}_w, \tag{1.22}$$

which is of the type of the “Fokker-Planck” equation where α, β are such that for a fix large N, $\mu - 1 = \dfrac{\alpha}{N}$ and $\sigma^2 = \dfrac{2\beta}{N}$, see [12]. Such a diffusion process is also called the Feller diffusion such that

$$dW_t = \alpha W_t dt + \sqrt{2\beta W_t} dB_t\,, \tag{1.23}$$

where (B_t) is a standard Brownian motion. We notice that when $\alpha = 0$ the Feller diffusion is a martingale.

1.5.3 The Wright-Fisher model and its diffusive limit

The Wright-Fisher model is proposed by Sewall Wright and Ronald Fisher in an attempt to explain the genetic drift in a diploid population. It can be constructed using the urn model language but we'll set it up in the context of a discrete time Markov chain as follows.

Consider a population of total number of individuals $2N$ and each individual belongs to exactly one of the two types A or a which were assumed to be alleles. Let X_n be a random variable which counts the frequency of individuals of gene type A, then the (neutral) Wright-Fisher model is essentially a time-homogeneous discrete time Markov chain $(X_n)_{n\geq 0}$ which assumes that the transition probabilities is

$$p_{i,j} \stackrel{\Delta}{=} P(X_{n+1} = \frac{j}{2N} \,|\, X_n = \frac{i}{2N}) = \binom{2N}{j}(\frac{i}{2N})^{j}(1 - \frac{i}{2N})^{2N-j}, \tag{1.24}$$

where $i, j \in \{0, 1, 2, \cdots, 2N\}$. Equation (1.24) tells us that conditioning on X_n, $2NX_{n+1}$ follows a binomial distribution with parameter $2N$ and X_n. It follows that

$$E[X_{n+1} \,|\, X_n] = X_n, \quad Var[2NX_{n+1} \,|\, X_n] = 2NX_n(1 - X_n).$$

And the change of the frequency satisfies

$$E[\Delta X|X_n] = 0, \quad E[(\Delta X)^2|X_n] = \frac{1}{2N}X_n(1 - X_n).$$

In view of equation (1.18), we can construct for large N a process $X_t^{(N)} = X_{\lfloor Nt\rfloor}$ which would have infinitesimal drift $\dfrac{E[\Delta X^{(N)}]}{\Delta t} = 0$ and infinitesimal variance $\dfrac{E[(\Delta X^{(N)})^2]}{\Delta t} =$

$\frac{1}{2}X^{(N)}(1-X^{(N)})$. This implies that as $N \to +\infty$, the limit process (X_t) would have

$$E[\Delta X_t | p_t = x] \simeq 0 \qquad E[(\Delta X_t)^2 | p_t = x] \simeq x(1-x)\Delta t\,,$$

and it is a diffusion process satisfying

$$dX_t = \sqrt{X_t(1-X_t)}dB_t\,.$$

The transition density function of such a process can be approximated by, for large N, the solution $p(t,x)$ to the Kolmogorov (backward) equation:

$$\frac{\partial p}{\partial t} = \frac{1}{2}[p(1-p)]_{xx},$$

with initial condition $p(0,x,y) = \delta_y$.

1.6 Main Results

As defined, in Chapter 3, the total mass of the system $|Z_t|$ (the L^1 norm of the vector Z_t) will fluctuate - following a random process. This is not Markovian, because its changes are dependent on the *whole configuration* and not only on the current state. However, we can see heuristically that it must follow an approximation of an exponential growth/decay with noise. It is natural to consider the total initial population as a scaling factor (size of the system), and suppose it is proportional to a large number $L > 0$. The updates of the system are done at exponential times with mean value proportional to the population. That means there is an embedded *time scaling* of a factor of L.

1.6.1 Scaling

We want to relate the *microscopic system* - with population $|Z_t| \sim O(L)$, evolving at speed $t_{micro} \sim O(L)$, with a *macroscopic system* $Z_t^L \sim O(1)$ and $t_{macro} \sim O(1)$, by setting

$$Z_t^L = L^{-1} Z_{Lt}\,, \quad t = t_{macro}\,, \qquad \text{with} \quad L^{-1} Z_0 \simeq z_0 \in \mathbb{R}^k\,. \tag{1.25}$$

Besides the initial condition scaling, we impose the assumption that the transition probabilities for Z_t satisfy a Lipschitz condition at the microscopic level see equation (3.14) , and this condition may possibly be relaxed, as conjectured in (9.2).

Mainly we are concerned with the limit of Z_t^L as $L \to \infty$. This is not a stability question ($t \to \infty$) but a *scaling limit* problem.

1.6.2 Description of results

Following the order they are derived in, the results are built upon result **1**, and then give a refined perspective of the micro- vs macro- levels.

1. Our first result is given by Theorem 3.4 which gives the *Fluid limit*, also known as a *hydrodynamic limit* in the literature for $(\widetilde{Z}_t^L)_{t \geq 0}$. This is a law of large numbers (LLN) at the level of the time-indexed process. We show that the distribution of types (colors) converges, as $L \to \infty$, to a *non-random*, i.e. deterministic, process $z_t \in \mathbb{R}^k$ uniquely characterized as the solution of an ordinary differential equation.

2. *Non-explosion* (Theorem 4.1) shows that both microscopically and macroscopically, the system does not become arbitrarily large, satisfying an exponential upper bound and there are not infinitely many jumps in a finite time interval.

3. *Non-extinction* plays at two levels. Microscopically (for fixed L), extinction is possible, but with a small probability: A concrete large deviations bound is provided, cf. Theorem 8.1. Macroscopically, as $L \to \infty$, the system does not vanish, i.e. the total mass is bounded away from zero at any finite time $t > 0$, see Corollary 8.2.

The non-extinction is crucial for the dynamical system's asymptotic analysis, for instance the results on stability. It is part of the hypothesis in the results of the original work [35]: Theorem 2.2, Corollary 2.4, Theorem 3.1, Corollary 3.2 and further, in [3], Lemma 1, Theorem 1, Theorem 4, Theorem 5, Proposition 1, as well as Theorems 8 and 10.

A significant contribution of this thesis is that, with our methods, we can prove rigorously that in the scaling limit, the conditional stipulation on non-extinction can be removed.

4. *Normalized scaling limit.* Statistically and often biologically, the *proportions of types* are the relevant quantity. We let

$$X_t^L = (\frac{Z_{Lt}^1}{|Z_{Lt}|}), \ldots, \frac{Z_{Lt}^k}{|Z_{Lt}|}) \in \Delta^{k-1}\,, \tag{1.26}$$

where Δ^{k-1} is the $(k-1)$-dimensional simplex, be the normalized population proportions vector. Under the same scaling - the total population is amplified by a factor of L and time is sped up $t \to Lt$ - we show that macroscopically X_t^L converges, again as a process, at the full trajectory level, to the deterministic solution of an ODE.

1.7 Main tools and method of proof

We first prove that the family of random processes $(\widetilde{Z}_t^L)_{t\geq 0}$, indexed by $L > 1$, is tight (compact as measure valued distributions on the space of càdlàg functions with the Skorokhod topology), see Theorem 5.8. Compactness implies there is a limit process. To prove convergence, in this case *weak convergence in probability*, we need to show that the limit is unique and deterministic. In fact, the martingale problem it satisfies is exactly the weak form of an evolution equation which is an ODE here. By analytic methods, it is uniquely determined as a weak solution of a dynamical system with affine bounds on the coefficients, that are also Lipschitz. Classical results in dynamical systems guarantee uniqueness and the existence of a strong global solution. This procedure concludes the identification of the limit.

It is paramount to ensure a global bound on the total population, which will be done in Section 5.2. Technical aspects, like working with a stopping time up to the possible extinction/blowup time are producing significant complications of the proof. However, these conditions can be removed, and that is part of the substance of the result.

This dissertation is organized as follows. In Chapter 2 we introduce notations and lay out the mathematical foundation of probability theory used in our study. Then, in Chapter 3, we set up the generalized urn model inspired by the replicator process and describe in a formal manner the main results. The whole Chapter 4 is reserved to showing the pure jump process we constructed from the replicator model is non-explosive. Chapter 5 proves the main bounds, technical lemmas and the tightness of the scaled processes $(\widetilde{Z}_t^L)_{t\geq 0}$, indexed by L. Chapter 6 is devoted to the proof of the weak limit of $(\widetilde{Z}_t^L)_{t\geq 0}$, identified as the unique solution of a multidimensional ODE; qualitative properties of the deterministic equation are also established. In chapter 7, the normalized process of proportions of types $(\widetilde{X}_t^L)_{t\geq 0}$ is investigated. We prove its tightness and derive its limiting ODE. In Chapter 8, we estimate and obtain a large deviation type bound for the probability that $(\widetilde{Z}_t^L)_{t\geq 0}$ goes to 0 in a finite time. Finally, Chapter 9 gives an outline of directions for possible future research.

Chapter 2

Preliminary Markov Processes Theory

We shall layout some basic notations, definitions and properties of continuous time Markov processes. In particular, we illustrate the construction of a pure jump process. Also we introduce infinitesimal generator and martingales which are necessary to study the generalized urn model. The following definitions and notations are mostly based on [5], [9], [22], [25],[33], and [34].

2.1 Probability spaces, random variables and stochastic processes

Definition 2.1. *Let X be a set. A σ-algebra or σ-field $\mathcal{F}$ on X is a collection of subsets of X satisfying:*

(a) $\emptyset \in \mathcal{F}$,

(b) $A \in \mathcal{F} \implies A^c \in \mathcal{F}$ *and*

(c) *if* $A_1, A_2, \cdots,$ *is a countable collection of elements of* $\mathcal{F}$, *then* $\bigcup_n A_n \in \mathcal{F}$.

Definition 2.2. *A measurable space* $(E, \mathcal{E})$ *is a set* $E \neq \emptyset$ *with a* σ*-algebra* $\mathcal{E}$ *on it.*

Definition 2.3. *A (positive) measure* μ *is a nonnegative countably additive real set function from* $\mathcal{E}$ *to the extended real line. If* $\mu(E) = 1$*,* μ *is called a probability measure and usually denoted by* P*.*

Definition 2.4. *A probability space is a triple* $(\Omega, \mathcal{F}, P)$*, where* $(\Omega, \mathcal{F})$ *is a measurable space with a probability measure* P *on it.*

Definition 2.5. *A filtration* $\{\mathcal{F}_i\}_{i\in I}$ *on a probability space* $(\Omega, \mathcal{F}, P)$*, where* I *is a totally ordered indexed set say under order "* $\prec$ *", is a family of* σ*-algebras that each* $\mathcal{F}_i \subset \mathcal{F}$ *and if* $i \prec j$ *in* I *then* $\mathcal{F}_i \subset \mathcal{F}_j$*.*

Definition 2.6. *A filtered probability space* $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in I}, P)$ *is a probability space* $(\Omega, \mathcal{F}, P)$ *with a filtration* $(\mathcal{F}_t)_{t\in I}$ *on it.*

Remark 2.1. *Usually the index set* I *is* $\mathbb{R}_+ = [0, \infty)$*,* $\mathbb{Z}_+ = \{0, 1, 2, \cdots\}$ *or their subsets and the total order "* $\prec$ *" is simply "* $<$ *". Intuitively we may think the parameter* $i \in I$ *as time and use* $t \in I$*; and* $I = \mathbb{R}_+$ *or* $I = \mathbb{Z}_+$ *represents discrete time or continuous time respectively. In the following, we will consider the continuous time case with the total order "* $<$ *", and we assume that either* $I = \mathbb{R}_+$ *or for a finite time* $T > 0$*,* $I = [0, T]$*.*

Most of the time, it is desirable to have a filtered space which satisfies the so-called *usual conditions*. To introduce that, we need the concept of completeness and right-continuity for a filtered space first.

Definition 2.7. *A probability space* $(\Omega, \mathcal{F}, P)$ *is complete if subsets of all measurable null sets are measurable, or equivalently, if* $A \in \mathcal{F}$ *and* $P(A) = 0$ *, then* $B \subset A$ *implies* $B \in \mathcal{F}$*. A filtered probability space* $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ *is complete if* $(\Omega, \mathcal{F}, P)$ *is complete and* $\mathcal{F}_0$ *contains all the null sets.*

Definition 2.8. *A filtration* $(\mathcal{F}_t)_{t\geq 0}$ *is called right-continuous if* $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s, \forall t \geq 0.$

Definition 2.9 (Usual conditions)**.** *A filtered space* $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ *is said to satisfy the usual conditions if it is complete and the filtration* $(\mathcal{F}_t)_{t\geq 0}$ *is right-continuous.*

Given a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F})_{t\geq 0}, P)$, we can always take the P-augmentation of it and obtain a filtered space which satisfies the usual condition. Also, this augmentation is "reasonable" in the sense that it is minimal and many properties of the original space are preserved in the new filtered space. In fact, the natural filtration of a large family of processes, levy processes for example, once completed, satisfies the usual conditions. One can refer [32] for a more detailed discussion about usual conditions.

A major object we consider in probability is the class of functions called random variables, or more generally random elements, which are simply measurable functions.

Definition 2.10. *Given two measurable spaces* $(E_1, \mathcal{E}_1)$ *and* $(E_2, \mathcal{E}_2)$*, a function* $f : E_1 \to E_2$ *is called measurable if* $\forall E'' \in \mathcal{E}_2$*, we have* $f^{-1}(E'') = E' \in \mathcal{E}_1$.

When specifying the σ-algebras with respect to which the function f is measurable is needed, we say f is $(\mathcal{E}_1/\mathcal{E}_2)$- measurable.

Let S be a locally compact complete separable metric space, for example $\mathbb{R}^d$ with the Euclidean metric, and consider it with its Borel σ - algebra, denoted by (S, Σ). The generality of this space is necessary, since we later on work with the metric space of probability measures, as well as the Skorokhod space of right-continuous left-limit paths, which can be endowed with metric topologies.

Definition 2.11. *A S-valued random element is a measurable map* $X : (\Omega, \mathcal{F}) \to (S, \Sigma)$ *between measurable spaces where* Ω *is called sample space and* S *is called state space.*

Remark 2.2. *In particular, for* $(S,\Sigma) = (\mathbb{R}^d, \mathbb{B}_{\mathbb{R}^d})$*, X is called a random variable if $d = 1$ and a d-dimensional (real-valued) random vector if $d > 1$ is an integer.*

An important type of random variable which we will need is *stopping time*.

Definition 2.12. *A random variable $\tau : \Omega \to [0, +\infty]$ is called a stopping time with respect to a filtration $(\mathcal{F}_t)_{t\geq 0}$ if $\forall t \in [0, +\infty]$, we have $\{\tau \leq t\} \in \mathcal{F}_t$.*

Note that a probability measure P is not required when defining a random element. But when one starts to consider probability concepts like distribution, expectation and more, a probability measure is a must, and we shall always assume a probability space $(\Omega, \mathcal{F}, P)$.

Definition 2.13. *The law, or distribution, for a random element $X : (\Omega, \mathcal{F}, P) \to (S, \Sigma)$ is the probability measure $P_X = P \circ X^{-1}$ on (S, Σ) induced by X, more precisely:*

$$P_X(A) = P(\{\omega : X(\omega) \in A\}), \ \forall A \in \Sigma.$$

Note that the distribution of a random element X is always well-defined since X is by definition measurable which implies $\{\omega : X(\omega) \in A\} \in \mathcal{F}$.

For a random variable we define its mathematical expectation.

Definition 2.14. *The expectation of a random variable $X : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathbb{B}_{\mathbb{R}})$ with respect to a probability measure P on $(\Omega, \mathcal{F})$ is defined by the Lebesgue integral.*

$$E[X] \triangleq \int_\Omega X(\omega) P(d\omega) \equiv \int_{\mathbb{R}} x \, P_X(dx). \tag{2.1}$$

Remark 2.3. *For a random vector, its expectation is defined component-wise as in (2.1).*

Given a sequence of random variables $(X_n)_{n\geq 1}$ and X defined on $(\Omega, \mathcal{F}, P)$ taking values in (S, Σ), we define the following modes of convergence.

Definition 2.15. *The sequence of random variables* $(X_n)_{n\geq 1}$ *is said to converge to* X *almost surely (a.s.), or in probability* 1*, if*

$$P(\{\omega \in \Omega \ : \ X_n(\omega) \to X(\omega)\}) = 1.$$

In this case, we write $X_n \xrightarrow{a.s.} X$.

Definition 2.16. *The sequence of random variables* $(X_n)_{n\geq 1}$ *is said to converge to* X *in probability, if* $\forall \epsilon > 0$*, we have*

$$\lim_{n\to+\infty} P(\{\omega \in \Omega : |X_n(\omega) - X(\omega)| > \epsilon\}) = 1.$$

In this case, we write $X_n \xrightarrow{P} X$.

Definition 2.17. *The sequence of random variables* $(X_n)_{n\geq 1}$ *is said to converge to* X *in distribution, or in law, if* $\forall f \in C_b(S)$*, we have*

$$E[f(X_n)] \to E[f(X)],$$

where $C_b(S)$ *denotes the set of continuous bounded functions* $f : S \to \mathbb{R}$.
In this case, we write $X_n \xrightarrow{D} X$*, or* $X_n \xrightarrow{L} X$.

Note that a random variable X is L^p integrable, denoted by $X \in L^p$, if $E[|X^p|] \leq +\infty$. With the additional assumption that $X, X_1, ..., X_n, ... \in L^p$, we define the L^p convergence.

Definition 2.18. *The sequence of random variables* $(X_n)_{n\geq 1}$ *is said to converge to* X *in* L^p*, denoted by* $X_n \xrightarrow{L^p} X$*, if* $\forall f \in C_b(S)$*, we have*

$$E[f(X_n)] \to E[f(X)],$$

where $C_b(S)$ denotes the set of continuous bounded functions $f : S \to \mathbb{R}$.

Definition 2.19. *Given measurable spaces $(\Omega, \mathcal{F})$ and (S, Σ), a S-valued stochastic process $(X_t)_{t\in I}$ is a collection of S-valued random elements $X_t : (\Omega, \mathcal{F}) \to (S, \Sigma)$ indexed by $t \in I$.*

Remark 2.4. *Given a stochastic process $(X_t)_{t\in I}$, it's very natural to think it as a two variables function $X : I \times \Omega \to S$ and sometimes one writes $X(t, \omega)$ instead of $X_t(\omega)$. We know for each $t \in I$, $X(t, \cdot) : \omega \longmapsto X_t(\omega)$ is a measurable function by definition. If we fix $\omega \in \Omega$, $X(\cdot, \omega) : t \longmapsto X_t(\omega)$ is also a function which is called a sample path at ω. However, we need a topology on the index set I before we can talk about measurability.*

Besides the assumption that $I = [0, \infty)$ or $I = [0, T]$ for a fixed $T > 0$ with the usual topology, we shall also assume that the state space S is $\mathbb{R}^d$, where d is a positive integer, equipped with the Borel σ-algebra $\mathbb{B}_{\mathbb{R}^d}$ generated by open balls from now on. We have several notions concerning measurability of a stochastic process.

Definition 2.20. *Given measurable spaces $(\Omega, \mathcal{F})$ and $(\mathbb{R}^d, \mathbb{B}_{\mathbb{R}^d})$, a stochastic process $(X_t)_{t\geq 0}$ is called* jointly measurable *if the function $X : [0, +\infty)\times\Omega \to \mathbb{R}^d$ is $\left(\mathbb{B}([0, +\infty)) \bigotimes \mathcal{F} / \mathbb{B}_{\mathbb{R}^d}\right)$ measurable where $\mathbb{B}([0, +\infty))$ is the Borel σ-algebra on $[0, +\infty)$ generated by the open sets of $[0, +\infty)$ induced from $\mathbb{R}$ and the operation "$\bigotimes$" represents the product σ-algebra .*

The next two concepts requires a filtered probability space.

Definition 2.21. *Given a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ and state space $(\mathbb{R}^d, \mathbb{B}_{\mathbb{R}^d})$, a stochastic process $(X_t)_{t\geq 0}$ is said to be adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$ if $\forall t \in [0, +\infty)$, the function $X_t : \Omega \to \mathbb{R}^d$ is $(\mathcal{F}_t \,/\, \mathbb{B}_{\mathbb{R}^d})$ measurable.*

A notion which is even stronger than jointly measurable and used more often is progressively measurable.

Definition 2.22. *Assumptions as in Definition 2.21, a stochastic process* $(X_t)_{t\geq 0}$ *is called* progressively measurable *if* $\forall t \in [0,+\infty)$*, the function* $[0,t]\times\Omega \to \mathbb{R}^d$ *given by* $(t,\omega) \longmapsto X_t(\omega)$ *is* $\left(\mathbb{B}([0,t]) \bigotimes \mathcal{F}_t / \mathbb{B}_{\mathbb{R}^d}\right)$ *measurable where* $\mathbb{B}([0,t])$ *is the Borel* σ*-algebra on* $[0,t]$.

Remark 2.5. *Definitions 2.20, 2.21 and 2.22 are closely related. For example, a jointly measurable and adapted process has a progressively measurable version. Also, an adapted process with respect to a filtration* $(\mathcal{F}_t)_{t\geq 0}$ *which has right-continuous or left-continuous sample paths is progressively measurably with respect to the same filtration, cf. [19].*

In particular, a RCLL stochastic process is progressively measurable with respect to its natural filtration (see related definitions below).

Definition 2.23 (RCLL)**.** *A stochastic process* $X : [0,+\infty)\times\Omega \to \mathbb{R}^d$ *is called RCLL, or càdlàg in French, if* $\forall\omega \in \Omega$*,* $X(\cdot,\omega) : t \longmapsto X_t(w)$ *is right-continuous with left limits.*

Note that the RCLL space is not necessarily a probabilistic concept. It is a metric space under various topologies, of which we are specializing only to the Skorokhod J1 topology.

Definition 2.24 (Version)**.** *Two stochastic process* $(X)_{t\geq 0}$ *and* $(X')_{t\geq 0}$ *are which are defined respectively on probability spaces* $(\Omega,\mathcal{F},P)$ *and* $(\Omega',\mathcal{F}',P')$ *with the same state space* (S,Σ) *are said to be versions, or modifications, of each other if they have the same finite dimensional distributions, or equivalently, for any finite sequence* $t_1,t_2,...,t_n$ *with each* $t_i \in [0,+\infty)$ *and measurable sets* $A_1,A_2,...,A_n$ *with each* $A_i \in \Sigma$*, we have*

$$\begin{aligned} &P\left(\{\omega\in\Omega\,|X_{t_1}(\omega)\in A_1, X_{t_2}(\omega)\in A_2,...,X_{t_n}(\omega)\in A_n\}\right) \\ &= P'(\{\omega'\in\Omega'\,|X'_{t_1}(\omega')\in A_1, X'_{t_2}(\omega')\in A_2,...,X'_{t_n}(\omega')\in A_n\}). \end{aligned} \tag{2.2}$$

Definition 2.25. *For a stochastic process* $(X_t)_{t\geq 0}$ *defined on* $(\Omega,\mathcal{F},P)$ *with state space* (S,Σ)*, the natural filtration, denoted by* $(\mathcal{F}_t^X)_{t\geq 0}$*, of* $(X_t)_{t\geq 0}$ *is the minimal filtration with* *respect to which the process is adapted. given by*

$$\mathcal{F}_t^X = \sigma\left\{X_s^{-1}(A)\,|\,0 \le s \le t, A \in \Sigma\right\}, \forall t \in [0,+\infty).$$

There is another view of a stochastic process $(X_t)_{t\geq 0} : I \times \Omega \to S$ which is very useful. As we noted earlier, each $\omega \in \Omega$ determines a sample path $X(\cdot, \omega)$ which is a function from I to S. If we denote S^I the set of all functions from I to S, alternatively we can view the stochastic process as a map $X : \Omega \to S^I$. Indeed, it turns out we can put an "appropriate" σ-algebra on the space S^I which is the smallest σ-algebra making all *coordinate maps*

$$\pi_t : S^I \to S,\ \pi_t(f) = f(t),\ \forall t \in I,\ \forall f \in S^I \tag{2.3}$$

measurable, say denoted by $\Sigma^{\mathbb{I}}$,

$$\Sigma^{\mathbb{I}} \triangleq \sigma\left\{\pi_t : t \in I\right\}, \tag{2.4}$$

in the sense that $X : \Omega \to S^I$ is $(\mathcal{F}/\Sigma^{\mathbb{I}})$- measurable if and only if X_t is $(\mathcal{F}/\Sigma)$- measurable for all $t \in I$. We refer the reader to Chapter II.3 of the book [34] for a detailed discussion of the σ-algebra $\Sigma^{\mathbb{I}}$ on the space S^I. The result is that we can always view a stochastic process $I \times \Omega \to S$ as a random element taking values in S^I, under a proper σ-algebra.

As we noted, a random element always induces a well-defined law (distribution) on the state space, we also have the law for a stochastic process:

Definition 2.26. *Given probability space $(\Omega, \mathcal{F}, P)$ and state space (S, Σ) the law for a stochastic process $X : I \times \Omega \to S$, or equivalently viewed as a random element $X : \Omega \to S^I$, is the probability measure P_X on the space $(S^I, \Sigma^{\mathbb{I}})$ induced by X given by $P_X = P \circ X^{-1}$.*

Remark 2.6. *A stochastic process* $(X_t)_{t\in I} : (\Omega, \mathcal{F}, P) \to (S, \Sigma)$ *always induces a law on* $(S^I, \Sigma^{\mathbb{I}})$ *given by* $P_X = P \circ X^{-1}$*, which is simply the law for it being view as a random element. Now consider the stochastic process of coordinate maps* $(\pi_t)_{t\in I}$ *given by equation (2.3) on the probability space* $(S^I, \Sigma^{\mathbb{I}}, \mu)$*, with the same state space* (S, Σ)*, and it can be viewed as a random element* $\pi : S^I \to S^I$*. Since* $\forall t \in I$ *and* $\forall f \in S^I$*,* $\pi_t(f) = f(t)$*, we know* $\pi(f) = f$ *and* π *is the identity map. Therefore, the law it induces on* S^I *would be* $P_\pi = \mu$*. If* $\mu = P_X$*, the coordinate maps process would has the same law with that of* $(X_t)_{t\in I}$*. In other words,* $(X_t)_{t\in I}$ *on* $(\Omega, \mathcal{F}, P)$ *and* $(\pi)_{t\in I}$ *on* $(S^I, \Sigma^{\mathbb{I}}, P \circ X^{-1})$ *are versions of each other.*

Definition 2.27. *The stochastic process* $(\pi_t)_{t\in I} : S^I \to S$ *defined as above on the probability space* $(S^I, \Sigma^{\mathbb{I}}, P_X)$ *is called the* canonical version *of the process* $(X_t)_{t\in I}$ *on* $(\Omega, \mathcal{F}, P)$*.*

We shall call the space S^I the *canonical space* and the stochastic process $(\pi_t)_{t\in I}$ the *canonical process*. We note that the existence of a canonical process on the canonical space is guaranteed by the Kolmogorov extension theorem (also known as Kolmogorov's consistency theorem, Daniell-Kolmogorov theorem), see [34].

A continuous time Markov process, as will be considered in this work (see definition 2.28), can be constructed on a much more specific sample space Ω, namely a set of paths continuous to the right and with left limits (RCLL, mentioned earlier). The most intuitive idea of such a process is illustrated by pure jump processes. The space of RCLL paths can be endowed with a metric, and becomes separable, complete, and locally compact and is called the Skorokhod space. One is referred to see [4]for more about the Skorokhod space.

In the case of $S = \mathbb{R}^d$, it is denoted by $D([0, \infty), \mathbb{R}^d)$, $I = [0, \infty)$ or $\Omega = D([0, T], \mathbb{R}^d)$, $I = [0, T]$ for a fixed $T > 0$ and $S = \mathbb{R}^d$, we can define a stochastic process $(X_t)_{0\leq t\leq T}$ such that $X(t, w) : I \times D(I, \mathbb{R}^d) \to \mathbb{R}^d$, $X(t, \omega) = w(t)$. Alternatively, it can be viewed as a random element $D(I, \mathbb{R}^d) \to D(I, \mathbb{R}^d)$ where $D(I, \mathbb{R}^d)$ is equipped with the Skorokhod J1 topology induced by the Skorokhod metric defined as follows, e.g., for $I = [0, T]$.

Let Λ_T denote the set of increasing homeomorphisms of $[0, T]$ onto itself, i.e.,

$$\{\lambda(t) : [0, T] \to [0, T] | \lambda(0) = 0, \lambda(T) = T, \lambda(t) \text{ is continuous and strictly increasing}\}. \tag{2.5}$$

For $\eta_n, \eta \in D([0, T], \mathbb{R}^d)$, the Skorokhod space J1 metric $\|\cdot\|_S$ is defined by

$$\|\eta_n - \eta\|_S = \inf_{\lambda \in \Lambda_T} \left\{ \sup_{0 \leqslant t \leqslant T} |\lambda(t) - t| \vee \sup_{0 \leqslant t \leqslant T} \|\eta_n(\lambda(t)) - \eta(t)\|_1 \right\} \tag{2.6}$$

and a sequence $\{\eta_n\} \in D([0, T], \mathbb{R}^d)$ converges to $\eta \in D([0, T], \mathbb{R}^d)$ in J_1 topology if $\exists \{\lambda_n\} \in \Lambda_T$ s.t. $\sup_{t \in [0,T]} |\lambda_n(t) - t| \to 0$ and $\sup_{t \in [0,T]} \|\eta_n(\lambda_n(t)) - \eta(t)\|_1 \to 0$ as $n \to \infty$. Here $\|\cdot\|_1$ denotes the L_1 norm on $\mathbb{R}^d$ and $x \vee y$ means $max\{x, y\}$.

We want to define a homogeneous continuous time Markov process with a countable state space and RCLL paths. We need some preparations. Throughout this paper, we'll assume the following when talking about a continuous Markov process.

In our work, the setup involves a state space S, which is a denumerable metric space, which is the case, as will be seen, for each process indexed by L.

However, our presentation includes the more general case $S = \mathbb{R}^d$, which is not denumerable. This is when we pass L to infinity and where limit process will live in. When necessary, we shall turn to the simplified notations of the discrete state space.

The following definition for a continuous Markov chain can be found in [25] with slight modifications. The sample space can be assumed to be the canonical space $\Omega = \{\omega \,|\, \omega : [0, \infty) \to S \text{ is RCLL}\}$ by construction. Define a stochastic process $X(t, \omega) = \omega(t)$ and for a shift operator $(\theta_s \omega)(t) = \omega(t + s)$. Alternatively, we may write $X(t, \omega)$ as $X_t(\omega)$ or simply X_t without specifying ω. Also we put a σ-algebra $\mathcal{F}$ on Ω such that it is smallest which makes the mapping $\omega \to w(t)$ measurable $\forall t \in [0, \infty)$.

Definition 2.28 (Markov process)**.** *A stochastic process $(X_t)_{t\geq 0}$ defined on a probability space $(\Omega, \mathcal{F}, P)$ is called a Markov process if there is*

(a) a collection of probability measures $\{P^x, x \in S\}$ on Ω, and

(b) a right continuous filtration $(\mathcal{F}_t)_{t\geq 0}$ on $(\Omega, \mathcal{F})$ with respect to which $(X_t)_{t\geq 0}$ is adapted to satisfying

$$P^x(X_0 = x) = 1$$

and for all bounded measurable Y on Ω

$$E^x[Y \circ \theta_s | \mathcal{F}_s] = E^{X_s}[Y] \quad P^x \, a.s., \forall x \in S \tag{2.7}$$

where E^x is the expectation corresponding to the measure P^x that

$$E^x[Y] = \int_\Omega Y \, dP^x. \tag{2.8}$$

Remark 2.7. *Alternatively, if the random variables $X(t, \omega)$ are jointly measurable on the probability space $(\Omega, \mathcal{F}, P)$, then $P^x(\cdot) = P(\cdot \,|\, X_0 = x)$.*

Remark 2.8. *Equation(2.7) is equivalent to the Markov property, see definition (2.29).*

Definition 2.29 (Markov property)**.** *Given a measurable space (S, Σ), a S-valued adapted stochastic process $(X_t)_{t\geq 0}$, with respect to a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P, (P^x)_{x\in S})$ is said to have the Markov property with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$ if, for each $A \in \Sigma$ and $0 \leq s \leq t$, $(\mathcal{F}_t)_{t\geq 0}$, we have*

$$P^x(X_t \in A \,|\, \mathcal{F}_s) = P^x(X_t \in A \,|\, X_s), \forall x \in S. \tag{2.9}$$

Remark 2.9. *Both sides of equation* (2.9) *are conditional probabilities, and* $P^x(\,\cdot\,|X_s)$ *is by definition* $P^x(\,\cdot\,|\,\sigma(X_s))$ *where* $\sigma(X_s)$ *is the* σ*-field generated by* X_s*.*

We will deal with homogeneous Markov processes, which means that their transition probabilities are time homogeneous.

Definition 2.30. *A function* $p : S \times \Sigma \to \mathbb{R}$ *is said to be a transition probability if:*
(a) *for each* $x \in S$, $B \to p(x, B)$ *is a probability measure where* $B \in \Sigma$ *is measurable, and*
(b)*for each* $B \in \Sigma$, $x \to p(x, B)$ *is a measurable function.*
We say that $(X_t)_{t\geq 0}$ *is a Markov process with respect to* $(\mathcal{F}_t)_{t\geq 0}$ *with transition probabilities* $p_{s,t}$ *if*

$$p_{s,t}(x, B) := P(X_t \in B\,|\,X_s = x)$$

where $0 \leq s \leq t$ *and* $B \in \Sigma$ *is measurable. In particular, for* $x, y \in S$*, if* S *is discrete, then we have*

$$p_{s,t}(x, \{y\}) = P(X_t \in \{y\}\,|\,X_s = x).$$

We will simply write y *for* $\{y\}$ *with the understanding as above and* $p_{s,t}(x, y)$ *will stand for* $p_{s,t}(x, \{y\})$:

$$p_{s,t}(x, y) = P(X_t = y\,|\,X_s = x).$$

Definition 2.31. *A Markov process* $(X_t)_{t\geq 0}$ *is called time homogeneous if its transition probabilities are time homogeneous, i.e., there exists a transition probability* $\bar{p}$ *such that*

$$\bar{p}_{t-s}(x, B) = p_{s,t}(x, B).$$

For a homogeneous Markov process, we have $\forall s, t \geq 0$ and $x \in S, \{y\} \in \Sigma$,

$$\bar{p}_t(x, y) = p_{s,s+t}(x, y) = P(X_{s+t} = y \,|\, X_s = x) = P(X_t = y \,|\, X_0 = x) = P^x(X_t = y)\,,$$

which says that the transition probabilities depend only on the positions x, y and the time difference $t = (s + t) - s$, but not on the starting time s. We will only consider time homogeneous Markov processes and use the notation $p_t(x, y)$ instead of $\bar{p}_t(x, y)$ from now on:

$$p_t(x, y) \stackrel{\Delta}{=} P^x(X_t = y) \tag{2.10}$$

Remark 2.10. *Technically, any time-inhomogeneous Markov process $(X_t)_{\geq 0}$ can be made time-homogeneous by considering the so-called space-time Markov chain $(Y_t)_{t\geq 0}$ where $Y_t = (X_t, t)$, this is a good perspective but it is not as useful as one may expect it to be in applications.*

Remark 2.11. *Under quite general conditions, for example a Feller-Dynkin process as defined in [34] III 6, a Markov process can be defined such that, with probability one, the sample paths $t \to X_t(\omega)$ are right continuous with left limits. If the state space S is discrete, one can imagine that the sample path cannot change in a continuous way but rather in the form of "jumps". In fact, this type of Markov process falls into the category of pure jump process.*

Definition 2.32. *A Markov process is called a pure jump process if it has right-continuous sample paths and it is constant between consequent jumps.*

Remark 2.12. *A pure jump process doesn't have to be time-homogeneous or have a countable state space, but it has by definition piecewise right continuous constant paths, which are trivially RCLL.*

Note that the RCLL condition for sample paths of a pure jump process $(X_t)_{t\geq 0}$ can be stated in another way that there is a sequence of increasing stopping times $(T_n)_{n\geq 0}$ such that $T_0 = 0$, X_t is constant on the interval $[T_n, T_{n+1})$ and $X_{T_n-} \neq X_{T_n}$ for all $n \geq 1$. More than that, for a time-homogeneous Markov process, the sojourn time (or holding time) at a state x is proved to be satisfying the memoryless property and thus exponentially distributed. The rate for the exponential distribution depends only on the state x and we will denote it by $\lambda(x)$. Since the stopping times sequence $(T_n)_{n\geq 0}$ is monotone increasing, we know it has an ω-by-ω limit which we shall denote by T_∞ . Usually it is desirable to have only finitely many jumps during any finite time in which case we say the pure jump process to be *regular* or have the *non-explosion* property.

Definition 2.33. *A pure jump process $(X_t)_{t\geq 0}$ is called regular or have the non-explosion property if its jump times $(T_n)_{n\geq 0}$ with point-wise limit $T_\infty \triangleq \lim_{n\to+\infty} T_n$ satisfying*

$$P(T_\infty = +\infty) = 1. \tag{Non-explosion}$$

Otherwise there is a positive probability that we may have a finite limit for the stopping times $P(T_\infty < +\infty) > 0$, and we call the jump process explosive.

With the above discussion and notations, we would like to mention that we have the following relations between a continuous time pure jump process and its embedded discrete chain (2.11).

Proposition 2.13 (Proposition 2.5 in appendix 1 of [22]). (a) *The skeleton chain defined by $\xi_n = X_{T_n}$ for $n \geq 0$ is a discrete time Markov chain with transition probabilities $p(x, B)$ given by*

$$p(x, B) = P(X_{T_1} \in B \mid X_0 = x) = P^x(X_{T_1} \in B). \tag{2.11}$$

(b) *Under P^x, T_1 has an exponential distribution whose parameter is denoted by $\lambda(x)$.*

Conditionally to the sequence $(\xi_n)_{n\geq 0}$, the variables $T_{j+1} - T_j$ are independent and have exponential distributions of parameter $\lambda(\xi_j)$.

(c) *For each continuous time homogeneous Markov chain, the associated transition probability $p(\cdot, \cdot)$ and jump rate $\lambda(\cdot)$ can be defined as in (a) and (b). Conversely, two continuous time homogeneous Markov chains having the same transition probability $p(\cdot, \cdot)$ and bounded jump rate $\lambda(\cdot)$ have the same distribution.*

Remark 2.14. *Proposition 2.13 assumes that the jump rates are bounded. However, the construction we give in the next section is consistent for arbitrary rates, with the possibility of explosion. Later on, in Chapter 3, Theorem 4.1 we will show that the process in this work is non - explosive.*

2.2 Construction of pure jump processes

Up to this point, we started from an existing pure jump process and we defined both the skeleton chain $p(\cdot, \cdot)$ and the associated jump rates $\lambda(\cdot)$.

Note that Proposition 2.13 suggests that in order to construct a time-homogeneous Markov process we only need to know the given transition probabilities and jump rates. We will construct a particular RCLL time-homogeneous Markov process $(X_t)_{t\geq 0}$ as an illustration of a pure jump process. This will also serve as a prototype for the pure jump process we are studying.

First we have the state space (S, Σ) where everything is to be built on. We recall that the space S a locally compact complete separable metric space like $\mathbb{R}^d$, which may be countable or uncountable.

We will suppose that the skeleton chain $(\xi_n)_{n\geq 0}$ as mentioned in Proposition 2.13 has transition probabilities $p(x, B)$ defined as in equation (2.11) and a jump rate parameter function $\lambda : S \to [0, +\infty)$. The chain and a family of exponential random variables, called *holding or waiting times*, are all measurable on a probability space $(\Omega, \mathcal{F}, P)$.

We want to construct a jump process working as follows. Starting at a position $X_0 \in S$, we stay at the same position until an exponential clock which is independent of the process with rate $\lambda(X_0)$ rings, and then we jump to a new random state, with distribution defined according to the prescribed transition probability $p(X_0, \cdot)$. To avoid the case that the clock rings but no real jump occurs, we shall assume $p(x, \{x\}) = 0, \forall x \in S$. This condition is natural, since it is automatically satisfied in Proposition 2.13, part (b).

Set $T_0 \equiv 0$ and denote T_1 the first jump time for our process, or more precisely $T_1 \triangleq inf\{t > T_0 \,|\, X_t \neq X_{T_0}\}$. The process would stay still (be constant) on the interval $[X_{T_0}, X_{T_1})$ until the first jump, i.e.,

$$X_t = X_{T_0},\ t \in [T_0, T_1). \tag{2.12}$$

And the position after the first jump would be determined according to the law

$$P(X_{T_1} \in B) = p(X_{T_0}, B),\ \forall B \in \mathcal{F}.$$

Define the n-th jump time recursively $T_n \triangleq inf\{t > T_{n-1} \,|\, X_t \neq X_{T_{n-1}}\}$. The sequence $(T_n)_{n\geq 0}$ are increasing stopping times adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$ since whether the event $\{T_n \leq t\}$ happens can be determined by observing the history the process $(X_s)_{s\geq 0}$ up to time t and we must have $\{T_n \leq t\} \in \mathcal{F}_t$.

Let's say after the n-th jump, we are at a new position X_{T_n}, and we will stay put until a new exponential clock, independent of the process and former clocks, with rate $\lambda(X_{T_n})$ rings.

Again, it means $(X_t)_{t\geq 0}$ is constant on the interval $[X_{T_n}, X_{T_{n+1}})$ until the next jump or the $(n+1)$-th jump, i.e.,

$$X_t = X_{T_n},\ t \in [T_n, T_{n+1}). \tag{2.13}$$

Also we require that the process jump according to the same time-independent transition probabilities,

$$P(X_{T_{n+1}} \in B) = p(X_{T_n}, B),\quad B \in \mathcal{F},\quad n \geq 0,$$

and this process goes on with the same rules.

Formulated more mathematically, the above described process is defined as follows. First we choose a start point $X_0 \in S$ and set $T_0 \equiv 0$. Let $(\eta_n)_{n\geq 1}$ be an i.i.d. sequence of random variables which are independent of the process $(X_t)_{t\geq 0}$ and have the unit exponential distribution $exp\{1\}$, i.e.,

$$P(\eta_1 \leq t) = \begin{cases} 1 - e^{-t}, & \text{if } t \geq 0; \\ 0\quad , & \text{if } t < 0. \end{cases}$$

We also have a jump rate function $\lambda : S \to [0, +\infty)$ that if for some time $t \in [0, +\infty)$ the process is at a position $X_t \in S$, the holding time before the next jump would have exponential distribution with rate $\lambda(X_t) \in [0, +\infty)$. If for some time t' we have $\lambda(X_{t'}) = 0$, then the process will not jump thereafter with probability 1 and we set $X_t = X_{t'}, t \in [t', +\infty)$. If $\lambda(X_{t'}) \neq 0$, with probability 1 there would be another jump in a finite time and the holding has an exponential distribution with rate $\lambda(X_{t'})$. Define the n-th jump time recursively $T_n \triangleq inf\{t > T_{n-1} \,|\, X_t \neq X_{T_{n-1}}\}$. In other words, conditioned on $X_{T_{n-1}}$, we define that, with probability 1,

$$T_n - T_{n-1} = \begin{cases} +\infty\ , & if \quad \lambda(X_{T_{n-1}}) = 0, \\ \dfrac{\eta_n}{\lambda(X_{T_{n-1}})}, & if \quad \lambda(X_{T_{n-1}}) \neq 0. \end{cases} \tag{2.14}$$

Equivalently, we can set formally $\dfrac{\eta_n}{\lambda(X_{T_{n-1}})} \triangleq +\infty$, if $\lambda(X_{T_{n-1}}) = 0$ and define recursively that $T_n = T_{n-1} + \dfrac{\eta_n}{\lambda(X_{T_{n-1}})}$ for $n \geq 1$. Note that the process doesn't change position between jumps that $X_t = X_{T_{n-1}}$ for $t \in [T_{n-1}, T_n)$ and at a jump time T_n, the next position X_{T_n} is a random variable having transition probabilities given by a prescribed law $p(x, \cdot)$ or more precisely

$$P(X_{T_n} \in B) = p(X_{T_{n-1}}, B). \tag{2.15}$$

Without strictly proving it we note that the above defined process is Markovian since for us to know what to do next, given the information for the current state $X_{t'}$ would be as good as given $(X_t)_{0 \leq t \leq t'}$, which is all the history up to t'. Also the process has RCLL sample paths and is constant between consecutive jumps by construction. All in all, we constructed a pure jump process indeed.

Now let's think about this process and how it works. We know a pure jump process is either regular or explosive (see definition 2.33). Let's say we pick any time $t \in [0, +\infty)$, if the pure jump process is regular, almost surely we have $T_n \leq t < T_{n+1}$ for some $n \geq 0$. Due to the memoryless property of exponential distribution, we know we will have an exponential clock with rate $\lambda(X_{T_n})$ for any time $t \in [T_n, T_{n+1})$. We wait until it rings and then jump to a new position $X_{T_{n+1}} \sim p(X_{T_n}, B)$, $n \geq 0$.

However, if the jump process is explosive, with positive probability we may encounter the case $t > T_\infty$ and we need to answer the question as what do we do next, if we want to define the process for all $t \in [0, +\infty)$. One standard way to fix it is to join an additional state $\Delta \notin S$ which is called a cemetery state. And we extend the pure jump process by defining

$$X_t^\Delta = \begin{cases} X_t, & \text{if} \quad T_n \leq t < T_{n+1}, \\ \Delta, & \text{if} \quad t \geq T_\infty. \end{cases} \tag{2.16}$$

This way, a pure jump process would be well-defined for all time $t \in [0, +\infty)$ whether it is regular or not.

Let us investigate the explosion phenomena more closely. For the type of pure jump process defined as we did, we have $T_0 = 0$ and $T_n = T_{n-1} + \dfrac{\eta_n}{\lambda(X_{T_{n-1}})}$ for $n \geq 1$ which gives

$$T_n = \sum_{i=1}^{n} \frac{\eta_i}{\lambda(X_{T_{i-1}})}, \tag{2.17}$$

and its limit is

$$T_\infty = \sum_{i=1}^{+\infty} \frac{\eta_i}{\lambda(X_{T_{i-1}})} \qquad a.s. \tag{2.18}$$

Note that $(\eta_i)_{i\geq 1}$ are i.i.d. unit exponential random variables, we have

$$E[T_\infty | \mathcal{F}_{T_{k-1}}] = \sum_{i=1}^{k} \frac{1}{\lambda(X_{T_{i-1}})}. \tag{2.19}$$

Then

$$E[T_\infty | \mathcal{F}_\infty] = \sum_{i=1}^{+\infty} \frac{1}{\lambda(X_{T_{i-1}})}. \tag{2.20}$$

where $\mathcal{F}_\infty = \sigma\left(\bigcup_{t\geq 0} \mathcal{F}_t\right) \subset \mathcal{F}$ is the σ-algebra generated by infinite union of all the $\mathcal{F}_t$ in the filtration $(\mathcal{F}_t)_{t\geq 0}$.

In general, the rates $\lambda(X_{T_{i-1}})$ may be dependent on each other as there is possibly a connection between jumps. However, if there is some uniform bound for the jump rate, for example, that $\sup_{x\in S} \lambda(x) < +\infty$, it follows that $P(\, E[T_\infty|\mathcal{F}_\infty] \, = +\infty) = 1$ and the corresponding jump process is regular; this is why many times people simply assume the jump rates to be bounded when discussing properties of pure jump processes.

In our work, the jump rates aren't uniformly bounded. Still, we show that the pure jump process under discussion is regular, see Theorem 4.1.

2.3 Infinitesimal generator and associated martingales

Now consider a continuous time homogeneous Markov process $(X_t)_{t\geq 0}$ with the transition probabilities $p_t(x,y)$ as in equation (2.10) taking values in a denumerable state space S. However, the presentation in this section is valid for a more general locally compact separable metric space S. We shall further assume $(X_t)_{t\geq 0}$ to be a Feller process (see definition (2.34) below). It is noteworthy that any regular pure jump process, including the processes considered in this work, belongs to this class [34, 24].

Denote $C_0(S)$ the collection of continuous functions f on S, $f : S \to \mathbb{R}$, which vanish at infinity. We shall regard S as a Banach space with sup norm

$$\|f\| = \sup_{z\in S} |f(z)|. \tag{2.21}$$

The transition operators $\{T(t), t \geq 0\}$ of a Markov process $(X_t)_{t\geq 0}$ is defined by

$$T(t)f(x) := E^x[f(X_t)], \forall f \in C_0(S). \tag{2.22}$$

Remark 2.15. *In the countable state space case, we have*

$$T(t)f(x) = \sum_{y\in S} p_t(x,y)f(y).$$

Definition 2.34 (Feller process)**.** *A time homogeneous Markov process $(X_t)_{t\geq 0}$ is said to be a strongly continuous Feller process if $\forall f \in C_0(S)$ and $t \geq 0$, we have*

(a) (Feller property) $x \to T(t)f(x) \in C_0(S)$, and

(b) $\lim_{t\downarrow 0} T(t)f(x) = f(x)$.

By the Markov property, these transition probabilities as defined in equation (2.10) satisfy the Chapman-Kolmogorov equations: $\forall x, y \in S$ and $s, t \geq 0$

$$p_{s+t}(x,y) = \sum_{z \in S} p_s(x,z) p_t(z,y). \tag{2.23}$$

It follows that the operators $\{T(t), t \geq 0\}$ form a semigroup that

$$T(s+t) = T(s)T(t), \forall s, t \geq 0. \tag{2.24}$$

Remark 2.16. *Indeed, under our construction, $\{T(t), t \geq 0\}$ is a strongly continuous, positive, contraction semigroup. This guarantees the validity of the Hille-Yosida Theorem.*

Now we introduce the infinitesimal generator for a Markov process $(X_t)_{t \geq 0}$ defined by

$$\mathscr{L} f \triangleq \lim_{t \downarrow 0} \frac{T(t)f - f}{t}, \tag{2.25}$$

and its domain denoted by $\mathcal{D}(\mathscr{L})$ is

$$\mathcal{D}(\mathscr{L}) \triangleq \left\{ f \in C_0(S) : \lim_{t \downarrow 0} \frac{T(t)f - f}{t} \text{ exists} \right\}. \tag{2.26}$$

Remark 2.17. *In particular, if $S = \mathbb{R}^d$, we have $C_c^2(\mathbb{R}^d) \subset \mathcal{D}(\mathscr{L}) \subset C_0(\mathbb{R}^d)$, and $C_c^2(\mathbb{R}^d)$ is dense in $C_0(\mathbb{R}^d)$.*

Next we need the concept of a martingale which is, roughly speaking, a stochastic process model for a fair game.

Definition 2.35 (Martingale)**.** *Given a probability space $(\Omega, \mathcal{F}, P)$ and a state space $(\mathbb{R}^d, \mathbb{B}_{\mathbb{R}^d})$, a stochastic process $X : I \times \Omega \to S$ is called a martingale with respect to a filtration $(\mathcal{F}_t)_{t \in I}$*

on $(\Omega, \mathcal{F}, P)$ *if*

(a) $(X_t)_{t\in I}$ *is adapted to* $(\mathcal{F}_t)_{t\in I}$ *,*

(b) $E[|X_t|] < +\infty$ *for each* $t \in I$*, and*

(c) $E[X_t\,|\,\mathcal{F}_s] = X_s$*,* $\forall s, t \in I$ *and* $s < t$*.*

For a continuous-time homogeneous Markov process (X_t) and a smooth test function $f(\cdot,\cdot)$ (defined precisely in the next Lemma), the processes

$$\mathscr{M}_t^f = f(t, X_t) - f(0, X_0) - \int_0^t [\frac{\partial f}{\partial s}(s, X_s) + \mathscr{L}_x f(s, X_s)]ds \tag{2.27}$$

and

$$\mathscr{N}_t^f = (\mathscr{M}_t^f)^2 - \int_0^t [\mathscr{L}_x f(s, X_s)^2 - 2f(x, X_s)\mathscr{L}_x f(s, X_s)]ds \tag{2.28}$$

where $\mathscr{L}_x$ means the infinitesimal generator is applied to the variable x of the test function $f(t, x)$, will be proven to be martingales.

Lemma 2.18. *For a Markov process* $(X_t)_{t\geq 0} \in D([0,\infty), \mathbb{R}^d)$ *and any test function* $f(t, x) \in C_c^{1,2}([0,\infty) \times \mathbb{R}^d, \mathbb{R})$*, let* $\mathcal{F}_t = \sigma(X_s, 0 \leq s < t)$ *be the natural filtration induced by* $(X_t)_{t\geq 0}$ *, the processes* $\mathscr{M}_t^f$ *and* $\mathscr{N}_t^f$ *as defined in (2.27) and (2.28) are* $\mathcal{F}_t$*-martingales.*

Remark 2.19. *Notice that for a Brownian motion, for example, the martingale (2.27) expresses the Ito formula, where the martingale part is given explicitly as a stochastic integral. Similarly, equation (2.28) would be exactly the so called Ito isometry formula.*

Next we shall obtain the corresponding martingales as defined above for a more concrete Markov process. Note that in the simple case $f(t, x) = f(x)$,we have $\mathscr{L}_x = \mathscr{L}$, and $\dfrac{\partial f}{\partial s} = 0$ and it follows that (2.27) and (2.28) become

$$\mathscr{M}_t^f = f(X_t) - f(X_0) - \int_0^t [\mathscr{L} f(X_s)]ds \tag{2.29}$$

and

$$\mathscr{N}_t^f = (\mathscr{M}_t^f)^2 - \int_0^t [\mathscr{L} f(X_s)^2 - 2f(X_s)\mathscr{L} f(X_s)]ds. \tag{2.30}$$

For a pure jump process as constructed in Section 2.2, it is known that, see for example [22], it can be characterized by specifying its generator

$$(\mathscr{L} f)(x) = \sum_{y \in S} \lambda(x)p(x,y)[f(y) - f(x)], \forall x \in S \tag{2.31}$$

where S, a subset of $\mathbb{R}^d$, is the state space and $f(x) \in C_c^2(\mathbb{R}^d;\mathbb{R})$ is a test function. Note that $f \in C_c^2(\mathbb{R}^d;\mathbb{R})$ implies $f^2 \in C_c^2(\mathbb{R}^d;\mathbb{R})$, we have

$$(\mathscr{L} f^2)(x) = \sum_{y \in S} \lambda(x)p(x,y)[f^2(y) - f^2(x)], \forall x \in S. \tag{2.32}$$

Substituting the results (2.31) and (2.32) into (2.29) and (2.30), after some simplifications, we obtain the differential equations for our pure jump Markov process on the discrete space S, i.e. the two processes below are $(\mathcal{F}_t)$-martingales

$$\mathscr{M}_t^f = f(X_t) - f(X_0) - \int_0^t \lambda(X_{s-}) \sum_{y \in S} p(X_{s-}, y)(f(y) - f(X_{s-}))ds \tag{2.33}$$

and

$$\mathscr{N}_t^f = (\mathscr{M}_t^f)^2 - \int_0^t \lambda(X_{s-}) \sum_{y \in S} p(X_{s-}, y)(f(y) - f(X_{s-}))^2 ds. \tag{2.34}$$

We will use extensively the above-mentioned martingales, together with other techniques like localization and truncation to obtain our main results.

Chapter 3

Setup of the Model

Given a probability space $(\Omega, \mathcal{F}, P)$ and an integer k $\geqslant$ 2, we shall define a stochastic process

$$Z_t(\omega) = \left(Z_t^1(\omega), Z_t^2(\omega), ..., Z_t^k(\omega)\right) \in S$$

modeling the k - type population evolution, where $S = \mathbb{Z}_+^k$.

Given a vector $Z = (Z^1, ..., Z^k) \in \mathbb{R}^k$, we define notations

$$|Z| = \sum_{i=1}^{k} |Z^i| \qquad \text{and} \qquad \alpha(Z) = \sum_{i=1}^{k} Z^i. \tag{3.1}$$

Note that for convenience we will use $|\cdot|$ for the L_1 norm in $\mathbb{R}^k$ instead of $\|\cdot\|_1$, which are more commonly seen, unless otherwise explicitly stated .

3.1 The microscopic process

We make the following assumptions. The set of all possible jumps J is given by:

$$J = \Big\{ w = (w_1, ..., w_k) \in \mathbb{Z}^k : \ -1 \leqslant w_i \leqslant m, \ \ i = 1, ..., k \Big\} . \tag{3.2}$$

We assume there exists a set of deterministic functions $\Pi : \mathbb{Z}^k_+ \times \mathbb{Z}^k_+ \longrightarrow \mathbb{R}_+$ with the property that, for all $z \in S$

$$\begin{aligned} \Pi(z, z') = 0 \quad \text{if} \quad z' - z \notin J \,, \\ \Pi(z, z) = 1 \,, \quad \text{if } z = 0 \quad \text{and} \quad \Pi(z, z) < 1 \quad \text{for all} \quad z \neq 0 \,, \\ \text{and} \quad \sum_{w \in J} \Pi(z, z + w) = 1 \,. \end{aligned} \tag{3.3}$$

Remark 3.1. *This is a generalization of the Replicator model. The requirement that at most two component of a jump w can be nonzero is not relevant, and the lower bound* -1 *for the jump size is obtained by normalization, but again, a finite lower bound would be sufficient. Notice that* $\sum_{w \in J \setminus \{0\}} \Pi(z, z + w) = 1 - \Pi(z, z) > 0$ *for* $z \neq 0$ *according to (3.3) which says that we have no absorbing state other than* 0.

Recall the construction of a pure jump process in (2.31). With that notation, let

$$p(z, z') = \frac{\Pi(z, z')}{1 - \Pi(z, z)} \,, \qquad \lambda(z) = |z|(1 - \Pi(z, z)) \,.$$

Definition 3.2. *The process* $(Z_t)_{t \geq 0}$ *is defined as the pure jump process with coefficients (also known as the coefficients of the* Q *- matrix of the infinitesimal generator)*

$$\lambda(z) p(z, z') = |z| \, \Pi(z, z') \,, \qquad z, z' \in \mathbb{Z}^k_+ \,, \quad z' - z = w \in J \tag{3.4}$$

i.e. with generator

$$\mathscr{L} f(z) = |z| \sum_{w \in J} \Pi(z, z+w)\Big(f(z+w) - f(z)\Big), \qquad f \in C_c^2(\mathbb{R}^k; \mathbb{R}). \tag{3.5}$$

The construction of the process is standard - see our own presentation in Section 2.2 based on the exponential clocks associated to each state. When the jump rates are bounded, see for example Liggett [25, 24], as well as Kipnis-Landim [22] in Appendix 1. It will be shown in Chapter 4 that this process is non-explosive, i.e. there are only finitely many jumps in any finite time interval; that the rates can be taken unbounded, linear, as above, and (3.8) is a proper martingale (it is a local martingale in any case), and finally that the test functions can be extended to smooth functions with linear growth (like projections), in Proposition 5.4.

Additionally, throughout the rest of this thesis, we assume (A1) and (A2) as defined in Assumptions 1.1 and 1.2, together with the notation (1.12), are in force. These are not necessary for the construction of the process, but essential for scaling, as the next section will describe.

Concretely, specializing the general construction of jump processes to our setup, if the process $Z_t = z \in \mathbb{Z}_+^k$ at time $t \geq 0$, then we have an exponential clock attached whose next ring time $\mathcal{T}$ is exponentially distributed with rate $\lambda(t, \omega) = |Z_t(\omega)|$. After a waiting time equal to $\mathcal{T}$, the process is updated to a new state $z' \in \mathbb{Z}_+^k$ with probability $\Pi(Z_t, z')$, as prescribed by (3.4).

As explained in Chapter 2, for any $t \in [0, \infty)$, $Z_t(\omega)$ is a random vector from $(\Omega, \mathcal{F})$ to $(\mathbb{R}^k, \mathbb{B}_{\mathbb{R}^k})$ and for any $\omega \in \Omega$, $Z_t(\omega)$ is a measurable function from $[0, \infty)$ to $\mathbb{Z}_+^k$ where $\mathbb{Z}_+^k$ is the positive cone $\mathbb{Z}_+^k = \{Z = (Z^1, ..., Z^k) \in \mathbb{Z} : Z^i \geqslant 0, i = 1, .., k\}$. Further we can define $\{\mathcal{F}_t\}_{t \geqslant 0}$ to be a filtration of $\mathcal{F}$ that $Z_t(\omega)$ is adapted to. In particular, $\mathcal{F}_t$ can be the natural filtration induced by this process $Z = \{Z(t) : t \geqslant 0\}$, i.e.

$$\forall\, t \geqslant 0\,, \qquad \mathcal{F}_t = \sigma\left(\left\{Z_s^{-1}(B) : 0 \leqslant s \leqslant t, B \in \mathcal{R}^k\right\}\right)\,.$$

Then for each $t \in [0, \infty)$, Z_t is $\mathcal{F}_t$-measurable.

It is always possible to augment the natural filtration to a filtration that satisfies the *usual conditions*. Henceforth we shall assume this is true throughout the paper.

The transition probabilities are time homogeneous and the resulting stochastic process $(Z_t)_{t\geq 0}$ has RCLL paths, meaning that its paths

$$\forall \omega \in \Omega \qquad t \longrightarrow Z_t(\omega) \in D([0, \infty), \mathbb{R}_+^k)\,, \tag{3.6}$$

where $D([0, \infty), \mathbb{R}_+^k)$ is the Skorokhod space of *right-continuous with left-limits* space in $\mathbb{R}^k$. Due to our construction, it turns out that the paths stay positive. The *cemetery state* is simply $0 \in \mathbb{Z}_+^k$ and we shall write

$$\tau_0(\omega) \stackrel{\Delta}{=} \inf\{t > 0 \,|\, Z_t(\omega) = 0\} \tag{3.7}$$

for the extinction time, which is a stopping time.

For a test function $f(\cdot, \cdot) \in C_c^{1,2}([0, \infty) \times \mathbb{R}^k; \mathbb{R})$, we have

$$f(t, X_t) - f(0, X_0) - \int_0^t \frac{\partial f}{\partial s}(s, X_s) + \mathscr{L} f(s, X_s) ds = \mathscr{M}_t^f \tag{3.8}$$

where $\mathscr{M}_t^f$ is a $\mathcal{F}_t$ - martingale and $\mathscr{L} = \mathscr{L}_x$ (the process is time - homogeneous and the generator acts on the space variable only) is the infinitesimal generator, defined in (2.25), corresponding to the pure jump process $(X_t)_{t\geq 0}$.

Remark 3.3. *The time - integrals in the martingale formulas are pathwise (in ω) Riemann - Stieltjes integrals. Since the jumps occur from $Z_{\tau-} \to Z_\tau$, where τ is a jump time, the integrands would formally carry the left-limit marker, for example $\mathscr{L} f(s, X_s-)$. However, Riemann-Stieltjes integrals coincide if the integrator function, here given by the Lebesgue measure on the time line ds, is continuous and we can drop the left limit symbol.*

In the case $f(t, x) = f(x)$, applying relation (3.8) to the process $(Z_t)_{t\geqslant 0}$, we have

$$f(Z_t) - f(Z_0) - \int_0^t \mathscr{L} f(Z_s)ds = \mathscr{M}_t^f. \tag{3.9}$$

Under our set-up, $(Z_t)_{t\geqslant 0}$ is a pure jump process, and for such a process we have

$$\mathscr{L} f(Z_t) = \lambda(Z_t) \sum_{w\in J} \Pi\,(Z_t, Z_t + w)\,(f(Z_t + w) - f(Z_t))\,, \tag{3.10}$$

where $\lambda(\cdot)$ is the jump rate function given by $\lambda(Z_t) = |Z_t|$. Plugging (3.10) into (3.9)we obtain

$$f(Z_t) - f(Z_0) - \int_0^t |Z_s| \sum_{w\in J} \Pi\,(Z_s, Z_s + w)\,[f(Z_s + w) - f(Z_s)]\,ds, \tag{3.11}$$

which is a martingale and we shall denote it by $\mathscr{M}_t^f$. Note that the summation in (3.11) is over all possible jumps w which we denoted by J.

3.2 Scaling and macroscopic equations

Our assumptions are made as in Assumption 1.1 and Assumption 1.2 with the understanding of the case $Z = 0$ as in Remark 1.3. We assume that there exist Lipschitz maps

$$\left\{p_w : \triangle^{k-1} \to [0,1] : w \in \mathbb{Z}^k, |w| \leqslant m\right\} \tag{3.12}$$

where $\triangle^{k-1}$ is the $k-1$ simplex:

$$\triangle^{k-1} = \left\{(x_1, ..., x_k) \in \mathbb{R}^k \middle| \sum_{i=1}^{k} x_i = 1, x_i \geqslant 0, \forall i \in \{1, ..., k\}\right\} \tag{3.13}$$

and a real number $a > 0$ s.t.

$$|Z|\,|p_w\,(Z/|Z|) - \Pi(Z, Z+w)| \leqslant a \tag{3.14}$$

for nonzero $Z \in \mathbb{Z}_+^k$ and $w \in \mathbb{Z}^k$ with $|w| \leqslant m$. Note that as $|J|$ is finite, once a set of Lipschitz maps p_w is given, we can conclude there exists a constant C_p for all such functions that

$$|p_w(X) - p_w(X')| \leqslant C_p|X - X'| \qquad \forall w \in J \quad \text{and} \quad \forall X, X' \in \triangle^{k-1}. \tag{3.15}$$

Suppose we can use the projection map $f(Z) = Z^i$ which requires justification and will be given later (see lemma 5.4) since $f \notin C_c^2(\mathbb{R}^k, \mathbb{R})$, we have that

$$Z_t^i - Z_0^i - \int_0^t |Z_s| \sum_{w \in J} \Pi\,(Z_s, Z_s + w)\, w_i ds, \tag{3.16}$$

is a martingale which we shall denote by $\mathscr{M}_t^i$.

Then we know from (3.16)

$$\begin{aligned} Z_t^i - Z_0^i - \int_0^t |Z_s| \sum_{w\in J} \left(\Pi\left(Z_s, Z_s + w\right) - p_w\left(\frac{Z_s}{|Z_s|}\right)\right) w_i ds \\ - \int_0^t |Z_s| \sum_{w\in J} p_w\left(\frac{Z_s}{|Z_s|}\right) w_i ds = \mathscr{M}_t^i. \end{aligned} \tag{3.17}$$

We want to investigate the case where Z_0 is large and the frequency of the initial distribution is fixed. In order to do that, let $L \geq 1$ be the scaling factor and pick a fixed $a_0 = (a_0^1, a_0^2, ..., a_0^k) \in \mathbb{R}_+^k \setminus \{0\}$, we define a process $Z_t^L = (Z_t^{L,1}, ..., Z_t^{L,k})$ depending on L. The update mechanism of Z_t^L is the same as that of Z_t, we only impose a initial condition that

$$Z_0^L = \left(\left\lfloor La_0^1 \right\rfloor, \left\lfloor La_0^2 \right\rfloor, ..., \left\lfloor La_0^k \right\rfloor\right). \tag{3.18}$$

Note that Z_0^L is a random vector and the initial condition should be understood to be true with probability one. Or we may simply assume it is true $\forall\omega$ which guarantees that $\frac{Z_0^L}{L} \to a_0$ as $L \to \infty$. Note that Z_t^L is a random vector and L is a scaling parameter which we will use throughout this paper, while when we write Z_t^i for $i = 0, .., k$, it is the i-th component of the random vector of Z_t. As for each Z_t^i, we have for each $Z_t^{L,i}, i = 1, 2, ..., k$ that

$$\begin{aligned} Z_t^{L,i} - Z_0^{L,i} = \int_0^t |Z_{s-}^L| \sum_{w\in J} \left(\Pi(Z_{s-}^L, Z_{s-}^L + w) - p_w\left(\frac{Z_{s-}^L}{|Z_{s-}^L|}\right)\right) w_i ds \\ + \int_0^t |Z_{s-}^L| \sum_{w\in J} p_w\left(\frac{Z_{s-}^L}{|Z_{s-}^L|}\right) w_i ds + \mathscr{M}_t^{L,i}, \end{aligned} \tag{3.19}$$

where $\mathscr{M}_t^{L,i}$ is a martingale. Define notations

$$\widetilde{Z}_t^{L,i} = \frac{Z_t^{L,i}}{L}, \ \widetilde{Z}_t^L = \frac{Z_t^L}{L} \quad \text{and} \quad \widetilde{\mathscr{M}}_t^{L,i} = \frac{\mathscr{M}_t^{L,i}}{L}; \tag{3.20}$$

dividing both sides of (3.19) by L and notice that

$$\frac{\widetilde{Z}^L_{s-}}{|\widetilde{Z}^L_{s-}|} = \frac{Z^L_{s-}}{|Z^L_{s-}|},$$

we obtain

$$\begin{aligned}\widetilde{Z}^{L,i}_t - \widetilde{Z}^{L,i}_0 = &\int_0^t |\widetilde{Z}^L_{s-}| \sum_{w\in J} \left(\Pi(L\widetilde{Z}^L_{s-}, L\widetilde{Z}^L_{s-} + w) - p_w\left(\frac{\widetilde{Z}^L_{s-}}{|\widetilde{Z}^L_{s-}|}\right)\right) w_i ds \\ &+ \int_0^t |\widetilde{Z}^L_{s-}| \sum_{w\in J} p_w\left(\frac{\widetilde{Z}^L_{s-}}{|\widetilde{Z}^L_{s-}|}\right) w_i ds + \widetilde{\mathscr{M}}^{L,i}_t.\end{aligned} \tag{3.21}$$

Note that for each fixed L, $\widetilde{\mathscr{M}}^{L,i}_t$ is still a martingale. By (3.14), we have that

$$\left|\Pi(Z^L_{s-}, Z^L_{s-} + w) - p_w\left(\frac{Z^L_{s-}}{|Z^L_{s-}|}\right)\right| \leqslant \frac{a}{|Z^L_{s-}|},$$

and thus

$$\left|\Pi(L\widetilde{Z}^L_{s-}, L\widetilde{Z}^L_{s-} + w) - p_w\left(\frac{\widetilde{Z}^L_{s-}}{|\widetilde{Z}^L_{s-}|}\right)\right| \leqslant \frac{a}{L|\widetilde{Z}^L_{s-}|}. \tag{3.22}$$

Therefore, using the fact that for all $w \in J$, $|w_i| \leqslant m$, we have that

$$\left|\int_0^t |\widetilde{Z}^L_{s-}| \sum_{w\in J} \left(\Pi(L\widetilde{Z}^L_{s-}, L\widetilde{Z}^L_{s-} + w) - p_w\left(\frac{\widetilde{Z}^L_{s-}}{|\widetilde{Z}^L_{s-}|}\right)\right) w_i ds\right| \leqslant \frac{am}{L} t. \tag{3.23}$$

Suppose the limit process of $\widetilde{Z}^{L,i}_t$ exists and is unique, which we shall denote by $\widetilde{Z}^i_t$, also note that the martingale part $\widetilde{\mathscr{M}}^{L,i}_t$ will fade away as $L \to \infty$, which we will show later, we can see from (3.23) that as $L \to \infty$, (3.21) becomes

$$\begin{aligned}\lim_{L\to\infty} \widetilde{Z}^{L,i}_t - \widetilde{Z}^{L,i}_0 &= \widetilde{Z}^i_t - \widetilde{Z}^i_0 \\ &= \lim_{L\to\infty} \int_0^t |\widetilde{Z}^L_{s-}| \sum_{w\in J} \left(\Pi(L\widetilde{Z}^L_{s-}, L\widetilde{Z}^L_{s-} + w) - p_w \left(\frac{\widetilde{Z}^L_{s-}}{|\widetilde{Z}^L_{s-}|} \right) \right) w_i ds \\ &+ \lim_{L\to\infty} \int_0^t |\widetilde{Z}^L_{s-}| \sum_{w\in J} p_w \left(\frac{\widetilde{Z}^L_{s-}}{|\widetilde{Z}^L_{s-}|} \right) w_i ds + \lim_{L\to\infty} \mathscr{M}^{L,i}_t \\ &= \lim_{L\to\infty} \int_0^t |\widetilde{Z}^L_{s-}| \sum_{w\in J} p_w \left(\frac{\widetilde{Z}^L_{s-}}{|\widetilde{Z}^L_{s-}|} \right) w_i ds.\end{aligned} \tag{3.24}$$

Suppose that we can change the order of limits and integration in (3.24), we will have

$$\widetilde{Z}^i_t - \widetilde{Z}^i_0 = \int_0^t |\widetilde{Z}_{s-}| \sum_{w\in J} p_w \left(\frac{\widetilde{Z}_{s-}}{|\widetilde{Z}_{s-}|} \right) w_i ds, \tag{3.25}$$

or

$$\frac{d\widetilde{Z}^i_t}{dt} = |\widetilde{Z}_t| \sum_{w} p_w \left(\frac{\widetilde{Z}_t}{|\widetilde{Z}_t|} \right) w_i, \tag{3.26}$$

assuming we can differentiate both sides because, as we will show later, the terms under integral is continuous. If by adding a certain initial condition, we can gurantee that $|\widetilde{Z}_t| > 0$ and $\frac{\widetilde{Z}_t}{|\widetilde{Z}_t|}$ makes sense in any finite time interval, and at the same time it maintains that $\widetilde{Z}^i_t \geq 0$, for all $i = 1, ..., k$, we can add all those equations up and get

$$\frac{d|\widetilde{Z}_t|}{dt} = |\widetilde{Z}_t| \sum_{w\in J} p_w \left(\frac{\widetilde{Z}_t}{|\widetilde{Z}_t|} \right) \alpha(w). \tag{3.27}$$

where $|\widetilde{Z}_t| = \sum_{i=1}^{k} |\widetilde{Z}^i_t|$ and $\alpha(w) = \sum_{i=1}^{k} w_i$ as defined in (3.1).

After completing the details of the proof, which are the object of Chapters 6 and 7, we obtain one of our main results.

Theorem 3.4. *The family of processes* $(\widetilde{Z}_t^L)_{t\geqslant 0}$ *from (3.20), defined on the RCLL paths space* $D([0,\infty),\mathbb{R}^k)$*, indexed by* $L > 0$ *, with initial states equal to* $\widetilde{Z}_0^L = L^{-1}\left([La_0^1],[La_0^2],...,[La_0^k]\right)$ *for a fixed* $a_0 = (a_0^1, a_0^2, ..., a_0^k) \in \mathbb{R}_+^k$*, is tight and converges in probability, uniformly in time, on any bounded time interval, to the deterministic process denoted by* $(\widetilde{Z}_t)_{t\geqslant 0}$*, which is uniquely characterized as the strong solution of the initial value problem*

$$\frac{d\widetilde{Z}_t}{dt} = |\widetilde{Z}_t| \sum_{w\in J} p_w \left(\frac{\widetilde{Z}_t}{|\widetilde{Z}_t|}\right) w\,, \qquad \widetilde{Z}_0 = a_0\,. \tag{3.28}$$

Proof. Theorem 5.8 guarantees the tightness of the process $(\widetilde{Z}_t^L)_{t\geqslant 0}$ and the time continuity of any limit process. The initial value problem has a unique strong global solution as shown in Proposition 6.6. Proposition 6.11 says that a weak limit of the process $(\widetilde{Z}_t^L)_{t\geqslant 0}$ is the solution to (3.28) and, by uniqueness, it must be identical to the unique strong solution of the ODE. Lemma 7.4 further proves that the convergence is in probability and uniform in time on any bounded time interval. □

Remark 3.5. *As will be seen in Proposition 6.6, and also discussed in Remark 1.3, the right-hand side of (3.28), namely* $|\widetilde{Z}| \sum_{w\in J} p_w\left(\frac{\widetilde{Z}}{|\widetilde{Z}|}\right)w$*, is extended continuously to* $\widetilde{Z} = 0$*, in which case the ODE becomes* $\frac{d\widetilde{Z}_t}{dt} = 0$*. In this form, we have a uniformly Lipschitz, affine (sublinenar) autonomous ODE, which will have a unique, global, strong solution.*

In fact, more is true.

Theorem 3.6. *If we start from a point* $a_0 \in \mathbb{R}_+^k \setminus \{0\}$*, then the process* $(\widetilde{Z}_t^L)_{t\geqslant 0}$ *has probability of extinction of order* $O(L^{-1})$ *and the deterministic limiting process* $(\widetilde{Z}_t)_{t\geqslant 0}$ *does not vanish in finite time.*

Proof. The proof is given in Lemma 7.3 and Corollary 8.2. □

We now move on to the process of population proportions.

Denote $\dfrac{\widetilde{Z}^i_t}{|\widetilde{Z}^i_t|}$ and $\dfrac{\widetilde{Z}_t}{|\widetilde{Z}_t|}$ by $\widetilde{X}^i_t$ and $\widetilde{X}_t$ respectively, we have that

$$\begin{aligned}\frac{d\widetilde{X}^i_t}{dt} &= \frac{1}{|\widetilde{Z}_t|}\cdot\frac{d\widetilde{Z}^i_t}{dt} - \frac{\widetilde{Z}^i_t}{|\widetilde{Z}_t|^2}\cdot\frac{d|\widetilde{Z}^i_t|}{dt} \\ &= \sum_{w\in J} p_w(\widetilde{X}_t)w_i - \widetilde{X}^i_t\sum_{w\in J} p_w(\widetilde{X}_t)\alpha(w) \\ &= \sum_{w\in J} p_w(\widetilde{X}_t)\left(w_i - \widetilde{X}^i_t\alpha(w)\right).\end{aligned} \tag{3.29}$$

Since (3.29) holds for all $i = 1, 2, ..., k$, we obtain the following.

Theorem 3.7. *The process of population proportions converges in probability to the solution of the autonomous dynamical system*

$$\frac{d\widetilde{X}_t}{dt} = \sum_{w\in J} p_w(\widetilde{X}_t)\left(w - \widetilde{X}_t\alpha(w)\right), \qquad \widetilde{X}_0 = \frac{a_0}{|a_0|}. \tag{3.30}$$

Proof. The proof is given in detail in Theorem 7.1 in Chapter 7. □

Chapter 4

Non-explosion

This chapter is dedicated to establish the result that the stochastic process we constructed as a pure jump process is regular, i.e., non-explosive.

Theorem 4.1. *There is no explosion for the pure jump processes* $(Z_t^L)_{t\geq 0}$ *and* $(\widetilde{Z}_t^L)_{t\geq 0}$.

Proof. The fact that $(Z_t^L)_{t\geq 0}$ and $(\widetilde{Z}_t^L)_{t\geq 0}$ are pure jump processes are clear since they are special cases of the type of pure jump process we constructed in Section 2.2 . Note that $\widetilde{Z}_t^L = \dfrac{Z_t^L}{L}$, it suffices to show that the family of jump process $(Z_t^L)_{t\geq 0}$ indexed by L is regular. We prove that this is true for an arbitrary index $L > 1$.

Define the jump times, as we did in Section 2.2 , that $T_0 = 0, T_n = inf\{t > T_{n-1} \,|\, Z_t^L \neq Z_{T_{n-1}}^L\}$. As we discussed in Section 2.2 , to determine whether the pure jump process $(Z_t^L)_{t\geq 0}$ is regular or explosive, we shall consider the limit of the jump times $(T_n)_{n\geq 0}$:

$$T_\infty = \sum_{i=1}^{+\infty} \frac{\eta_i}{\lambda(Z_{T_{i-1}}^L)}, \tag{4.1}$$

where $(\eta_i)_{i\geq 1}$ are i.i.d. exponential with unit rate which are also independent of the process $(Z_t^L)_{t\geq 0}$. Note that we set $\dfrac{1}{\lambda(Z_{T_{i-1}}^L)} = +\infty$ if $\lambda(Z_{T_{i-1}}^L) = 0$ indicating the next jump will take forever to occur. Also in this case we would have only finitely many jumps and $T_\infty = +\infty$ which are consistent with non-explosion.

Since we assumed that the jumps are bounded in the sense that there is a constant m' such that $|Z^L_{T_n} - Z^L_{T_{n-1}}| \leq m'$ and the rate for the exponential clock at any time t is $\lambda(Z^L_t) = |Z^L_t|$, also recall that $(Z^L_t)_{t\geq 0}$ is just $(Z_t)_{t\geq 0}$ imposed with an initial conditions that for a fixed $a_0 = (a^1_0, a^2_0, ..., a^k_0) \in \mathbb{R}^k_+ \setminus \{0\}$, $Z^L_0 = \left([La^1_0], [La^2_0], ..., [La^k_0]\right)$ which implies $|Z^L_0| \leq L|a_0|$, we obtain estimates for the jump rates:

$$0 \leq \lambda(Z^L_{T_n}) \leq |Z^L_0| + nm' \leq L|a_0| + nm'. \tag{4.2}$$

If after a finite many jumps we have $|Z^L_{T_n}| = 0$ for some $n \geq 0$, we know the process will stay at 0 thereafter by construction and there is non-explosion. We need only consider the case of non-extinction, i.e., $|Z_{T_n}| > 0, \forall n \geq 0$.

Define $\lambda_i = L|a_0| + im', i = 0, 1, ...,$ we have by (4.2) and non-negativeness of η_i that

$$T_\infty \geq \sum_{i=1}^{+\infty} \frac{\eta_i}{\lambda_{i-1}}, \quad a.s.. \tag{4.3}$$

Note that for a random variable $\dfrac{\eta_i}{\lambda_{i-1}}$, where $(\eta_i)_{i\geq 1}$ are i.i.d. unit exponential random variables as mentioned before, we have

$$E[exp\{-\frac{\eta_i}{\lambda_{i-1}}\}] = \frac{1}{1 + \frac{1}{\lambda_{i-1}}}. \tag{4.4}$$

Besides, $0 \leq E[exp\{-T_\infty\}] \leq 1$ always exists by the bounded convergence theorem.

Moreover, by equations (4.2), (4.4) and independence of $(\eta_i)_{i\geq 1}$, we have

$$\begin{aligned} E[exp\{-T_\infty\}] \leq \quad & E[exp\{-\sum_{i=1}^{+\infty}\frac{\eta_i}{\lambda_{i-1}}\}] \\ &= \prod_{i=1}^{+\infty} E[exp\{-\frac{\eta_i}{L|a_0|+im'}\}] \\ &= \prod_{i=1}^{+\infty}(1+\frac{1}{L|a_0|+im'})^{-1}. \end{aligned} \tag{4.5}$$

Note that $\sum_{i=1}^{+\infty}\frac{1}{L|a_0|+im'} = +\infty$ which implies $\prod_{i=1}^{+\infty}(1+\frac{1}{L|a_0|+im'}) = +\infty$, we obtain from equation (4.5) that

$$E[exp\{-T_\infty\}] = 0,$$

and thus

$$P(T_\infty = +\infty) = 1.$$

This is the desired result. □

Chapter 5

Tightness of $\widetilde{Z}_t^L$

To show the limit as $L \to \infty$ of the process $(\widetilde{Z}_t^L)_{t\geq 0}$ exists and is unique, we will prove the tightness of the family of processes $(\widetilde{Z}_t^L)_{t\geq 0}$ indexed by L, in this chapter. To do that, we prove the tightness of $(\widetilde{Z}_t^{L,i})_{0\leq t\leq T}$ for an arbitrary but fixed $T \geq 0$. The main result of this chapter is Theorem 5.8.

Note that two conditions for tightness [22] of a family of stochastic processes $\{y^L(\cdot)\}_{L>0}$ indexed by L with values in S seen as measures on the Skorokhod space $D([0,T],S)$ which ensure that any limit point belongs to $C([0,T],S)$ are

$$\text{there exists a M} > 0 \text{ such that} \limsup_{L\to\infty} P\left(\|y^L(0)\| > M\right) = 0 \tag{5.1}$$

and

$$\forall \epsilon > 0,\ \lim_{\delta\to 0}\limsup_{L\to\infty} P\left(\sup_{\substack{s,t\in[0,T]\\ 0<|t-s|<\delta}} \|y^L(t) - y^L(s)\| > \epsilon\right) = 0. \tag{5.2}$$

By (3.18) we can see that $\lim_{L\to\infty} \widetilde{Z}_0^L = \lim_{L\to\infty} \frac{Z_0^L}{L} = a_0$, and if we choose M big enough we have

$\limsup_{L\to\infty} p\left(\|\widetilde{Z}_0^{L,i}\| > M\right) = 0$. Therefore, condition (5.1) is satisfied for $(\widetilde{Z}_t^{L,i})_{0\leqslant t\leqslant T}$. To prove condition (5.2), it is equivalent to show, by a variant of Aldous criterion [1] that

$$\forall\epsilon > 0,\ \lim_{\delta\to 0}\limsup_{L\to\infty} P\left(\sup_{\substack{0\leqslant s<s+t\leqslant T\\ 0<t<\delta}} \|\widetilde{Z}_{s+t}^{L,i} - \widetilde{Z}_s^{L,i}\| > \epsilon\right) = 0. \tag{5.3}$$

As we mentioned earlier in Chapter 3, we'd like to use a martingale relation (3.16) for $(Z_t^L)_{t\geqslant 0}$ and further obtain results for $(\widetilde{Z}_t^L)_{t\geqslant 0}$. However, we need to justify why the general martingale relations as in (2.33) and (2.34) where the test functions family were supposed to be $C_c^2(\mathbb{R}^d)$ and be chosen to be projection maps $f(z) = z^i, i = 1, 2, ..., k$ in our case. In particular, to show that $\widetilde{Z}_t^{L,i} - \widetilde{Z}_0^{L,i} - \int_0^t |\widetilde{Z}_s^L| \sum_{w\in J} \Pi\left(L\widetilde{Z}_s^L, L\widetilde{Z}_s^L + w\right) w_i ds$ is a martingale, we need to show it has finite expectation. We shall show a stronger result, namely that $E\left[\sup_{0\leqslant t\leqslant T} |\widetilde{Z}_{t^-}^L|^2\right]$ is uniformly bounded, as a first step.
First we define a sequence of stopping times and use truncation techniques so that we can use a martingale relation for the stopped process to do the estimates.

5.1 Localization of the martingales

Define the stopping times $\tau_N = \inf\{t > 0, ||Z_t^L| > N\}$. Since $(Z_t^L)_{t\geq 0}$ is non-explosive, as was shown in the previous chapter, meaning it has finitely many jumps in any finite time almost surely, also its jump size is bounded by a constant $m' > 0$ by assumption, we have

$$\lim_{N\to+\infty} \tau_N = +\infty\, a.s..$$

Define

$$f_N(z) = \begin{cases} z_i, & \text{if } |z| \le N + 2m' \\ 0, & \text{if } |z| \ge N + 2m' + 1 \\ \text{smooth}, & \text{if } N + 2m' < |z| < N + m' + 1. \end{cases} \tag{5.4}$$

Note that by “smooth" we mean the function $f_N(z)$ shall connect the part where $|z| \le N+2m'$ and the part where $|z| \ge N + 2m' + 1$ in a smooth way that $f_N(z)$ is smooth ($f_N \in C_c^2$ is what we need). Basically, $f_N(z)$ is a truncation of f which agrees with in $|Z| \le N + 2m'$ and smoothly approaches to 0 in an unit interval and stays there afterwards. Clearly $f_N \to f$ pointwise. Also note that $f_N \in C_c^2(\mathbb{R}^d, \mathbb{R})$, and we have

$$f_N(Z^L_{t\wedge\tau_N}) = Z^{L,i}_{t\wedge\tau_N}, \qquad f_N(Z^L_{t\wedge\tau_N} + w) = Z^{L,i}_{t\wedge\tau_N} + w_i, \forall w \in J, \tag{5.5}$$

because the process has RCLL paths and bounded jumps that

$$|Z^L_{t\wedge\tau_N}| \le N + m', \; |Z^L_{t\wedge\tau_N} + w| \le N + 2m'.$$

Therefore, $\forall t \in [0, +\infty), \forall w \in J$, we have

$$\lim_{N\to+\infty} f_N(Z^L_{t\wedge\tau_N}) = Z^{L,i}_t, \text{ and } \lim_{N\to+\infty} f_N(Z^L_{t\wedge\tau_N} + w) = Z^{L,i}_t + w_i \quad \omega - a.s.. \tag{5.6}$$

Remark 5.1. *Up to the time τ_N, the jump rates $\lambda(Z^L_s) = |Z^L_s|$ are bounded by N. Technically, we use the martingales (2.33) - (2.34) obtained in the general construction of pure jump processes with bounded coefficients.*

Applying the truncation functions as test functions to the process $(\widetilde{Z}^{L,i}_s)_{s\ge0}$, we obtain the following martingales.

$$\begin{aligned}
& f_N(Z^L_{t\wedge\tau_N}) - f_N(Z^L_0) - \int_0^{t\wedge\tau_N} |Z^L_s| \sum_{w\in J} \Pi\left(Z^L_s, Z^L_s + w\right)\left(f_N(Z^L_s + w) - f_N(Z^L_s)\right) ds \\
= \quad & Z^{L,i}_{t\wedge\tau_N} - Z^{L,i}_0 - \int_0^{t\wedge\tau_N} |Z^L_s| \sum_{w\in J} \Pi\left(Z^L_s, Z^L_s + w\right) w_i ds \triangleq \mathscr{M}^{L,i}_t.
\end{aligned} \tag{5.7}$$

Note that $\widetilde{Z}^{L,i}_t$ is just a scaled process of $Z^{L,i}_t$, a similar martingale relation also holds for $\widetilde{Z}^{L,i}_t$ that we have proven that the following is a martingale

$$\widetilde{Z}^{L,i}_{t\wedge\tau_N} - \widetilde{Z}^{L,i}_0 - \int_0^{t\wedge\tau_N} |\widetilde{Z}^L_s| \sum_{w\in J} \Pi\left(L\widetilde{Z}^L_s, L\widetilde{Z}^L_s + w\right) w_i ds \triangleq \widetilde{\mathscr{M}}^{L,i}_t, \tag{5.8}$$

together with its corresponding quadratic variation which can be obtained from (2.34).

5.2 Uniform bound on the total number of particles $|\widetilde{Z}^L_t|$

For ease of notation, we shall work on the process without indicating the stopping time. The key point is that for each fixed N, we develop the bounds from Proposition 5.2, and establish that *they do not depend on N*. By letting $N \to \infty$, the bounds are proven for the process without stopping, and immediately are showing that the martingales can be extended to smooth linear functions, beyond compactly supported functions.

Proposition 5.2. *For all $T > 0$, $E\left[\sup_{0\leqslant t\leqslant T} |\widetilde{Z}^L_{t-}|^2\right]$ is uniformly bounded for all $L > 1$. And as a consequence, $E\left[\sup_{0\leqslant t\leqslant T} |\widetilde{Z}^L_{t-}|\right]$ is also uniformly bounded for $L > 1$.*

More precisely, for any $L > 1$, we obtain the estimates

$$E\left[\sup_{0\leqslant t\leqslant T}|\widetilde{Z}^L_{t^-}|^2\right] \leqslant \left[6k^2mCT + 3kM\right]\cdot\exp\left\{k^2CT[6m+3CT]\right\} := C_2(k,m,T)\,, \tag{5.9}$$

$$E\left[\sup_{0\leqslant t\leqslant T}|\widetilde{Z}^L_{t^-}|\right] \leqslant \frac{1}{2}\left(1 + C_2(k,m,T)\right) := C_1(k,m,T)\,. \tag{5.10}$$

Remark 5.3. *1) Since $T > 0$ is fixed, but arbitrary, we can remove the left-side limits markers in the proposition without loss of generality. 2) The bounds do not depend on N, which allows removing the localization. This step is important, but technical. 3) The bounds do not depend on L, which is the main significance of this proposition. This is the most important bound we need for tightness.*

Proof. As noted above, even relation (5.8) holds up to a stopping time, we shall write

$$\widetilde{Z}^{L,i}_t = \widetilde{Z}^{L,i}_0 + \int_0^t |\widetilde{Z}^L_{s-}|\sum_{w\in J}\Big(\Pi(L\widetilde{Z}^L_{s-}, L\widetilde{Z}^L_{s-}+w)\Big)w_i ds + \widetilde{\mathscr{M}}^{L,i}_t. \tag{5.11}$$

By the inequality $(a+b+c)^2 \leqslant 3(a^2+b^2+c^2)$ for real a, b, c, we have that

$$(\widetilde{Z}^{L,i}_t)^2 \leqslant 3\left[(\widetilde{Z}^{L,i}_0)^2 + \left(\int_0^t |\widetilde{Z}^L_{s-}|\sum_{w\in J}\Big(\Pi(L\widetilde{Z}^L_{s-}, L\widetilde{Z}^L_{s-}+w)\Big)w_i ds\right)^2 + (\widetilde{\mathscr{M}}^{L,i}_t)^2\right]. \tag{5.12}$$

Applying $\sup\limits_{0\leqslant t\leqslant t'}$ on both sides of (5.12) for $0 \leqslant t' \leqslant T$, we further have

$$\begin{aligned}\sup_{0\leqslant t\leqslant t'}(\widetilde{Z}^{L,i}_t)^2 \leqslant{}& 3\left[\sup_{0\leqslant t\leqslant t'}(\widetilde{Z}^{L,i}_0)^2 + \sup_{0\leqslant t\leqslant t'}(\widetilde{\mathscr{M}}^{L,i}_t)^2\right]\\ &+ 3\left[\sup_{0\leqslant t\leqslant t'}\left(\int_0^t |\widetilde{Z}^L_{s-}|\sum_{w\in J}\Big(\Pi(L\widetilde{Z}^L_{s-}, L\widetilde{Z}^L_{s-}+w)\Big)w_i ds\right)^2\right].\end{aligned} \tag{5.13}$$

Applying expectation on both sides of (5.13) and using $\sup\limits_{0\leqslant t\leqslant t'}(\widetilde{Z}^{L,i}_{t-})^2 \leqslant \sup\limits_{0\leqslant t\leqslant t'}(\widetilde{Z}^{L,i}_t)^2$ by the fact that $\widetilde{Z}_{t-}^{L,i}$ is right continuous with left limits(RCLL), we get

$$
\begin{aligned}
E\left[\sup_{0\leqslant t\leqslant t'}(\widetilde{Z}_{t-}^{L,i})^2\right] &\leqslant E\left[\sup_{0\leqslant t\leqslant t'}(\widetilde{Z}_{t}^{L,i})^2\right] \\
&\leqslant E\left[3\sup_{0\leqslant t\leqslant t'}(\widetilde{Z}_{0}^{L,i})^2\right] + E\left[3\sup_{0\leqslant t\leqslant t'}(\widetilde{\mathscr{M}}_{t}^{L,i})^2\right] \\
&\quad + E\left[3\sup_{0\leqslant t\leqslant t'}\left(\int_0^t |\widetilde{Z}_{s-}^{L}|\sum_{w\in J}\Big(\Pi(L\widetilde{Z}_{s-}^{L}, L\widetilde{Z}_{s-}^{L}+w)\Big)w_i ds\right)^2\right].
\end{aligned}
\tag{5.14}
$$

We want to find estimates for the right side terms of (5.14). First we note that the term $\sup_{0\leqslant t\leqslant t'}(\widetilde{Z}_{0}^{L,i})^2$ is just $(\widetilde{Z}_{0}^{L,i})^2$ and with our assumption $\lim_{L\to\infty}\widetilde{Z}_{0}^{L} = a_0$ there is nothing to estimate as $L \to \infty$. Since $\left|\sum_{w\in J}\Big(\Pi(L\widetilde{Z}_{s-}^{L}, L\widetilde{Z}_{s-}^{L}+w)\Big)w_i\right|$ is uniformly bounded by a constant C for all L and thus $\left|\sum_{w\in J}\Big(\Pi(L\widetilde{Z}_{s-}^{L}, L\widetilde{Z}_{s-}^{L}+w)\Big)w_i{}^2\right|$ by mC, we have that:

$$
\begin{aligned}
&E\left[\sup_{0\leqslant t\leqslant t'}\left(\int_0^t |\widetilde{Z}_{s-}^{L}|\sum_{w\in J}\Big(\Pi(L\widetilde{Z}_{s-}^{L}, L\widetilde{Z}_{s-}^{L}+w)\Big)w_i ds\right)^2\right] \\
&\leqslant C^2 E\left[\sup_{0\leqslant t\leqslant t'}\left(\int_0^t |\widetilde{Z}_{s-}^{L}| ds\right)^2\right] \\
&= C^2 E\left[\left(\int_0^{t'} |\widetilde{Z}_{s-}^{L}| ds\right)^2\right] \qquad \text{(nonnegativeness of } |\widetilde{Z}_{s-}^{L}|\text{)} \\
&\leqslant C^2 E\left[\left(\int_0^{t'} \sup_{0\leqslant t\leqslant s}|\widetilde{Z}_{t^-}^{L}| ds\right)^2\right] \\
&\leqslant C^2 E\left[\int_0^{t'} \left(\sup_{0\leqslant t\leqslant s}|\widetilde{Z}_{t^-}^{L}|\right)^2 ds\right] T \qquad \text{(since } 0\leqslant t'\leqslant T\text{)} \\
&= C^2 E\left[\int_0^{t'} \sup_{0\leqslant t\leqslant s}|\widetilde{Z}_{t^-}^{L}|^2 ds\right] T.
\end{aligned}
\tag{5.15}
$$

Therefore, we get the estimate that

$$
\begin{aligned}
& E\left[3 \sup_{0\leqslant t\leqslant t'}\left(\int_0^t |\widetilde{Z}^L_{s-}| \sum_{w\in J}\left(\Pi(L\widetilde{Z}^L_{s-}, L\widetilde{Z}^L_{s-}+w)\right) w_i ds\right)^2\right] \\
\leqslant\ & 3C^2T\cdot E\left[\int_0^{t'} \sup_{0\leqslant t\leqslant s} |\widetilde{Z}^L_{t-}|^2 ds\right].
\end{aligned}
\tag{5.16}
$$

Now we proceed to estimate $E\left[3 \sup_{0\leqslant t\leqslant t'} (\widetilde{\mathscr{M}}^{L,i}_t)^2\right]$. Note that $|\widetilde{\mathscr{M}}^{L,i}_t|$ is a submartingale, using Doob's martingale inequality for p=2, we have

$$
\left(E\left[\left(\sup_{0\leqslant t\leqslant t'} |\widetilde{\mathscr{M}}^{L,i}_t|\right)^2\right]\right)^{\frac{1}{2}} \leqslant 2\cdot\left(E\left[|\widetilde{\mathscr{M}}^{L,i}_{t'}|^2\right]\right)^{\frac{1}{2}}
\tag{5.17}
$$

or squaring both sides

$$
E\left[\left(\sup_{0\leqslant t\leqslant t'} |\widetilde{\mathscr{M}}^{L,i}_t|\right)^2\right] \leqslant 4\cdot E\left[|\widetilde{\mathscr{M}}^{L,i}_{t'}|^2\right].
\tag{5.18}
$$

Since $\widetilde{\mathscr{M}}^{L,i}_{t'}$ is the martingale corresponding to the process $\widetilde{Z}^{L,i}_{t'}$, we have

$$
\begin{aligned}
E\left[(\widetilde{\mathscr{M}}^{L,i}_{t'})^2\right] &= E\left[\int_0^{t'} L|\widetilde{Z}^L_{s-}| \sum_{w\in J}\left(\Pi(L\widetilde{Z}^L_{s-}, L\widetilde{Z}^L_{s-}+w)\right)\left(\frac{w_i}{L}\right)^2 ds\right] \\
&\leqslant \frac{mC}{L} E\left[\int_0^{t'} |\widetilde{Z}^L_{s-}| ds\right] \\
&\leqslant \frac{mC}{L} E\left[\int_0^{t'} \sup_{0\leqslant t\leqslant s} |\widetilde{Z}^L_{t-}| ds\right] \\
&\leqslant \frac{mC}{2L} E\left[\int_0^{t'} [(\sup_{0\leqslant t\leqslant s} |\widetilde{Z}^L_{t-}|)^2 + 1] ds\right] \\
&\leqslant \frac{mCT}{2L} + \frac{mC}{2L}\cdot E\left[\int_0^{t'} \sup_{0\leqslant t\leqslant s} |\widetilde{Z}^L_{t-}|^2 ds\right].
\end{aligned}
\tag{5.19}
$$

Therefore, we obtain from (5.18) and (5.19) that

$$E\left[3\sup_{0\leqslant t\leqslant t'}(\widetilde{\mathscr{M}}_t^{L,i})^2\right]\leqslant\frac{6mCT}{L}+\frac{6mC}{L}\cdot E\left[\int_0^{t'}\sup_{0\leqslant t\leqslant s}|\widetilde{Z}_{t-}^L|^2ds\right]. \tag{5.20}$$

Substituting result (5.16) and (5.20) into (5.14), we get

$$\begin{aligned}
&E\left[\sup_{0\leqslant t\leqslant t'}(\widetilde{Z}_{t-}^{L,i})^2\right]\\
&\leqslant 3E\left[\sup_{0\leqslant t\leqslant t'}(\widetilde{Z}_0^{L,i})^2\right]+3C^2TE\left[\int_0^{t'}\sup_{0\leqslant t\leqslant s}|\widetilde{Z}_{t-}^L|^2ds\right]\\
&+\frac{6mCT}{L}+\frac{6mC}{L}\cdot E\left[\int_0^{t'}\sup_{0\leqslant t\leqslant s}|\widetilde{Z}_{t-}^L|^2ds\right]\\
&=\frac{6mCT}{L}+3E\left[\sup_{0\leqslant t\leqslant t'}(\widetilde{Z}_0^{L,i})^2\right]+\left[\frac{6mC}{L}+3C^2T\right]\cdot E\left[\int_0^{t'}\sup_{0\leqslant t\leqslant s}|\widetilde{Z}_{t-}^L|^2ds\right].
\end{aligned} \tag{5.21}$$

Adding the above result for $i=1,2,\cdots,k$, we get

$$\begin{aligned}
&\sum_{i=1}^k E\left[\sup_{0\leqslant t\leqslant t'}(\widetilde{Z}_{t-}^{L,i})^2\right]\\
&\leqslant\frac{6kmCT}{L}+3\sum_{i=1}^k E\left[\sup_{0\leqslant t\leqslant t'}(\widetilde{Z}_0^{L,i})^2\right]\\
&+\left[\frac{6kmC}{L}+3kC^2T\right]\cdot E\left[\int_0^{t'}\sup_{0\leqslant t\leqslant s}|\widetilde{Z}_{t-}^L|^2ds\right].
\end{aligned} \tag{5.22}$$

Since $|\widetilde{Z}^L_{t-}| = \sum_{i=1}^k |\widetilde{Z}^{L,i}_{t-}|$, we have that $\sum_{i=1}^k |\widetilde{Z}^{L,i}_{t-}|^2 \leqslant |\widetilde{Z}^L_{t-}|^2 \leqslant k \sum_{i=1}^k |\widetilde{Z}^{L,i}_{t-}|^2$ which implies

$$\begin{aligned}
&E\left[\sup_{0\leqslant t\leqslant t'} |\widetilde{Z}^L_{t-}|^2\right] \leqslant k\sum_{i=1}^k E\left[\sup_{0\leqslant t\leqslant t'} (\widetilde{Z}^{L,i}_{t-})^2\right] \\
&\leqslant \frac{6k^2mCT}{L} + 3k\sum_{i=1}^k E\left[\sup_{0\leqslant t\leqslant t'} (\widetilde{Z}^{L,i}_0)^2\right] + k^2C\left[\frac{6m}{L} + 3CT\right] E\left[\int_0^{t'} \sup_{0\leqslant t\leqslant s} |\widetilde{Z}^L_{t^-}|^2 ds\right] \qquad (5.23) \\
&\leqslant \frac{6k^2mCT}{L} + 3kE\left[\sup_{0\leqslant t\leqslant t'} |\widetilde{Z}^L_0|^2\right] + k^2C\left[\frac{6m}{L} + 3CT\right] \cdot \int_0^{t'} E\left[\sup_{0\leqslant t\leqslant s} |\widetilde{Z}^L_{t^-}|^2\right] ds.
\end{aligned}$$

Define $U_L(t') = E\left[\sup_{0\leqslant t\leqslant t'} |\widetilde{Z}^L_{t^-}|^2\right]$ and note that $E\left[\sup_{0\leqslant t\leqslant t'} |\widetilde{Z}^L_0|^2\right] = U_L(0)$, we obtain from (5.23) that

$$U_L(t') \leqslant \frac{6k^2mCT}{L} + 3kU_L(0) + k^2C\left[\frac{6m}{L} + 3CT\right] \cdot \int_0^{t'} U_L(s)ds. \qquad (5.24)$$

By Gronwall's inequality for the integral form, given L, we have from (5.24) that

$$U_L(t') \leqslant \left[\frac{6k^2mCT}{L} + 3kU_L(0)\right] \cdot \exp\left\{k^2C\left[\frac{6m}{L} + 3CT\right]t'\right\}. \qquad (5.25)$$

As we are considering the case $0 \leqslant t' \leqslant T$, for a given L, we get an uniform bound for all $U_L(t')$, i.e.,

$$U_L(t') \leqslant \left[\frac{6k^2mCT}{L} + 3kU_L(0)\right] \cdot \exp\left\{k^2C\left[\frac{6m}{L} + 3CT\right]T\right\}, \forall\, 0 \leqslant t' \leqslant T. \qquad (5.26)$$

Since we have the additional assumption that $\lim_{L\to\infty} \widetilde{Z}^L_0 = a_0$, we know that $U_L(0)$ must be uniformly bounded by a constant M; and we know that L is large, say $L \geqslant 1$ at least, we obtain an uniform bound for all $E\left[\sup_{0\leqslant t\leqslant t'} |\widetilde{Z}^L_{t-}|^2\right]$:

$$U_L(t') \leqslant \left[6k^2mCT + 3kM\right] \cdot \exp\left\{k^2CT[6m + 3CT]\right\}, \forall\, 0 \leqslant t' \leqslant T, \forall L \geqslant 1. \tag{5.27}$$

Since $t' \leq T$, we get

$$E\left[\sup_{0\leqslant t\leqslant T} |\widetilde{Z}^L_{t-}|^2\right] \leqslant \left[6k^2mCT + 3kM\right] \cdot \exp\left\{k^2CT[6m + 3CT]\right\}, \forall L > 1. \tag{5.28}$$

Note that this bound is uniform and has nothing to do with the time $t \in [0, T]$. It follows that it doesn't depend on stopping time τ_N we introduced. If we worked with the process with stopping time, we shall pass N to infinity now, and as $\lim_{N\to+\infty} \tau_N = +\infty\, a.s.$, we would obtain the same bound as above. Therefore, Proposition 5.2 is proved. □

5.3 Delocalization of the martingales

Until now, we used local martingales up to time τ_N to prove universal bounds for the total number of particles. This bound (uniform square integrability) allows us to extend the set of test functions. To show the modulus of continuity part of tightness, done in Section 5.4, we need the martingale relations in (2.29) and (2.30) to hold for the process $(\widetilde{Z}^L_t)_{t\geqslant 0}$ where the test function can be projections maps. This proof follows the same steps (5.4)-(5.8).

Lemma 5.4. *We can extend the test functions from $f \in C^2_c(\mathbb{R}^k, \mathbb{R})$ to include projection functions $f(z) = z_i, \forall z \in \mathbb{R}^k, i = 1, \cdots, k$, and the corresponding martingale relation for Z^L_t and $\widetilde{Z}^L_t$ still hold. More precisely, we have martingales*

$$f(Z^L_t) - f(Z^L_0) - \int_0^t |Z^L_s| \sum_{w\in J} \Pi\left(Z^L_s, Z^L_s + w\right)\left(f(Z^L_s + w) - f(Z^L_s)\right) ds \tag{5.29}$$

which shall be denoted by $\mathscr{M}_t^{f,Z^L}$, *and*

$$(\mathscr{M}_t^{f,Z_t^L})^2 - \int_0^t |Z_s^L| \sum_{w \in J} \Pi\left(Z_s^L, Z_s^L + w\right)\left(f(Z_s^L + w) - f(Z_s^L)\right)^2 ds \tag{5.30}$$

which shall be denoted by $\mathscr{N}_t^{f,Z^L}$ *for* $i = 1, 2, ..., k$ *and all index* $L > 1$, *where the test function* $f : \mathbb{R}^k \to \mathbb{R}$ *can be all functions in* C_c^2 *as well as projection maps.*

Remark 5.5. *In particular, if* $f(z) = z_i, i = 1, 2, ..., k$, *being projection maps, we have martingales*

$$(Z_i^L)_t - (Z_i^L)_0 - \int_0^t |Z_s^L| \sum_{w \in J} \Pi\left(Z_s^L, Z_s^L + w\right) w_i ds \tag{5.31}$$

which we shall denote by $\mathscr{M}_t^{Z_i^L}$, *and*

$$(\mathscr{M}_t^{Z_i^L})^2 - \int_0^t |Z_s^L| \sum_{w \in J} \Pi\left(Z_s^L, Z_s^L + w\right) w_i^2 ds \tag{5.32}$$

which we shall denote by $\mathscr{N}_t^{Z_i^L}$ *for* $i = 1, 2, ..., k$, *and all indexes* $L > 1$.

Proof. Note that $(\widetilde{Z}_t^L)_{t\geq 0}$ is merely a scaled version of $(Z_t^L)_{t\geq 0}$, it suffices to prove the lemma for the latter. Recall that the process $(Z_t^L)_{t\geq 0}$ is a non-explosive (Theorem 4.1) pure jump process. Such a process has an infinitesimal generator as in (2.31), and we have martingales (5.29) and (5.30) for $f \in C_c^2(\mathbb{R}^d, \mathbb{R})$. The goal is to show we can choose f to be projection maps and the martingale relations stated above still hold. Recall the initial condition for Z_t^L is given by equation (3.18) that

$$Z_0^L = \left([La_0^1], [La_0^2], ..., [La_0^k]\right)$$

where $a_0 = (a_0^1, a_0^2, ..., a_0^k) \in \mathbb{R}_+^k \setminus \{0\}$ is fixed.

Define the stopping times τ_N and truncation functions $f_N(\cdot)$ as we did in Section 5.1 that $\tau_N = \inf\{t > 0, ||Z_t^L| > N\}$ and f_N as in (5.4).

Fix a time horizon $T > 0$. For $0 \leq s < t \leq T$ we can stop at time $\tau_N \wedge T$ (uniformly bounded by T) and from the Optional Stopping Theorem we have the martingale

$$\begin{aligned} & f_N(Z^L_{t\wedge\tau_N}) - f_N(Z^L_0) - \int_0^{t\wedge\tau_N} |Z^L_s| \sum_{w\in J} \Pi\left(Z^L_s, Z^L_s + w\right)\left(f_N(Z^L_s + w) - f_N(Z^L_s)\right) ds \\ =& f(Z^L_{t\wedge\tau_N}) - f(Z^L_0) - \int_0^{t\wedge\tau_N} |Z^L_s| \sum_{w\in J} \Pi\left(Z^L_s, Z^L_s + w\right)\left(f(Z^L_s + w) - f(Z^L_s)\right) ds \triangleq M^N_t \,. \end{aligned} \tag{5.33}$$

We shall use Lemma 5.6 to prove our result.

Step 1 - almost sure convergence. Using (5.6) we see that the first two terms of M_t^N converge a.s.

$$\lim_{N\to+\infty} \left(f_N(Z^L_{t\wedge\tau_N}) - f_N(Z^L_0)\right) = \lim_{N\to+\infty} \left((Z^L_{t\wedge\tau_N}) - f(Z^L_0)\right) = f(Z_t) - f(Z_0)\,.$$

The time integral part also converges almost surely. The integrand is a product of a bounded function $0 \leq \Pi(\cdot,\cdot) \leq 1$ and RCLL functions, since the composition with an RCLL function with a continuous function is RCLL. It is the case that the integrand denoted by $b(s,\omega)$ for now, is bounded a.s. on the interval $[0,T]$, implying that the integral $t \to \int_0^t b(s,\omega)ds$, where

$$b(s,\omega) \triangleq |Z^L_s(\omega)| \sum_{w\in J} \Pi\left(Z^L_s(\omega), Z^L_s(\omega) + w\right)\left(f(Z^L_s(\omega) + w) - f(Z^L_s(\omega))\right)$$

is a continuous mapping. We mention that $\Pi(\cdot,\cdot)$ is bounded but we didn't require it to be continuous, yet the above integral, $t \to \int_0^t b(s,\omega)ds$, is continuous in t.

Since $t \wedge \tau_N \to t$ as $N \to \infty$ we have a.s. convergence for the time integral as well. This shows that

$$\lim_{N\to+\infty} M_t^N = M_t \qquad a.s.$$

where M_t^N is defined in (5.33) and

$$M_t \triangleq f(Z_t^L) - f(Z_0^L) - \int_0^t |Z_s^L| \sum_{w\in J} \Pi\left(Z_s^L, Z_s^L + w\right)\left(f(Z_s^L + w) - f(Z_s^L)\right) ds.$$

Step 2 - the uniform integrability. We need to show that for all $t \in [0, T]$,

$$E[\sup_{N>0} |M_t^N|] < \infty\,. \tag{5.34}$$

We shall show the stronger property that there exists $B_T(\omega) \geq 0$ such that

$$|M_t^N| \leq B_T(\omega) \quad \text{and} \quad E[B_T(\omega)] < \infty\,. \tag{5.35}$$

We can assume the functions $|f_N(z)| \leq |z| + 2m' + 1$. Inspecting the first two terms of M_t^N we see they are bounded by $2[\sup_{t\in[0,T]} |Z_t^L| + 2m' + 1]$. The integral term is bounded by $m'|J|T \sup_{t\in[0,T]} |Z_t^L|$. Take $B_T(\omega) := 4m' + 2 + (2 + m'|J|T) \sup_{t\in[0,T]} |Z_t^L(\omega)|$ and the conclusion is an immediate consequence of the uniform bound in time of $\sup_{t\in[0,T]} |Z_t^L(\omega)|$ proven in Proposition 5.2 in Section 5.2.

Notice that L is fixed here, and the bound is actually stronger, holding uniformly as well in L for $\widetilde{Z}_t^L = Z_t^L/L$ but this fact is not used in the present proof. □

Lemma 5.6. *Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ be a filtered probability space, $(M_t^N(\omega))$ be a family of $\mathcal{F}_t$ - martingales indexed by $N > 0$ such that (i) $\lim_{N\to\infty} M_t^N(\omega) = M_t(\omega)$ almost surely for $\omega \in \Omega$ and (ii) $E[\sup_{N>0} |M_t^N(\omega)|] < \infty$, for all $t > 0$. Then M_t is a $\mathcal{F}_t$ - martingale.*

Proof. Denote $C_t(\omega) = \sup_{N>0} |M_t^N(\omega)|$. By dominated convergence theorem, it follows that the a.s. limit M_t is integrable. It remains to show that $E[M_t \,|\, \mathcal{F}_s] = M_s$, for $0 \leq s \leq t$. Take $\Phi(\omega)$ a bounded measurable function with respect to $\mathcal{F}_s$. We have to prove that

$$E[M_t(\omega)\Phi(\omega)] = E[M_s\Phi(\omega)]\,. \tag{5.36}$$

It is true that for each $N > 0$ we have

$$E[M_t^N(\omega)\Phi(\omega)] = E[M_s^N\Phi(\omega)]\,. \tag{5.37}$$

We verify the conditions of Lebesgue Dominated Convergence Theorem. Convergence a.s. is true by hypothesis (i), as Φ does not depend on N. It is immediate that $|M_t^N(\omega)\Phi(\omega)| \leq ||\Phi||C_t(\omega)$ and the right-hand side bound does not depend on N and is integrable by (ii). Here $||\Phi||$ is an upper bound for Φ. A similar bound is valid for the time s. Let $N \to \infty$ in (5.37). Using dominated convergence, we have equality on both sides, proving (5.36). □

5.4 Uniform continuity condition

The bound (5.3) is proven in the next Theorem.

Theorem 5.7. *$\forall\, T > 0, L > 1$, let $(\widetilde{Z}_t^{L,i})_{0\leqslant t\leqslant T}$ be the random processes defined as in the model setup, for all $0 \leqslant s < s+t < s+\delta \leqslant T$, we have*

$$\lim_{\delta\to 0}\lim_{L\to\infty} E \sup_{0\leqslant t\leqslant\delta} \left|\widetilde{Z}_{s+t}^{L,i} - \widetilde{Z}_s^{L,i}\right|^2 = 0, \tag{5.38}$$

and it follows that condition (5.3) *is satisfied.*

Proof. By lemma 5.4, pick $0 \le s < s+t \le T$ for a fixed $T > 0$, we have

$$\begin{aligned}\widetilde{Z}^{L,i}_{s+t} - \widetilde{Z}^{L,i}_{s} = &\int_s^{s+t} |\widetilde{Z}^L_{s-}| \sum_{w \in J} \left(\Pi(L\widetilde{Z}^L_{s-}, L\widetilde{Z}^L_{s-} + w) - p_w\left(\frac{\widetilde{Z}^L_{s-}}{|\widetilde{Z}^L_{s-}|}\right)\right) w_i ds \\ &+ \int_s^{s+t} |\widetilde{Z}^L_{s-}| \sum_{w \in J} p_w\left(\frac{\widetilde{Z}^L_{s-}}{|\widetilde{Z}^L_{s-}|}\right) w_i ds + \widetilde{\mathscr{M}}^{L,i}_{s+t} - \widetilde{\mathscr{M}}^{L,i}_{s}.\end{aligned}$$

To show (5.38), it suffices to prove the following:

$$\lim_{\delta\to 0}\lim_{L\to\infty} E \sup_{0 \leqslant t \leqslant \delta} \left| \int_s^{s+t} |\widetilde{Z}^L_{s-}| \sum_{w \in J} \left(\Pi(L\widetilde{Z}^L_{s-}, L\widetilde{Z}^L_{s-} + w) - p_w\left(\frac{\widetilde{Z}^L_{s-}}{|\widetilde{Z}^L_{s-}|}\right)\right) w_i ds \right|^2 = 0, \tag{5.39}$$

$$\lim_{\delta\to 0}\lim_{L\to\infty} E \sup_{0 \leqslant t \leqslant \delta} \left| \int_s^{s+t} |\widetilde{Z}^L_{s-}| \sum_{w \in J} p_w\left(\frac{\widetilde{Z}^L_{s-}}{|\widetilde{Z}^L_{s-}|}\right) w_i ds \right|^2 = 0, \tag{5.40}$$

and

$$\lim_{\delta\to 0}\lim_{L\to\infty} E \sup_{0 \leqslant t \leqslant \delta} \left| \widetilde{\mathscr{M}}^{L,i}_{s+t} - \widetilde{\mathscr{M}}^{L,i}_{s} \right|^2 = 0. \tag{5.41}$$

Equation (5.39) is trivially true by (3.23). Note that $\left| \sum_{w \in J} p_w\left(\frac{\widetilde{Z}^L_{s-}}{|\widetilde{Z}^L_{s-}|}\right) w_i \right|$ is uniformly bounded for all L and $E\left[\sup_{0 \leqslant t \leqslant \delta} \left| \int_s^{s+t} |\widetilde{Z}^L_{s-}| ds \right|^2\right] \leqslant \delta^2 E\left[\sup_{0 \leqslant t \leqslant T} |\widetilde{Z}^L_{t-}|^2\right]$, it is sufficient to show that $E \sup_{0 \leqslant t \leqslant T} |\widetilde{Z}^L_{t-}|^2$ is bounded for (5.40) to be true. To show (5.41), note that $\widetilde{\mathscr{M}}^{L,i}_{s+t} - \widetilde{\mathscr{M}}^{L,i}_{s}$ is a martingale and $|\widetilde{\mathscr{M}}^{L,i}_{s+t} - \widetilde{\mathscr{M}}^{L,i}_{s}|^2$ is a nonnegative submartingale, we can use Doob's maximal inequality to estimate that

$$\begin{aligned} & E \sup_{0 \leqslant t \leqslant \delta} \left| \widetilde{\mathscr{M}}^{L,i}_{s+t} - \widetilde{\mathscr{M}}^{L,i}_{s} \right|^2 \leqslant 4E \left| \widetilde{\mathscr{M}}^{L,i}_{s+\delta} - \widetilde{\mathscr{M}}^{L,i}_{s} \right|^2 \\ = \quad & 4E \left| \int_s^{s+\delta} |\widetilde{Z}^L_{s-}| \sum_{w \in J} \left(\Pi(L\widetilde{Z}^L_{s-}, L\widetilde{Z}^L_{s-} + w) \right) {w_i}^2 ds \right| \leqslant C\delta E \sup_{0 \leqslant s \leqslant T} |\widetilde{Z}^L_{s-}| ds \end{aligned}$$

for some positive constant C since $\left|\sum_{w\in J}\left(\Pi(L\widetilde{Z}^L_{s-}, L\widetilde{Z}^L_{s-}+w)\right)w_i{}^2\right|$ is uniformly bounded for all L and i. Note that since $\sup_{0\leqslant t\leqslant T}|\widetilde{Z}^L_t| \leqslant 1+\sup_{0\leqslant t\leqslant T}|\widetilde{Z}^L_t|^2$, we have

$$E[\sup_{0\leqslant t\leqslant T}|\widetilde{Z}^L_t|] \leqslant 1+E[\sup_{0\leqslant t\leqslant T}|\widetilde{Z}^L_t|^2]\,. \tag{5.42}$$

The bounds needed to prove the theorem are (5.40) and (5.41), which have been reduced to the uniform bound in L on $E[\sup_{0\leqslant t\leqslant T}|\widetilde{Z}^L_t|^2]$. This is the result of Proposition 5.2, proved in Section 5.2. □

We finally have the tightness result.

Theorem 5.8. *Under the assumption that* $Z^L_0 = \left([La^1_0],[La^2_0],...,[La^k_0]\right)$ *for a fixed* $a_0 = (a^1_0,a^2_0,...,a^k_0)\in\mathbb{R}^k_+$*, the family of random processes* $(\widetilde{Z}^L_t)_{t\geqslant 0}$ *as defined in* (3.20) *indexed by* $L>1$ *is tight and its limit is continuous.*

Proof. Under the assumption of Theorem (5.8), condition (5.1) is guaranteed and Theorem (5.7) also holds. Theorem (5.7) says condition (5.3) holds, which is equivalent to condition (5.2). Condition (5.1) and (5.2) in together proves that $(\widetilde{Z}^{L,i}_t)_{0\leqslant t\leqslant T}$ is tight for any fixed $T>0$ and its limit is continuous. Since $T>0$ is arbitrary, this implies the tightness of $(\widetilde{Z}^{L,i}_t)_{t\geqslant 0}$ and thus the tightness of $(\widetilde{Z}^L_t)_{t\geqslant 0}$. □

Chapter 6

Scaling Limit

Note that we have shown the tightness of process $(\widetilde{Z}_t^L)_{t\geqslant 0}$, which means that for any subsequence (in L) of $(\widetilde{Z}_t^L)_{t\geqslant 0}$ we have a further subsequence that is convergent in distribution. We want show that $(\widetilde{Z}_t^L)_{t\geqslant 0}$ itself is convergent in distribution and it has a unique continuous limit process which solves an ODE in the classical sense. First, we show that $(\widetilde{Z}_t^L)_{0\leq t\leq T}$ must satisfy the following proposition.

Proposition 6.1. $\forall\, T > 0$, *for a test function* $f \in C_c^2(\mathbb{R}^k, \mathbb{R})$, *we have*

$$\lim_{L\to\infty} E\left[\sup_{0\leqslant t\leqslant T}\left|f(\widetilde{Z}_t^L) - f(\widetilde{Z}_0^L) - \int_0^t L|\widetilde{Z}_{s-}^L| \sum_{w\in J} p_w\left(\frac{\widetilde{Z}_{s-}^L}{\left|\widetilde{Z}_{s-}^L\right|}\right)\left[f\left(\widetilde{Z}_{s-}^L + \frac{w}{L}\right) - f(\widetilde{Z}_{s-}^L)\right] ds\right|\right] = 0\,. \tag{6.1}$$

Proof. $\forall\, 0 \leqslant t \leqslant T$, and a test function $f \in C_c^2(\mathbb{R}^k, \mathbb{R})$, we have by Ito's formula that

$$\begin{aligned}
& f(\widetilde{Z}_t^L) - f(\widetilde{Z}_0^L) \\
= & \int_0^t L|\widetilde{Z}_{s-}^L| \sum_{w \in J} \Pi\left(L\widetilde{Z}_{s-}^L, L\widetilde{Z}_{s-}^L + w\right)\left[f\left(\widetilde{Z}_{s-}^L + \frac{w}{L}\right) - f(\widetilde{Z}_{s-}^L)\right] ds + \widetilde{\mathscr{M}}_t^{f,L} \\
= & \int_0^t L|\widetilde{Z}_{s-}^L| \sum_{w \in J} \left[\Pi\left(L\widetilde{Z}_{s-}^L, L\widetilde{Z}_{s-}^L + w\right) - p_w\left(\frac{\widetilde{Z}_{s-}^L}{|\widetilde{Z}_{s-}^L|}\right)\right]\left[f\left(\widetilde{Z}_{s-}^L + \frac{w}{L}\right) - f(\widetilde{Z}_{s-}^L)\right] ds \\
+ & \int_0^t L|\widetilde{Z}_{s-}^L| \sum_{w \in J} p_w\left(\frac{\widetilde{Z}_{s-}^L}{|\widetilde{Z}_{s-}^L|}\right)\left[f\left(\widetilde{Z}_{s-}^L + \frac{w}{L}\right) - f(\widetilde{Z}_{s-}^L)\right] ds + \widetilde{\mathscr{M}}_t^{f,L}.
\end{aligned} \tag{6.2}$$

By (3.22), we have $\left|\Pi\left(L\widetilde{Z}_{s-}^L, L\widetilde{Z}_{s-}^L + w\right) - p_w\left(\frac{\widetilde{Z}_{s-}^L}{|\widetilde{Z}_{s-}^L|}\right)\right| \leqslant \frac{a}{L|\widetilde{Z}_{s-}^L|}$ and thus

$$\begin{aligned}
& \left|f(\widetilde{Z}_t^L) - f(\widetilde{Z}_0^L) - \int_0^t L|\widetilde{Z}_{s-}^L| \sum_{w \in J} p_w(\frac{\widetilde{Z}_{s-}^L}{|\widetilde{Z}_{s-}^L|})\left[f\left(\widetilde{Z}_{s-}^L + \frac{w}{L}\right) - f(\widetilde{Z}_{s-}^L)\right] ds\right| \\
\leqslant & \left|\int_0^t L|\widetilde{Z}_{s-}^L| \sum_{w \in J} \left[\Pi\left(L\widetilde{Z}_{s-}^L, L\widetilde{Z}_{s-}^L + w\right) - p_w(\frac{\widetilde{Z}_{s-}^L}{|\widetilde{Z}_{s-}^L|})\right]\left[f\left(\widetilde{Z}_{s-}^L + \frac{w}{L}\right) - f(\widetilde{Z}_{s-}^L)\right] ds\right| \\
+ & \left|\widetilde{\mathscr{M}}_t^{f,L}\right| \\
\leqslant & \left|\int_0^t L|\widetilde{Z}_{s-}^L| \cdot \sum_{w \in J} \frac{a}{L|\widetilde{Z}_{s-}^L|} \cdot \left[f\left(\widetilde{Z}_{s-}^L + \frac{w}{L}\right) - f(\widetilde{Z}_{s-}^L)\right] ds\right| + \left|\widetilde{\mathscr{M}}_t^{f,L}\right| \\
= & a\left|\int_0^t \sum_{w \in J} \left[f\left(\widetilde{Z}_{s-}^L + \frac{w}{L}\right) - f(\widetilde{Z}_{s-}^L)\right] ds\right| + \left|\widetilde{\mathscr{M}}_t^{f,L}\right|.
\end{aligned} \tag{6.3}$$

Therefore,

$$\begin{aligned}
& E \sup_{0 \leqslant t \leqslant T}\left[\left|f(\widetilde{Z}_t^L) - f(\widetilde{Z}_0^L) - \int_0^t L|\widetilde{Z}_{s-}^L| \sum_{w \in J} p_w(\frac{\widetilde{Z}_{s-}^L}{|\widetilde{Z}_{s-}^L|})\left[f\left(\widetilde{Z}_{s-}^L + \frac{w}{L}\right) - f(\widetilde{Z}_{s-}^L)\right] ds\right|\right] \\
\leqslant & a\, E\left[\sup_{0 \leqslant t \leqslant T}\left|\int_0^t \sum_{w \in J} \left[f\left(\widetilde{Z}_{s-}^L + \frac{w}{L}\right) - f(\widetilde{Z}_{s-}^L)\right] ds\right|\right] + E\left[\sup_{0 \leqslant t \leqslant T}\left|\widetilde{\mathscr{M}}_t^{f,L}\right|\right].
\end{aligned} \tag{6.4}$$

We need to estimate

$$E\left[\sup_{0\leqslant t\leqslant T}\left|\int_0^t \sum_{w\in J}\left[f\left(\widetilde{Z}^L_{s-}+\frac{w}{L}\right)-f(\widetilde{Z}^L_{s-})\right]ds\right|\right], \tag{6.5}$$

$$E\left[\sup_{0\leqslant t\leqslant T}\left|\widetilde{\mathscr{M}}_t^{f,L}\right|\right] \tag{6.6}$$

and their limits as $L\to\infty$.

By (6.4), Lemma (6.2) and Lemma (6.3), which are proven immediately after, we have that

$$\begin{aligned}&\lim_{L\to\infty} E\sup_{0\leqslant t\leqslant T}\left|f(\widetilde{Z}^L_t)-f(\widetilde{Z}^L_0)-\int_0^t L|\widetilde{Z}^L_{s-}|\sum_{w\in J}p_w\left(\frac{\widetilde{Z}^L_{s-}}{|\widetilde{Z}^L_{s-}|}\right)\left[f\left(\widetilde{Z}^L_{s-}+\frac{w}{L}\right)-f(\widetilde{Z}^L_{s-})\right]ds\right|\\ &\leqslant \lim_{L\to\infty} aE\left[\sup_{0\leqslant t\leqslant T}\left|\int_0^t\sum_{w\in J}\left[f\left(\widetilde{Z}^L_{s-}+\frac{w}{L}\right)-f(\widetilde{Z}^L_{s-})\right]ds\right|\right]+\lim_{L\to\infty}E\left[\sup_{0\leqslant t\leqslant T}\left|\widetilde{\mathscr{M}}_t^{f,L}\right|\right]=0.\end{aligned} \tag{6.7}$$

This finishes the proof of Proposition (6.1). □

Lemma 6.2.

$$\lim_{L\to\infty}E\left[\sup_{0\leqslant t\leqslant T}\left|\int_0^t\sum_{w\in J}\left[f\left(\widetilde{Z}^L_{s-}+\frac{w}{L}\right)-f(\widetilde{Z}^L_{s-})\right]ds\right|\right]=0.$$

Proof. Note that $f\in C^2_c(\mathbb{R}^k)$, which implies that f is global Lipschitz and there is a constant $K_f>0$ depending on f such that $\forall x,y\in\mathbb{R}^k$, $|f(x)-f(y)|\leqslant K_f|x-y|$.
It follows that $\left|\int_0^t a\sum_{w\in J}\left[f\left(\widetilde{Z}^L_{s-}+\frac{w}{L}\right)-f(\widetilde{Z}^L_{s-})\right]ds\right| \leqslant \left|\int_0^t a\sum_{w\in J}\frac{K_f w}{L}ds\right|$. Recall that the number of possible jumps $w\in J$ is finite, bounded by a positive number N that is independent of L and each jump $|w|\leqslant m$, we have

$$E\sup_{0\leqslant t\leqslant T}\left|\int_0^t a\sum_{w\in J}\left[f\left(\widetilde{Z}^L_{s-}+\frac{w}{L}\right)-f(\widetilde{Z}^L_{s-})\right]ds\right|\leqslant\frac{aK_f mNT}{L}. \tag{6.8}$$

Therefore,

$$\lim_{L\to\infty} E\left[\sup_{0\leqslant t\leqslant T}\left|\int_0^t a\sum_{w\in J}\left[f\left(\widetilde{Z}^L_{s-}+\frac{w}{L}\right)-f(\widetilde{Z}^L_{s-})\right]ds\right|\right]\leqslant \lim_{L\to\infty}\frac{aK_f mNT}{L}=0. \tag{6.9}$$

□

Lemma 6.3.

$$\lim_{L\to\infty} E\left[\sup_{0\leqslant t\leqslant T}\left|\widetilde{\mathscr{M}}^{f,L}_t\right|\right]=0.$$

Proof. To estimate $E\left[\sup_{0\leqslant t\leqslant T}\left|\widetilde{\mathscr{M}}^{f,L}_t\right|\right]$, we work on $E\left[\sup_{0\leqslant t\leqslant T}\left|\widetilde{\mathscr{M}}^{f,L}_t\right|^2\right]$ first.

$$\begin{aligned}
&E\left[\sup_{0\leqslant t\leqslant T}\left(\widetilde{\mathscr{M}}^{f,L}_t\right)^2\right]\\
\leqslant&4E\left(\widetilde{\mathscr{M}}^{f,L}_T\right)^2 \qquad\qquad \text{(by Doob's maximal inequality where p=2)}\\
=&4E\int_0^T L|\widetilde{Z}^L_{s-}|\sum_{w\in J}\Pi\left(L\widetilde{Z}^L_{s-},L\widetilde{Z}^L_{s-}+w\right)\left[f\left(\widetilde{Z}^L_{s-}+\frac{w}{L}\right)-f(\widetilde{Z}^L_{s-})\right]^2ds\\
\leqslant&4E\int_0^T L|\widetilde{Z}^L_{s-}|\sum_{w\in J}K^2\frac{|w|^2}{L^2}ds\\
\leqslant&\frac{4K^2m^2N}{L}E\int_0^T|\widetilde{Z}^L_{s-}|ds.
\end{aligned}$$

Recall that we already proved that $E\left[\sup_{0\leqslant s\leqslant T}|\widetilde{Z}^L_{s-}|^2\right]$ is uniformly bounded in $L\geq 1$, which implies that

$$\int_0^T E\left[\sup_{0\leqslant s\leqslant T}|\widetilde{Z}^L_{s-}|^2\right]ds=E\int_0^T\left[\sup_{0\leqslant s\leqslant T}|\widetilde{Z}^L_{s-}|^2\right]ds\geqslant E\int_0^T\left[|\widetilde{Z}^L_{s-}|^2\right]ds.$$

Therefore,

$$E\int_0^T |\widetilde{Z}^L_{s-}|ds \leqslant E\int_0^T \left[|\widetilde{Z}^L_{s-}|^2+1\right]ds$$

is bounded, say by M_1, for all L. We can then conclude that

$$E\sup_{0\leqslant s\leqslant T}\left|\widetilde{\mathscr{M}}^{f,L}_t\right| \leqslant \left[E\sup_{0\leqslant s\leqslant T}\left(\mathscr{M}^{f,L}_t\right)^2\right]^{\frac{1}{2}} \leqslant 2Km\left(\frac{NM_1}{L}\right)^{\frac{1}{2}}. \tag{6.10}$$

which proves that

$$\lim_{L\to\infty}E\left[\sup_{0\leqslant t\leqslant T}\left|\widetilde{\mathscr{M}}^{f,L}_t\right|\right] \leqslant \lim_{L\to\infty}2Km\left(\frac{NM_1}{L}\right)^{\frac{1}{2}} = 0.$$

□

By using a linear approximation of the test function $f \in C^2_c(\mathbb{R}^k,\mathbb{R})$, we can further obtain the following Proposition.

Proposition 6.4. *Under the same conditions as* (6.1)*, we further have that*

$$\lim_{L\to\infty}E\left[\sup_{0\leqslant t\leqslant T}\left|f(\widetilde{Z}^L_t)-f(\widetilde{Z}^L_0)-\int_0^t|\widetilde{Z}^L_{s-}|\sum_{w\in J}\left[p_w\left(\frac{\widetilde{Z}^L_{s-}}{|\widetilde{Z}^L_{s-}|}\right)\left(\sum_{i=1}^k\frac{\partial f}{\partial Z^i}(\widetilde{Z}_{s-})w_i\right)\right]ds\right|\right]=0. \tag{6.11}$$

Proof. For a test function $f \in C^2_c(\mathbb{R}^k,\mathbb{R})$, $\forall Z = (Z^1,Z^2,...,Z^k) \in \mathbb{R}^k$, we have $f(Z) = f(Z^1,Z^2,...,Z^k) \in \mathbb{R}$ and by Taylor's formula

$$f\left(\widetilde{Z}^L_{s-}+\frac{w}{L}\right)-f(\widetilde{Z}^L_{s-}) = \sum_{i=1}^k\frac{\partial f}{\partial Z^i}(\widetilde{Z}^L_{s-})\frac{w_i}{L}+\varepsilon_{f,L}(\widetilde{Z}^L_{s-},w)$$

where $\varepsilon_{f,L}(\widetilde{Z}^L_{s-},w)$ is an error part which satisfies that for large L and a constant depending on the test function f,

$$\left|\varepsilon_{f,L}(\widetilde{Z}^L_{s-},w)\right| \leqslant \frac{C_f}{L^2}. \tag{6.12}$$

Then

$$\begin{aligned}
& E\sup_{0\leqslant t\leqslant T}\left|f(\widetilde{Z}^L_t)-f(\widetilde{Z}^L_0)-\int_0^t L|\widetilde{Z}^L_{s-}|\sum_{w\in J}\left[p_w\left(\frac{\widetilde{Z}^L_{s-}}{\left|\widetilde{Z}^L_{s-}\right|}\right)\left(\sum_{i=1}^k\frac{\partial f}{\partial Z^i}\frac{w_i}{L}\right)\right]ds\right| \\
\leqslant \quad & E\sup_{0\leqslant t\leqslant T}\left|f(\widetilde{Z}^L_t)-f(\widetilde{Z}^L_0)-\int_0^t L|\widetilde{Z}^L_{s-}|\sum_{w\in J}p_w\left(\frac{\widetilde{Z}^L_{s-}}{\left|\widetilde{Z}^L_{s-}\right|}\right)\left[f\left(\widetilde{Z}^L_{s-}+\frac{w}{L}\right)-f(\widetilde{Z}^L_{s-})\right]ds\right| \\
+ \quad & E\sup_{0\leqslant t\leqslant T}\left|\int_0^t L|\widetilde{Z}^L_{s-}|\sum_{w\in J}p_w\left(\frac{\widetilde{Z}^L_{s-}}{\left|\widetilde{Z}^L_{s-}\right|}\right)\varepsilon_{f,L}(\widetilde{Z}^L_{s-},w)\right|.
\end{aligned}$$

Passing L to infinity, we can conclude from (6.12) and Proposition (6.1) that

$$\lim_{L\to\infty}E\sup_{0\leqslant t\leqslant T}\left|f(\widetilde{Z}^L_t)-f(\widetilde{Z}^L_0)-\int_0^t L|\widetilde{Z}^L_{s-}|\sum_{w\in J}\left[p_w\left(\frac{\widetilde{Z}^L_{s-}}{\left|\widetilde{Z}^L_{s-}\right|}\right)\left(\sum_{i=1}^k\frac{\partial f}{\partial Z^i}(\widetilde{Z}^L_{s-})\frac{w_i}{L}\right)\right]ds\right| = 0. \tag{6.13}$$

This finishes the proof of Proposition (6.4). □

We want to show that a similar result to that in proposition (6.4) also holds for any limit process, which requires that the functional involved is continuous and bounded. However, despite it is indeed continuous, it is not bounded due to the fact that there is an unbounded term under integration. To proceed, we first prove the following lemma for the case in which the boundedness is no longer an issue.

Lemma 6.5. *Let $\eta \in D([0,T],\mathbb{R}^k)$, $g \in C_b(\mathbb{R}^k,\mathbb{R})$ and $f \in C^1_b(\mathbb{R}^k,\mathbb{R})$ with bounded continuous first-order partial derivatives. Then the functional from $D[0,T]$ to $\mathbb{R}$ defined by*

$$\sup_{0\leqslant t\leqslant T}\left|f(\eta(t))-f(\eta(0))-\int_0^t g(\eta(s_-))ds\right| \tag{6.14}$$

is continuous and bounded. In particular, f can be a $C^1_c(\mathbb{R}^k,\mathbb{R})$ function.

Proof. Define

$$\Phi_\eta(t) = f(\eta(t)) - f(\eta(0)) - \int_0^t g(\eta(s_-))ds$$

where $t \in [0, T]$ and η, f are as given in the lemma. If we can show that

$$\eta \mapsto \Phi_\eta \text{ is a continuous map } D([0,T],\mathbb{R}^k) \to D([0,T],\mathbb{R}) \text{ in the Skorohod J1 topology,} \tag{I}$$

and

$$q : D([0,T],\mathbb{R}) \to \mathbb{R} \text{ where } q(\eta) = \sup_{0\leqslant t\leqslant T} |\eta(t)| \text{ is continuous,} \tag{II}$$

then $q(\Phi_\eta) : D([0,T],\mathbb{R}^k) \to \mathbb{R}$ is continuous as the composition of two continuous function. To make notations clear, despite we commented that we will use L_1 norm for $\mathbb{R}^k$ throughout this paper and for simplicity we use $|\cdot|$ instead of $||\cdot||_1$, in this proof we will use $||\cdot||_1$ for L_1 norm in $\mathbb{R}^k$ and $|\cdot|$ for absolute value in $\mathbb{R}$.

Also let Λ_T denote the set of increasing homeomorphisms of $[0,T]$ onto itself as defined in (2.5) and $||\cdot||_S$ be the Skorokhod space J_1 metric defined in (2.6).

First we prove (I), that is, $\eta \mapsto \Phi_{\eta(t)}$ is continuous from $D([0,T],\mathbb{R}^k)$ to $D([0,T],\mathbb{R})$ in the Skorokhod topology J_1. For a sequence $\eta_n(t) \to \eta(t)$, we need to show that $\Phi_{\eta_n} \to \Phi_\eta$ in J_1. $\eta_n(t) \to \eta(t)$ implies that there exists a sequence of increasing homeomorphisms $\lambda_n(t)$ such that $\sup_{0\leqslant t\leqslant T} ||\eta_n(\lambda_n(t)) - \eta(t)||_1 \to 0$; it would be sufficient if for the same sequence $\lambda_n(t)$, we have $\sup_{0\leqslant t\leqslant T} |\Phi_{\eta_n}(\lambda_n(t)) - \Phi_\eta(t)| \to 0$. By definition, we have

$$\Phi_{\eta_n}(\lambda_n(t)) = f(\eta_n(\lambda_n(t))) - f(\eta_n(0)) - \int_0^{\lambda_n(t)} g(\eta_n(s_-))ds,$$

and direct calculations give

$$\begin{aligned}
&\sup_{0\leqslant t\leqslant T} |\Phi_{\eta_n}(\lambda_n(t)) - \Phi_\eta(t)| \\
&\leqslant 2 \sup_{0\leqslant t\leqslant T} |f(\eta_n(\lambda_n(t))) - f(\eta(t))| \\
&+ \sup_{0\leqslant t\leqslant T} \left| \int_0^{\lambda_n(t)} g(\eta_n(s_-))ds - \int_0^t g(\eta_n(s_-))ds \right| \\
&\leqslant 2 \sup \|\nabla f\|_1 \sup_{0\leqslant t\leqslant T} \|\eta_n(\lambda_n(t)) - \eta(t)\|_1 \\
&+ \sup_{0\leqslant t\leqslant T} \int_0^{\lambda_n(t)} |g(\eta_n(s_-)) - g(\eta(s_-))|\, ds + \sup_{0\leqslant t\leqslant T} \left| \int_{\lambda_n(t)}^t g(\eta(s_-))ds \right|.
\end{aligned} \tag{6.15}$$

Since we assumed that f has continuous bounded first-order partial derivatives, $\|\nabla f\|_1$ is bounded. Also g is assumed to be bounded, continuing (6.15) we get

$$\begin{aligned}
&\sup_{0\leqslant t\leqslant T} |\Phi_{\eta_n}(\lambda_n(t)) - \Phi_\eta(t)| \\
&\leqslant 2 \sup \|\nabla f\|_1 \sup_{0\leqslant t\leqslant T} \|\eta_n(\lambda_n(t)) - \eta(t)\|_1 + \sup_{0\leqslant t\leqslant T} |g(\eta(t_-))| \sup_{0\leqslant t\leqslant T} |\lambda_n(t) - t| \\
&+ \sup_{0\leqslant t\leqslant T} \int_0^{\lambda_n(t)} |g(\eta_n(s_-)) - g(\eta(s_-))| ds.
\end{aligned} \tag{6.16}$$

For the first two terms we have $\sup_{t\in[0,T]} |\lambda_n(t) - t| \to 0$ and $\sup_{t\in[0,T]} \|\eta_n(\lambda_n(t)) - \eta(t)\|_1 \to 0$ as $n \to \infty$; and for the last term we can use dominated convergence theorem since $\eta_n(s) \to \eta(s)$ at the points of continuity of $\eta(s)$. Letting n go to infinity, we get

$$\lim_{n\to\infty} \sup_{0\leqslant t\leqslant T} |\Phi_{\eta_n}(\lambda_n(t)) - \Phi_\eta(t)| = 0, \tag{6.17}$$

which proves (I). Now we proceed to prove (II), that is,

$q : D([0, T], \mathbb{R}) \to \mathbb{R}$ where $q(\eta) = \sup_{0\leqslant t\leqslant T} |\eta(t)|$ is continuous.

Note that $\forall \eta(t), \eta_n(t) \in D([0,T],\mathbb{R})$, we have

$$
\begin{aligned}
|q(\eta) - q(\eta')| &= \left| \sup_{0\leqslant t\leqslant T} |\eta(t)| - \sup_{0\leqslant t\leqslant T} |\eta'(t)| \right| \\
&= \left| \sup_{0\leqslant t\leqslant T} |\eta(t)| - \sup_{0\leqslant t\leqslant T} |\eta'(\lambda(t))| \right| \\
&\leqslant \sup_{0\leqslant t\leqslant T} ||\eta(t)| - |\eta'(\lambda(t))|| \\
&\leqslant \sup_{0\leqslant t\leqslant T} |\eta(t) - \eta'(\lambda(t))|
\end{aligned}
\tag{6.18}
$$

where $\lambda(t) \in \Lambda_T$ as defined in (2.5) . Taking infimum over all possible $\lambda(t)$ we conclude that

$$
|q(\eta) - q(\eta')| \leqslant \inf_{\lambda\in\Lambda_T} \sup_{0\leqslant t\leqslant T} |\eta(t) - \eta'(\lambda(t))| \leqslant \|\eta - \eta'\|_{S}.
$$

As both (I) and (II) are proved, we have the continuity of $q(\Phi_\eta)$. Boundedness of $q(\Phi_\eta)$ is readily seen since we assumed that both g and f are bounded. Note that functions in $C_c^1(\mathbb{R}^k,\mathbb{R})$ are bounded continuous which also have bounded continuous derivatives, $f \in C_c^1(\mathbb{R}^k,\mathbb{R})$ would be sufficient for the above proof. This finishes the proof of Lemma (6.5).

□

Next we will show that the part under integration in proposition (6.4) is Lipschitz, which implies continuity, and also a corresponding ODE system has a unique classical solution. First note that if Lipschitz maps p_w as defined in (3.12) that satisfy (3.14) exist, we can extend them from the $k-1$ simplex $\triangle^{k-1}$ as defined in (3.13) to the $k-1$ sphere S^{k-1} in the ℓ_1 norm in $\mathbb{R}^k$, namely

$$
S^{k-1} = \left\{ (x_1, ..., x_k) \in \mathbb{R}^k \middle| \sum_{i=1}^{k} |x_i| = 1, x_i \geqslant 0, \forall i \in \{1, ..., k\} \right\} \tag{6.19}
$$

such that $\forall (x_1, ..., x_k) \in S^{k-1}$, an extended $\widetilde{p}_w$ is defined by

$$\widetilde{p}_w(x_1, ..., x_k) = p_w(|x_1|, ..., |x_k|) \quad \forall w \in J. \tag{6.20}$$

These extended $\widetilde{p}_w$ coincide with p_w in $\triangle^{k-1}$ and remain Lipschitz; moreover, $\forall w \in J$ and $\forall X = (x_1, ..., x_k), X' = (x_1^{'}, ..., x_k^{'}) \in \triangle^{k-1}$ we also have

$$\begin{aligned}
&\left|\widetilde{p}_w(x_1, ..., x_k) - \widetilde{p}_w(x_1^{'}, ..., x_k^{'})\right| \\
&= \left|p_w(|x_1|, ..., |x_k|) - p_w(|x_1^{'}|, ..., |x_k^{'}|)\right| \\
&\leqslant C_p \sum_{i=1}^{k} \left||x_i| - |x_i^{'}|\right| \\
&\leqslant C_p \sum_{i=1}^{k} \left|x_i - x_i^{'}\right| \\
&= C_p \, |X - X'|
\end{aligned} \tag{6.21}$$

where C_p is the same constant as in (3.15). We will use p_w in the extended sense as $\widetilde{p}_w$ thereafter. Let $Z = (Z^1, ..., Z^k) \in \mathbb{R}^k$ where $k \geqslant 2$ is a given integer and w, J as defined in (1.7), define

$$\begin{aligned}
g_i(Z) &= g_i(Z^1, ..., Z^k) \\
&= \begin{cases} |Z| \sum_{w \in J} p_w\left(\dfrac{Z}{|Z|}\right) w_i, & \text{if } Z \neq 0 \\ 0, & \text{if } Z = 0, \end{cases}
\end{aligned} \tag{6.22}$$

for $i = 1, 2, ..., k$, and

$$g(Z) = (g_1(Z), ..., g_k(Z)), \tag{6.23}$$

we claim the following.

Proposition 6.6. *Let $g_i(Z)$ where $Z = (Z^1, ..., Z^k) \in \mathbb{R}^k$ be as given in* (6.22) *for $i = 1, 2, ..., k$, then each $g_i(Z)$ is Lipschitz in Z and the ODE system defined by*

$$\frac{dZ_t^i}{dt} = g_i(Z_t),$$

where $Z_t = (Z_t^1, Z_t^2, ..., Z_t^k) \in \mathbb{R}^k$ with initial condition $Z_0 = a_0 = (a_0^1, a_0^2, ..., a_0^k) \in \mathbb{R}_+^k$ has a unique classical solution for $t \in [0, \infty)$.

Proof. Using the given notations, we have

$$\begin{aligned} g_i(Z) &= |Z| \sum_{w \in J} p_w(\frac{Z}{|Z|}) w_i \\ &= (\sum_{i=1}^k |Z^i|) \sum_{w \in J} p_w(\frac{Z^1}{|Z|}, ..., \frac{Z^k}{|Z|}) w_i, \end{aligned}$$

for nonzero $Z \in \mathbb{R}^k$. To guarantee the existence and uniqueness of a classical solution to the ODE system, it suffices to show that each $g_i(Z)$ is Lipschitz in $Z^1, ..., Z^k$. Without loss of generality, we only need to show it is so for g_1.

$\forall Z = (Z^1, ..., Z^k), Z_* = (Z_*^1, ..., Z_*^k) \in \mathbb{R}^k$, if one of them, say $Z_* = 0$, then since the number of jumps $w \in J$ is finite, denoted by $|J|$, and each jump w satisfies $|w_i| \leqslant |w| \leqslant m$ for a positive number m and p_w takes values in $[0, 1]$, we have

$$|g_1(Z) - g_1(Z^*)| = \left| g_1(Z^1, ..., Z^k) \right| \leqslant m|J||Z| = m|J||Z - Z^*|.$$

If both Z and Z_* are nonzero,

$$
\begin{aligned}
&\left|g_1(Z^1,...,Z^k) - g_1(Z_*^1,...,Z_*^k)\right| \\
&= \left|\left(\sum_{j=1}^{k}|Z^j|\right)\sum_{w\in J} p_w\left(\frac{Z^1}{|Z|},...,\frac{Z^k}{|Z|}\right)w_1 - \left(\sum_{j=1}^{k}|Z_*^j|\right)\sum_{w\in J} p_w\left(\frac{Z_*^1}{|Z_*|},...,\frac{Z_*^k}{|Z_*|}\right)w_1\right| \\
&\leqslant \left(\sum_{j=1}^{k}\left||Z^j| - |Z_*^j|\right|\right)\left|\sum_{w\in J} p_w\left(\frac{Z^1}{|Z|},...,\frac{Z^k}{|Z|}\right)w_1\right| \\
&+ \left(\sum_{j=1}^{k}|Z_*^j|\right)\left|\sum_{w\in J}\left(p_w\left(\frac{Z^1}{|Z|},...,\frac{Z^k}{|Z|}\right) - p_w\left(\frac{Z_*^1}{|Z_*|},...,\frac{Z_*^k}{|Z_*|}\right)\right)w_1\right| \\
&\leqslant m|J||Z - Z_*| + m|Z_*|\left|\sum_{w\in J}\left(p_w\left(\frac{Z^1}{|Z|},...,\frac{Z^k}{|Z|}\right) - p_w\left(\frac{Z_*^1}{|Z_*|},...,\frac{Z_*^k}{|Z_*|}\right)\right)\right|.
\end{aligned}
$$

By (6.21), We have

$$
\begin{aligned}
&\left|\sum_{w\in J}\left(p_w\left(\frac{Z^1}{|Z|},...,\frac{Z^k}{|Z|}\right) - p_w\left(\frac{Z_*^1}{|Z_*|},...,\frac{Z_*^k}{|Z_*|}\right)\right)\right| \\
&\leqslant |J|C_p\left(\sum_{j=1}^{k}\left|\frac{Z^j}{|Z|} - \frac{Z_*^j}{|Z_*|}\right|\right) \\
&\leqslant |J|C_p(|Z||Z_*|)^{-1}\left(\sum_{j=1}^{k}\left||Z_*|Z^j - |Z|Z_*^j\right|\right).
\end{aligned}
$$

Since

$$
\begin{aligned}
\left||Z_*|Z^j - |Z|Z_*^j\right| &= \left|(|Z_*| - |Z| + |Z|)Z^j - |Z|(Z_*^j - Z^j + Z^j)\right| \\
&= \left|Z^j(|Z_*| - |Z|) + |Z|(Z^j - Z_*^j)\right| \\
&\leqslant 2|Z|\sum_{j=1}^{k}\left|Z_*^j - Z^j\right| \\
&= 2|Z||Z - Z_*|,
\end{aligned}
$$

we get

$$\begin{aligned}
&\left|\sum_{w\in J}\left(p_w\left(\frac{Z^1}{|Z|},...,\frac{Z^k}{|Z|}\right)-p_w\left(\frac{Z_*^1}{|Z_*|},...,\frac{Z_*^k}{|Z_*|}\right)\right)\right| \\
\leqslant&|J|C_p(|Z||Z_*|)^{-1}\left(2k\,|Z|\Big|Z_*-Z\Big|\right) \\
=&2|J|kC_p|Z_*|^{-1}\Big|Z_*-Z\Big|,
\end{aligned}$$

and further

$$\begin{aligned}
&\left|g_1(Z^1,...,Z^k)-g_1(Z_*^1,...,Z_*^k)\right| \\
\leqslant& m|J||Z-Z_*|+m|Z_*|\left|\sum_{w\in J}\left(p_w\left(\frac{Z^1}{|Z|},...,\frac{Z^k}{|Z|}\right)-p_w\left(\frac{Z_*^1}{|Z_*|},...,\frac{Z_*^k}{|Z_*|}\right)\right)\right| \\
\leqslant& m|J|(1+2kC_p)\Big|Z_*-Z\Big|,
\end{aligned}$$

which means that g_1 is indeed Lipschitz in $Z^1,...,Z^k$ and our lemma is proved. □

Note that J and w in Proposition (6.6) are as mentioned in (1.7), however, examining the proof, one can see that all we need is the fact that $|J|$ is finite and $|w_i|\leq m$ for all $i=1,...,k.$, but not the part that at most two components of a jump can be nonzero. Let $f\in C_b^1(\mathbb{R}^k,\mathbb{R})$ such that $f'\in C_b(\mathbb{R}^k,\mathbb{R})$ as in Lemma (6.5), then $\left|\sum_{i=1}^{k}\frac{\partial f}{\partial Z^i}(Z)w_i\right|$ is uniformly bounded for a fixed f, by similar reasoning as in Proposition (6.6) we can obtain the following result.

Lemma 6.7. *Let $Z=(Z^1,...,Z^k)\in\mathbb{R}_+^k$ and w, J, p_w are the same as given in our setup. For $f\in C_b^1(\mathbb{R}^k,\mathbb{R})$ s.t. $f'\in C_b(\mathbb{R}^k,\mathbb{R})$, define*

$$\begin{aligned}
g_f(Z)=&g_f(Z^1,...,Z^k) \\
=&\begin{cases}|Z|\sum_{w\in J}\left[p_w(\frac{Z}{|Z|})\left(\sum_{i=1}^{k}\frac{\partial f}{\partial Z^i}(Z)w_i\right)\right], & \text{if } Z\neq 0 \\ 0, & \text{if } Z=0,\end{cases}
\end{aligned}\tag{6.24}$$

then $g_f(Z)$ is Lipschitz and thus continuous.

Now we are ready to show that any limit process of $\left(\widetilde{Z}_t^L\right)_{0\leqslant t\leqslant T}$ must satisfy a relation similar to that in proposition (6.4).

Proposition 6.8. *Any limit point of $\left(\widetilde{Z}_t^L\right)_{0\leqslant t\leqslant T}$, denoted by $\left(\widetilde{Z}_t\right)_{0\leqslant t\leqslant T}$, satisfies*

$$E\left[\sup_{0\leqslant t\leqslant T}\left|f(\widetilde{Z}_t)-f(\widetilde{Z}_0)-\int_0^t|\widetilde{Z}_s|\sum_{w\in J}\left[p_w\left(\frac{\widetilde{Z}_s}{\left|\widetilde{Z}_s\right|}\right)\left(\sum_{i=1}^k\frac{\partial f}{\partial Z^i}(\widetilde{Z}_s)w_i\right)\right]ds\right|\right]=0, \tag{6.25}$$

where $f\in C_c^2(\mathbb{R}^k,\mathbb{R})$.

Proof. For $M\geqslant 0$, define for $z\in\mathbb{R}$

$$h_M(z)=\begin{cases}1, & |z|\leqslant M\\ 0, & |z|\geqslant M+1\\ -(|z|-M)+1, & M<z<M+1\end{cases}$$

and define for $\eta(\cdot)\in D\left([0,T],\mathbb{R}^k\right)$,

$$H_M\left(\eta(\cdot)\right)=\sup_{0\leqslant t\leqslant T}\left|f\left(\eta(t)\right)-f\left(\eta(0)\right)-\int_0^t h_M\left(|\eta(s_-)|)g_f\left(\eta(s_-)\right)ds\right)\right|, \tag{6.26}$$

where for a fixed $f\in C_c^2(\mathbb{R}^k,\mathbb{R})$, g_f is the same as defined in Lemma (6.7) and

$$g_f\left((\eta(s_-)\right)=|\eta(s_-)|\sum_{w\in J}p_w\left(\frac{\eta(s_-)}{|\eta(s_-)|}\right)\left(\sum_{i=1}^k\frac{\partial f}{\partial Z^i}(\eta(s_-))w_i\right). \tag{6.27}$$

For $\left(\widetilde{Z}_t^L\right)_{0\leqslant t\leqslant T} \in D\left([0,T],\mathbb{R}^k\right)$, we have

$$\begin{aligned} H_M(\widetilde{Z}^L(\cdot)) &= \sup_{0\leqslant t\leqslant T}\left| f\left(\widetilde{Z}_t^L\right) - f\left(\widetilde{Z}_0^L\right) - \int_0^t h_M\left(|\widetilde{Z}_{s-}^L|\right) g_f\left(\widetilde{Z}_{s-}^L\right) ds\right| \\ &\leqslant \sup_{0\leqslant t\leqslant T}\left| f\left(\widetilde{Z}_t^L\right) - f\left(\widetilde{Z}_0^L\right) - \int_0^t g_f\left(\widetilde{Z}_{s-}^L\right) ds\right| \\ &\quad + \sup_{0\leqslant t\leqslant T}\left|\int_0^t \left(1 - h_M\left(|\widetilde{Z}_{s-}^L|\right)\right) g_f\left(\widetilde{Z}_{s-}^L\right)\right|. \end{aligned} \tag{6.28}$$

Taking expectation on both sides, we get

$$\begin{aligned} & E\sup_{0\leqslant t\leqslant T}\left| f\left(\widetilde{Z}_t^L\right) - f\left(\widetilde{Z}_0^L\right) - \int_0^t h_M\left(|\widetilde{Z}_{s-}^L|\right) g_f\left(\widetilde{Z}_{s-}^L\right) ds\right| \\ \leqslant\ & E\sup_{0\leqslant t\leqslant T}\left| f\left(\widetilde{Z}_t^L\right) - f\left(\widetilde{Z}_0^L\right) - \int_0^t g_f\left(\widetilde{Z}_{s-}^L\right) ds\right| \\ +\ & E\sup_{0\leqslant t\leqslant T}\left|\int_0^t \left(1 - h_M\left(|\widetilde{Z}_{s-}^L|\right)\right) g_f\left(\widetilde{Z}_{s-}^L\right)\right|. \end{aligned} \tag{6.29}$$

By proposition (6.4), we have

$$\lim_{L\to\infty} E\sup_{0\leqslant t\leqslant T}\left| f\left(\widetilde{Z}_t^L\right) - f\left(\widetilde{Z}_0^L\right) - \int_0^t g_f\left(\widetilde{Z}_{s-}^L\right) ds\right| = 0. \tag{6.30}$$

Note that $f \in C_c^2$, we know each $\dfrac{\partial f}{\partial Z^i}$ is bounded for a fixed f, and since $|w_i| \leqslant m$, $p_w \leqslant 1$, we know $\sum\limits_{w\in J} p_w\left(\frac{\widetilde{Z}_{s_-}^L}{|\widetilde{Z}_{s_-}^L|}\right)\left(\sum\limits_{i=1}^k \frac{\partial f}{\partial Z^i}(\widetilde{Z}_{s_-}^L)w_i\right)$ is bounded by some constant C_f that doesn't depend on L, we have that

$$\left|\left(1 - h_M\left(|\widetilde{Z}_{s-}^L|\right)\right) g_f\left(\widetilde{Z}_{s-}^L\right)\right| \leqslant C_f \sup_{0\leqslant t\leqslant T}\left|\widetilde{Z}_{t-}^L\right| \sup_{0\leqslant s\leqslant T}\left|1 - h_M(|\widetilde{Z}_{s-}^L|)\right|.$$

Note that we have

$$0 \leqslant \mathbb{1}_{\{|\widetilde{Z}_{s-}^L|\leq M\}} \leqslant h_M(|\widetilde{Z}_{s-}^L|) \leqslant 1,$$

by definition of h_M, which implies

$$0 \leqslant 1 - h_M(|\widetilde{Z}^L_{s-}|) \leqslant \mathbb{1}_{\{|\widetilde{Z}^L_{s-}|>M\}} \leqslant \mathbb{1}_{\{\sup\limits_{0\leq s\leq T} |\widetilde{Z}^L_{s-}|>M\}}.$$

Therefore,

$$E\left[\sup_{0\leqslant t\leqslant T}\left|\int_0^t \left(1 - h_M\left(|\widetilde{Z}^L_{s-}|\right)\right) g_f\left(\widetilde{Z}^L_{s-}\right)\right|\right] \leqslant C_f T E\left[\sup_{0\leqslant t\leqslant T}\left|\widetilde{Z}^L_{t-}\right| \mathbb{1}_{\{\sup\limits_{0\leq s\leq T} |\widetilde{Z}^L_s| > M\}}\right]. \quad (6.31)$$

By Cauchy-Schwartz inequality,

$$\begin{aligned} & \left(E\left[\sup_{0\leqslant t\leqslant T}\left|\widetilde{Z}^L_{t-}\right| \mathbb{1}_{\{\sup\limits_{0\leq s\leq T} |\widetilde{Z}^L_s| > M\}}\right]\right)^2 \\ \leqslant \quad & E\left[\left(\sup_{0\leqslant t\leqslant T}\left|\widetilde{Z}^L_{t-}\right|\right)^2\right] E\left[\left(\mathbb{1}_{\{\sup\limits_{0\leq s\leq T} |\widetilde{Z}^L_s| > M\}}\right)^2\right] \quad (6.32) \\ = \quad & E\left[\sup_{0\leqslant t\leqslant T}\left|\widetilde{Z}^L_{t-}\right|^2\right] P(\sup_{0\leq s\leq T} |\widetilde{Z}^L_s| > M). \end{aligned}$$

Since $E\left[\sup\limits_{0\leqslant t\leqslant T} |\widetilde{Z}^L_{t^-}|^2\right]$ is uniformly bounded, say by $K > 0$, for all $L > 1$ according to Proposition (5.2), and by Chebyshev's inequality

$$P(\sup_{0\leq t\leq T} |\widetilde{Z}^L_{t-}| > M) \leqslant M^{-2} E\left[\left(\sup_{0\leqslant t\leqslant T} |\widetilde{Z}^L_{t^-}|\right)^2\right] \leqslant M^{-2} K. \quad (6.33)$$

By (6.32) and (6.33), we get from (6.31) that

$$E\left[\sup_{0\leqslant t\leqslant T}\left|\int_0^t \left(1 - h_M\left(|\widetilde{Z}^L_{s-}|\right)\right) g_f\left(\widetilde{Z}^L_{s-}\right)\right|\right] \leqslant C_f T K M^{-1}. \quad (6.34)$$

By (6.30) and (6.34), passing L to infinity we obtain from (6.29) that

$$\lim_{L\to\infty} E\left[\sup_{0\leqslant t\leqslant T}\left|f\left(\widetilde{Z}_t^L\right)-f\left(\widetilde{Z}_0^L\right)-\int_0^t h_M\left(|\widetilde{Z}_{s-}^L|\right)g_f\left(\widetilde{Z}_{s-}^L\right)ds\right|\right]\leqslant C_f TKM^{-1}, \tag{6.35}$$

or writing in a simpler form

$$\lim_{L\to\infty} E\left[H_M\left(\widetilde{Z}^L(\cdot)\right)\right]\leqslant C_f TKM^{-1}. \tag{6.36}$$

Note that for a fixed M, h_M is Lipschitz and bounded, g_f is Lipschitz by Lemma (6.7), then $h_M g_f$ is bounded and Lipschitz which is in $C_b(\mathbb{R}^k)$, also we assumed that $f\in C_c^2(\mathbb{R}^k,\mathbb{R})\subset C_c^1(\mathbb{R}^k,\mathbb{R})$, thus by Lemma (6.5) , $H_M\left(\eta(\cdot)\right)$ is a continuous and bounded functional. Therefore, any limit process $\left(\widetilde{Z}_t\right)_{0\leqslant t\leqslant T}$ of $\left(\widetilde{Z}_t^L\right)_{0\leqslant t\leqslant T}$ would also satisfy (6.36) , i.e.,

$$E\left[H_M\left(\widetilde{Z}(\cdot)\right)\right]\leqslant C_f TKM^{-1}. \tag{6.37}$$

Recall that the conditions (5.1) and (5.2) we proved in Chapter 5 are strong enough not only to guarantee tightness but also that a limit point belongs to $C([0,T],\mathbb{R}^k)$ almost surely, which must be in $D([0,T],\mathbb{R}^k)$ and this justifies our usage of the notation $H_M\left(\widetilde{Z}(\cdot)\right)$. Moreover, since a continuous function $\widetilde{Z}_t$ on a bounded interval $[0,T]$ must be bounded and $\widetilde{Z}_{t-}=\widetilde{Z}_t$, for $M\geqslant \operatorname*{ess\,sup}_{0\leq t\leq T}|\widetilde{Z}_t|$, by definition of h_M we know $h_M(|Z_{t-}|)=1$ a.s. and

$$H_M\left(\widetilde{Z}(\cdot)\right)=\sup_{0\leqslant t\leqslant T}\left|f\left(\widetilde{Z}_t\right)-f\left(\widetilde{Z}_0\right)-\int_0^t g_f\left(Z_s\right)ds\right|. \tag{6.38}$$

Passing M to infinity, it follows from (6.37) and (6.38) that

$$E\left[\sup_{0\leqslant t\leqslant T}\left|f\left(\widetilde{Z}_t\right)-f\left(\widetilde{Z}_0\right)-\int_0^t g_f\left(\widetilde{Z}_s\right)ds\right|\right]=0. \tag{6.39}$$

□

Note that we wanted $f \in C_c^2(\mathbb{R}^k, \mathbb{R})$ in proposition (6.8) since it is needed in proposition (6.4), which we cited to obtain result (6.30). However, this result actually holds for more general functions. For our purpose, we claim that it holds for f being projection maps.

Lemma 6.9. *A limit process* $\left(\widetilde{Z}_t\right)_{0\leqslant t\leqslant T}$ *of the family of processes* $\left(\widetilde{Z}_t^L\right)_{0\leqslant t\leqslant T}$ *indexed by L satisfies that*

$$\sup_{0\leqslant t\leqslant T}\left|\widetilde{Z}_t^i - \widetilde{Z}_0^i - \int_0^t g_i(\widetilde{Z}_s)ds\right| = 0,\ a.s. \tag{6.40}$$

for all $i = 1, ..., k$*, where* g_i *is as defined in* (6.22).

Proof. By equation (6.39), we have $\sup_{0\leqslant t\leqslant T}\left|f\left(\widetilde{Z}_t\right) - f\left(\widetilde{Z}_0\right) - \int_0^t g_f\left(\widetilde{Z}_s\right)ds\right| = 0$ almost surely . Let's say for a set $N_1 \subset \Omega$ with $P(N_1) = 0$, that we have $\sup_{0\leqslant t\leqslant T}\left|f\left(\widetilde{Z}_t\right) - f\left(\widetilde{Z}_0\right) - \int_0^t g_f\left(\widetilde{Z}_s\right)ds\right| = 0$ ω-by-ω on $\Omega \setminus N_1$. Recall that the two tightness condition we proved are strong enough to guarantee a continuous limit process a.s., there exists a set $N_2 \subset \Omega$ that $P(N_2) = 0$ and $(\widetilde{Z}_t)_{0\leq t\leq T}$ is continuous ω-by-ω on $\Omega \setminus N_2$. Then on the set $\Omega \setminus (N_1 \cup N_2)$, we have a continuous $(\widetilde{Z}_t)_{0\leq t\leq T}$ which satisfies $\sup_{0\leqslant t\leqslant T}\left|f\left(\widetilde{Z}_t\right) - f\left(\widetilde{Z}_0\right) - \int_0^t g_f\left(\widetilde{Z}_s\right)ds\right| = 0, \forall f \in C_c^2(\mathbb{R}^k, \mathbb{R})$. $\forall \omega' \in \Omega \setminus (N_1 \cup N_2)$, $(\widetilde{Z}_t(\omega'))$ is a continuous function of t on the interval $[0, T]$ and thus bounded by a constant $M_{\omega'}$ depending on ω', we can chose a C_c^2 function $f_{\omega'}$ such that $f_{\omega'}(z) = z_i$ for $|z| \leq M_{\omega'}$, and we have $\sup_{0\leqslant t\leqslant T}\left|\widetilde{Z}_t^i(\omega') - \widetilde{Z}_0^i(\omega') - \int_0^t g_i\left(\widetilde{Z}_s(\omega')\right)ds\right| = 0$. This can be done for all $\omega' \in \Omega \setminus (N_1 \cup N_2)$ and we claim on the set $\Omega \setminus (N_1 \cup N_2)$, which is of probability 1, equation (6.40) holds for any $i \in \{1, 2, ..., k\}$. This proves the lemma. □

By lemma 6.9, we can conclude that $\sup_{0\leqslant t\leqslant T}\left|\widetilde{Z}_t^i - \widetilde{Z}_0^i - \int_0^t |\widetilde{Z}_s| \sum_{w\in J}\left[p_w\left(\frac{\widetilde{Z}_s}{|\widetilde{Z}_s|}\right)w_i\right]\right| = 0$ with probability 1 and $\left(\widetilde{Z}_t\right)_{0\leqslant t\leqslant T}$ is a weak solution to the ODE system

$$\frac{d\widetilde{Z}_t^i}{dt} = g_i(\widetilde{Z}_t), \quad \text{i=1,2,...,k} \tag{6.41}$$

or writing in vector form

$$\frac{d\widetilde{Z}_t}{dt} = g(\widetilde{Z}_t), \tag{6.42}$$

where g is as defined in (6.23).

Recall that we first conclude in theorem (5.8) the tightness of the family of processes $(\widetilde{Z}_t^L)_{0\leqslant t\leqslant T}$, which means any subsequence of $(\widetilde{Z}_t^L)_{0\leqslant t\leqslant T}$ has a further subsequence that converges weakly. Then we showed that any limit of this kind must be a weak solution to the same ODE system (6.41). By proposition (6.6), this limit ODE has a unique strong solution, which implies that the weak solution must be the same as the strong solution and thus unique. One may ask does the original family of processes itself converge weakly to the solution of the limit ODE? Note that the space $D([0,T],\mathbb{R}^k)$ the family of processes $(\widetilde{Z}_t^L)_{0\leqslant t\leqslant T}$ lives in is a metric space induced by the Skorokhod J1 metric and the limit process is in $C([0,T],\mathbb{R}^k) \subset D([0,T],\mathbb{R}^k)$, the answer is affirmative by the following lemma.

Lemma 6.10. *Given a sequence $\{Z_n\}$ in a metric space (X,d), if every subsequence $\{Z_{n_k}\}$ of $\{Z_n\}$ has a further subsequence that converges to the same limit point $z \in X$, then the sequence $\{Z_n\}$ is convergent and it converges to z.*

Proof. Suppose the opposite that $\{Z_n\}$ is not convergent. In particular, it cannot converge to any point in X. Then $\forall x \in X, \exists \epsilon > 0$ s.t. $\exists$ subsequence $\{Z_{n_k}\}$ of $\{Z_n\}$ satisfying that $d(Z_{n_k}, x) > \epsilon, \forall n_k$. Now choose x to be the limit point z given, we must have $d(Z_{n'_k}, z) > 0$, $\forall\, n'_k$ for some subsequence $\{Z_{n'_k}\}$. But $\{Z_{n'_k}\}$ must have a further subsequence that converges to z by assumption, and we cannot have $d(Z_{n'_k}, z) > \epsilon$, $\forall\, n_k$. This is a contradiction. Thus $\{Z_n\}$ is convergent. Clearly z is a limit of $\{Z_n\}$ and now it is the limit. □

Therefore, we can conclude the main result of this chapter.

Proposition 6.11. *Given $0 < T < +\infty$, the family of processes $(\widetilde{Z}^L_t)_{0\leqslant t\leqslant T} \in D([0,T],\mathbb{R}^k)$ indexed by L, with initial conditions $\widetilde{Z}^L_0 = L^{-1}\left([La^1_0],[La^2_0],...,[La^k_0]\right)$ for a fixed $a_0 = (a^1_0,a^2_0,...,a^k_0) \in \mathbb{R}^k_+$, converges weakly to a unique limit process, denoted by $(\widetilde{Z}_t)_{0\leqslant t\leqslant T}$, that solves the initial value problem for the ODE*

$$\frac{d\widetilde{Z}_t}{dt} = |\widetilde{Z}_t| \sum_{w\in J} p_w \left(\frac{\widetilde{Z}_t}{|\widetilde{Z}_t|}\right) w, \quad \widetilde{Z}_0 = a_0,$$

in the classical sense.

Note that proposition (6.11) holds for any $T > 0$, which means $(\widetilde{Z}_t)_{0\leqslant t\leqslant T}$ satisfies the ODE with the given initial condition in any bounded interval $[0,T]$, and it must be a solution to the ODE for all $t \geq 0$.

Chapter 7

Weak Limit of $\widetilde{X}_t^L$

Note that we have shown the limit process $(\widetilde{Z}_t)_{t\geqslant 0}$ of $(\widetilde{Z}_t^L)_{t\geqslant 0}$, with initial conditions $\widetilde{Z}_0^L = L^{-1}\left([La_0^1], [La_0^2], ..., [La_0^k]\right)$ for some fixed real vector $a_0 = (a_0^1, a_0^2, ..., a_0^k) \in \mathbb{R}_+^k \setminus \{0\}$, is the solution to an ODE initial value problem. Next we want to investigate a new family of processes defined by

$$\widetilde{X}_t^L = \begin{cases} \dfrac{\widetilde{Z}_t^L}{|\widetilde{Z}_t^L|} \;, & \text{if } |\widetilde{Z}_t^L| > 0, \\ \\ 0 \;, & \text{if } |\widetilde{Z}_t^L| = 0. \end{cases} \tag{7.1}$$

We want to show that this family of "normalized" process is also tight, and characterize its limit behavior by an ODE. An intuitive guess is its limit might be

$$\widetilde{X}_t = \begin{cases} \dfrac{\widetilde{Z}_t}{|\widetilde{Z}_t|} \;, & \text{if } |\widetilde{Z}_t| > 0, \\ \\ 0 \;, & \text{if } |\widetilde{Z}_t| = 0. \end{cases} \tag{7.2}$$

This turned out to be the case with the same initial condition under which we deduced the ODE for $(\widetilde{Z}_t)_{0\leqslant t\leqslant T}$ for a finite $0 < T < \infty$. Note that we need to deal with it carefully since the $|\widetilde{Z}_t^L|$ term is in the denominator and it might go to 0 as L goes to infinity.

Here is the main result of this chapter.

Theorem 7.1. *Given $0 < T < \infty$, the family of processes $(\widetilde{X}_t^L)_{0\leqslant t\leqslant T} \in D([0,T], \mathbb{R}^k)$ indexed by L and defined in (7.1), with initial conditions $\widetilde{Z}_0^L = L^{-1}\left([La_0^1], [La_0^2], ..., [La_0^k]\right)$ for a fixed $a_0 = (a_0^1, a_0^2, ..., a_0^k) \in \mathbb{R}_+^k \setminus \{0\}$, converges in probability as $L \to \infty$, uniformly in time to a process $(\widetilde{X}_t)_{0\leq t\leq T}$. More precisely*

$$\forall \epsilon > 0,\ \lim_{L\to\infty} P\left(\sup_{0\leqslant t\leqslant T} \left|\widetilde{X}_t^L - \widetilde{X}_t\right| > \epsilon\right) = 0,$$

where $(\widetilde{X}_t)_{0\leqslant t\leqslant T}$ is the deterministic process that solves the ODE initial value problem

$$\frac{d\widetilde{X}_t}{dt} = \sum_{w\in J} p_w(\widetilde{X}_t)\left(w - \widetilde{X}_t\alpha(w)\right),\ \widetilde{X}_0 = \frac{a_0}{|a_0|},$$

in the classical sense which is also the weak limit of $(\widetilde{X}_t^L)_{0\leqslant t\leqslant T}$.

Proof. This result follows from Proposition 7.5, Corollary 7.8 and Proposition 7.9, which are done in the rest of the chapter. □

First, we need the following proposition.

Proposition 7.2. *Given $a_0 = (a_0^1, a_0^2, ..., a_0^k) \in \mathbb{R}_+^k$, in a finite time $0 < T < \infty$, $\exists\, b_0 > 0$ such that*

$$\lim_{L\to\infty} P\left(\inf_{t\in[0,T]} |\widetilde{Z}_t^L| < b_0\right) = 0, \tag{7.3}$$

where $\widetilde{Z}_0^L = L^{-1}\left([La_0^1], [La_0^2], ..., [La_0^k]\right)$.

To show Proposition 7.2 , we need a lemma first.

Lemma 7.3. *Given* $a_0 = (a_0^1, a_0^2, ..., a_0^k) \in \mathbb{R}_+^k$, *in a finite time* $0 < T < \infty$, $\exists \epsilon > 0$ *such that the solution* $(\widetilde{Z}_t)_{t \geqslant 0}$ *to the initial value problem of (6.42) with* $\widetilde{Z}_0 = a_0$ *satisfies that*

$$\inf_{0 \leqslant t \leqslant T} |\widetilde{Z}_t| > \epsilon > 0. \tag{7.4}$$

Proof. Let g_i be as defined in (6.22), i.e.,

$$\begin{aligned} g_i(Z) &= g_i(Z^1, ..., Z^k) \\ &= \begin{cases} |Z| \sum_{w \in J} p_w \left(\frac{Z}{|Z|} \right) w_i, & \text{if } Z \neq 0 \\ 0, & \text{if } Z = 0 \end{cases} \end{aligned}$$

note that $\left| d|\widetilde{Z}_t^i| \right| \leqslant \left| d\widetilde{Z}_t^i \right|$, we have

$$\left| d|\widetilde{Z}_t| \right| = \left| d \sum_{i=1}^k |\widetilde{Z}_t^i| \right| \leqslant \sum_{i=1}^k \left| d\widetilde{Z}_t^i \right| = \sum_{i=1}^k \left| g_i(\widetilde{Z}_t) \right| |dt| = \left| g(\widetilde{Z}_t) \right| |dt|.$$

Since $\left| g_i(\widetilde{Z}_t) \right| \leqslant m|J||\widetilde{Z}_t|$, we have $\left| d|\widetilde{Z}_t| \right| \geqslant -mk|J||\widetilde{Z}_t||dt|$. With the initial condition $\widetilde{Z}_0 = a_0 \in \mathbb{R}_+^k$, we have $\left| \widetilde{Z}_0 \right| = |a_0| > 0$, and in a finite time $0 \leqslant t \leqslant T$ we get

$$\left| \widetilde{Z}_t \right| \geqslant \left| \widetilde{Z}_0 \right| \exp\{-mk|J|t\} \geqslant |a_0| \exp\{-mk|J|T\}. \tag{7.5}$$

Now pick $0 < \epsilon < |a_0| \exp\{-mk|J|T\}$, we can conclude that $\inf_{0 \leqslant t \leqslant T} |\widetilde{Z}_t| > \epsilon > 0$. □

Next we proceed to show Proposition 7.2 .

Proof. Construct a continuous function $l_\epsilon : \mathbb{R} \to \mathbb{R}$ such that for $\epsilon > 0$,

$$l_\epsilon(x) = \begin{cases} 1, & \text{if } x \leq \dfrac{\epsilon}{2} \\ 0, & \text{if } x \geq \epsilon \\ \textit{linear}, & \text{if } \dfrac{\epsilon}{2} < x < \epsilon. \end{cases} \tag{7.6}$$

Note that for our proof to work, $l_\epsilon(x)$ doesn't have to be defined linear for $\frac{\epsilon}{2} < x < \epsilon$, as long as it is continuous at end points and $0 \leqslant l_\epsilon(x) \leqslant 1$.

Recall we proved in (II) that $q : D([0,T], \mathbb{R}) \to \mathbb{R}$ where $q(\eta) = \sup_{0 \leqslant t \leqslant T} |\eta(t)|$is continuous. A similar argument applies and we know $\inf_{0 \leqslant t \leqslant T} |\eta(t)| : D([0,T], \mathbb{R}^k) \to \mathbb{R}$ is also continuous. Therefore, $l_\epsilon(\inf_{0 \leqslant t \leqslant T} |\eta(t)|) : D([0,T], \mathbb{R}^k) \to \mathbb{R}$ is a continuous and bounded function. Since $(\widetilde{Z}_t^L)_{0 \leqslant t \leqslant T} \xrightarrow{D} (\widetilde{Z}_t)_{0 \leqslant t \leqslant T}$, we have $E\left[l_\epsilon(\inf_{0 \leqslant t \leqslant T} |\widetilde{Z}_t^L|)\right] \to E\left[l_\epsilon(\inf_{0 \leqslant t \leqslant T} |\widetilde{Z}_t|)\right]$.
Note that for $0 \leqslant Z < \infty$, we have

$$\mathbb{1}_{[0,\frac{\epsilon}{2})}(Z) \leqslant l_\epsilon(Z) \leqslant \mathbb{1}_{[0,\epsilon)}(Z). \tag{7.7}$$

It follows that

$$P(\inf_{0 \leqslant t \leqslant T} |\widetilde{Z}_t^L| < \frac{\epsilon}{2}) = E\left[\mathbb{1}_{[0,\frac{\epsilon}{2})}(\inf_{0 \leqslant t \leqslant T} |\widetilde{Z}_t^L|)\right] \leqslant E\left[[l_\epsilon(\inf_{0 \leqslant t \leqslant T} |\widetilde{Z}_t^L|)\right], \tag{7.8}$$

$$E\left[[l_\epsilon(\inf_{0 \leqslant t \leqslant T} |\widetilde{Z}_t|)\right] \leqslant E\left[\mathbb{1}_{[0,\epsilon)}(\inf_{0 \leqslant t \leqslant T} |\widetilde{Z}_t|)\right] = P(\inf_{0 \leqslant t \leqslant T} |\widetilde{Z}_t| < \epsilon), \tag{7.9}$$

and as $L \to \infty$, we have

$$\begin{aligned}
&\limsup_{L\to\infty} P(\inf_{0\leqslant t\leqslant T} |\widetilde{Z}_t^L| < \frac{\epsilon}{2}) \\
&\leqslant \limsup_{L\to\infty} E\left[[l_\epsilon(\inf_{0\leqslant t\leqslant T} |\widetilde{Z}_t^L|)\right] \\
&= E\left[[l_\epsilon(\inf_{0\leqslant t\leqslant T} |\widetilde{Z}_t|)\right] \\
&\leqslant P(\inf_{0\leqslant t\leqslant T} |\widetilde{Z}_t| < \epsilon).
\end{aligned} \tag{7.10}$$

Now pick $0 < \epsilon < |a_0| \exp\{-mk|J|T\}$ and $b_0 = \dfrac{\epsilon}{2}$, by Lemma 7.3 we can conclude $\lim_{L\to\infty} P(\inf_{0\leqslant t\leqslant T} |\widetilde{Z}_t^L| < b_0) = 0$. □

Next we proceed to show the tightness of $(\widetilde{X}_t^L)_{0\leqslant t\leqslant T}$. As in Chapter 5 , we need to prove two conditions:

$$\text{there exists a M} > 0 \text{ such that} \limsup_{N\to\infty} P\left(|\widetilde{X}_0^L| > M\right) = 0, \tag{X1}$$

and

$$\forall \epsilon > 0,\ \lim_{\delta\to 0} \limsup_{L\to\infty} P\left(\sup_{\substack{s,t\in[0,T] \\ 0<|t-s|<\delta}} \left|\widetilde{X}_t^L - \widetilde{X}_s^L\right| > \epsilon\right) = 0. \tag{X2}$$

Note that by definition, we have $|\widetilde{X}_t^L| \leqslant 1$ and if we choose $M > 1$ condition (X1) is trivially satisfied. So the key part is to prove condition (X2). The following lemma about $(\widetilde{Z}_t^L)_{0\leqslant t\leqslant T}$ can be useful.

Lemma 7.4. $\forall \epsilon > 0,\ \lim_{L\to\infty} P\left(\sup_{0\leqslant t\leqslant T} \left|\widetilde{Z}_t^L - \widetilde{Z}_t\right| > \epsilon\right) = 0.$

Note that we assumed that we are given $a_0 = (a_0^1, a_0^2, ..., a_0^k) \in \mathbb{R}_+^k$ and initial condition $\widetilde{Z}_0^L = L^{-1}\left([La_0^1], [La_0^2], ..., [La_0^k]\right)$ without explicitly mentioning it every time

Proof. We will show that $\limsup_{L\to\infty} P\left(\sup_{0\leqslant t\leqslant T} \left|\widetilde{Z}_t^L - \widetilde{Z}_t\right| > \epsilon\right) = 0.$

Construct a continuous function $h_\epsilon : \mathbb{R} \to \mathbb{R}$ such that for $\epsilon > 0$,

$$h_\epsilon(x) = \begin{cases} 0, & \text{if } x \le \dfrac{\epsilon}{2} \\ 1, & \text{if } x \ge \epsilon \\ \mathit{linear}, & \text{if } \dfrac{\epsilon}{2} < x < \epsilon. \end{cases} \tag{7.11}$$

Note that $h_\epsilon(x)$ is continuous bounded and we have

$$\mathbb{1}_{(\epsilon,\infty)}(Z) \leqslant h_\epsilon(Z) \leqslant 1. \tag{7.12}$$

Let $\eta_t^L = \widetilde{Z}_t^L - \widetilde{Z}_t, 0 \leqslant t \leqslant T$, then $\eta_t^L \in \mathcal{D}([0,T],\mathbb{R}^k)$. By (7.12), we have that

$$P(\sup_{0\leqslant t\leqslant T} |\eta^L(t)| > \epsilon) = E[\mathbb{1}_{(\epsilon,\infty)}(\sup_{0\leqslant t\leqslant T} |\eta^L(t)|)] \leqslant E[h_\epsilon(\sup_{0\leqslant t\leqslant T} |\eta^L(t)|)]. \tag{7.13}$$

Recall that we proved in (II) $q : D([0,T],\mathbb{R}) \to \mathbb{R}$ where $q(\eta) = \sup_{0\leqslant t\leqslant T} |\eta(t)|$ is continuous. It is also the case when $\eta \in D([0,T],\mathbb{R}^k)$ by a similar argument. Consequently, $h_\epsilon(\sup_{0\leqslant t\leqslant T} |\eta(t)|) : D([0,T],\mathbb{R}^k) \to \mathbb{R}$ is a continuous and bounded function. Note that $(\eta_t^L)_{0\leqslant t\leqslant T} \xrightarrow{D} (0)_{0\leqslant t\leqslant T}$ since $(\widetilde{Z}_t^L)_{0\leqslant t\leqslant T} \xrightarrow{D} (\widetilde{Z}_t)_{0\leqslant t\leqslant T}$. Taking limsup over (7.13) we can conclude that

$$\begin{aligned} &\limsup_{L\to\infty} P(\sup_{0\leqslant t\leqslant T} |\widetilde{Z}_t^L - \widetilde{Z}_t| > \epsilon) \\ =& \limsup_{L\to\infty} P(\sup_{0\leqslant t\leqslant T} |\eta^L(t)| > \epsilon) \\ \leqslant& \limsup_{L\to\infty} E[h_\epsilon(\sup_{0\leqslant t\leqslant T} |\eta^L(t)|)] \\ =& E[h_\epsilon(\sup_{0\leqslant t\leqslant T} |0|)] \\ =& 0\,. \end{aligned} \tag{7.14}$$

□

Next we show the convergence in probability of $\widetilde{X}_t^L$, uniformly in time $t \in [0, T]$.

Proposition 7.5. $\forall \epsilon' > 0,\ \lim_{L\to\infty} P\left(\sup_{0\leqslant t\leqslant T}\left|\widetilde{X}_t^L - \widetilde{X}_t\right| > \epsilon'\right) = 0.$

Proof. It suffices to show that $\limsup_{L\to\infty} P\left(\sup_{0\leqslant t\leqslant T}\left|\widetilde{X}_t^L - \widetilde{X}_t\right| > \epsilon'\right) = 0.$
By Lemma 7.3, we know that for a given initial condition $\widetilde{Z}_0 = a_0 \in \mathbb{R}_+^k$, $\exists \epsilon > 0$ such that $\inf_{0\leqslant t\leqslant T} |\widetilde{Z}_t| > \epsilon > 0$. First suppose $|\widetilde{X}_t^L| \neq 0$ or simply $|\widetilde{Z}_t^L| \neq 0$ and we have

$$\begin{aligned}
&\left|\widetilde{X}_t^L - \widetilde{X}_t\right| \\
&= \left|\frac{\widetilde{Z}_t^L}{\left|\widetilde{Z}_t^L\right|} - \frac{\widetilde{Z}_t}{\left|\widetilde{Z}_t\right|}\right| \\
&= \left|\frac{\widetilde{Z}_t^L\left|\widetilde{Z}_t\right| - \widetilde{Z}_t\left|\widetilde{Z}_t^L\right|}{\left|\widetilde{Z}_t^L\right|\left|\widetilde{Z}_t\right|}\right| \\
&= \left|\frac{\widetilde{Z}_t^L\left|\widetilde{Z}_t\right| - \widetilde{Z}_t^L\left|\widetilde{Z}_t^L\right| + \widetilde{Z}_t^L\left|\widetilde{Z}_t^L\right| - \widetilde{Z}_t\left|\widetilde{Z}_t^L\right|}{\left|\widetilde{Z}_t^L\right|\left|\widetilde{Z}_t\right|}\right| \\
&\leqslant \left|\frac{\widetilde{Z}_t^L(|\widetilde{Z}_t| - |\widetilde{Z}_t^L|)}{|\widetilde{Z}_t^L||\widetilde{Z}_t|}\right| + \left|\frac{\widetilde{Z}_t^L - \widetilde{Z}_t}{|\widetilde{Z}_t|}\right| \\
&\leqslant 2\frac{|\widetilde{Z}_t^L - \widetilde{Z}_t|}{|\widetilde{Z}_t|} \\
&\leqslant \frac{2|\widetilde{Z}_t^L - \widetilde{Z}_t|}{\epsilon}.
\end{aligned}$$

By Proposition 7.2, we know $\exists\, b_0 > 0$ such that

$$\lim_{L\to\infty} P\left(\inf_{t\in[0,T]} |\widetilde{Z}_t^L| < b_0\right) = 0.$$

For this b_0, we further have

$$
\begin{aligned}
&\left|\widetilde{X}_t^L - \widetilde{X}_t\right| \\
&= \left|\widetilde{X}_t^L - \widetilde{X}_t\right| \mathbb{1}_{\{|\widetilde{Z}_t^L|<b_0\}} + \left|\widetilde{X}_t^L - \widetilde{X}_t\right| \mathbb{1}_{\{|\widetilde{Z}_t^L|\geqslant b_0\}} \\
&\leqslant 2\mathbb{1}_{\{|\widetilde{Z}_t^L|<b_0\}} + \frac{2|\widetilde{Z}_t^L - \widetilde{Z}_t|}{\epsilon}\mathbb{1}_{\{|\widetilde{Z}_t^L|\geqslant b_0\}},
\end{aligned}
\tag{7.15}
$$

and it follows that

$$
\begin{aligned}
&P\left(\sup_{0\leqslant t\leqslant T}\left|\widetilde{X}_t^L - \widetilde{X}_t\right| > \epsilon'\right) \\
&\leqslant P(\sup_{0\leqslant t\leqslant T}\left|2\mathbb{1}_{\{|\widetilde{Z}_t^L|<b_0\}}\right| > \frac{\epsilon'}{2}) + P(\sup_{0\leqslant t\leqslant T}\left|\frac{2|\widetilde{Z}_t^L - \widetilde{Z}_t|}{\epsilon}\mathbb{1}_{\{|\widetilde{Z}_t^L|\geqslant b_0\}}\right| > \frac{\epsilon'}{2}).
\end{aligned}
\tag{7.16}
$$

Taking limsup, by Proposition 7.2 and Lemma 7.4 respectively, we have that

$$
\limsup_{L\to\infty} P\left(\sup_{0\leqslant t\leqslant T}\left|2\mathbb{1}_{\{|\widetilde{Z}_t^L|<b_0\}}\right| > \frac{\epsilon'}{2}\right) \leqslant \limsup_{L\to\infty} P\left(\inf_{0\leqslant t\leqslant T}|\widetilde{Z}_t^L| < b_0\right) = 0,
$$

and

$$
\limsup_{L\to\infty} P\left(\sup_{0\leqslant t\leqslant T}\left|\frac{2|\widetilde{Z}_t^L - \widetilde{Z}_t|}{\epsilon}\mathbb{1}_{\{|\widetilde{Z}_t^L|\geqslant b_0\}}\right| > \frac{\epsilon'}{2}\right) \leqslant \limsup_{L\to\infty} P\left(\sup_{0\leqslant t\leqslant T}\frac{2|\widetilde{Z}_t^L - \widetilde{Z}_t|}{\epsilon} > \frac{\epsilon'}{2}\right) = 0.
$$

Therefore,

$$
\limsup_{L\to\infty} P\left(\sup_{0\leqslant t\leqslant T}\left|\widetilde{X}_t^L - \widetilde{X}_t\right| > \epsilon'\right) \leqslant 0,
$$

and it must be

$$
\lim_{L\to\infty} P\left(\sup_{0\leqslant t\leqslant T}\left|\widetilde{X}_t^L - \widetilde{X}_t\right| > \epsilon'\right) = 0.
$$

□

We comment that once we have Proposition 7.5 , condition (X2) holds.

Corollary 7.6.

$$\forall \epsilon > 0,\ \lim_{\delta \to 0} \limsup_{L \to \infty} P\left(\sup_{\substack{s,t \in [0,T] \\ 0 < |t-s| < \delta}} \left| \widetilde{X}_t^L - \widetilde{X}_s^L \right| > \epsilon \right) = 0.$$

Proof. Note that

$$|\widetilde{X}_t^L - \widetilde{X}_s^L| \leqslant \left| \widetilde{X}_t^L - \widetilde{X}_t \right| + \left| \widetilde{X}_s^L - \widetilde{X}_s \right| + \left| \widetilde{X}_s - \widetilde{X}_t \right| .$$

we have

$$\begin{aligned} & P\left(\sup_{\substack{s,t \in [0,T] \\ 0 < |t-s| < \delta}} \left| \widetilde{X}_t^L - \widetilde{X}_s^L \right| > \epsilon \right) \\ \leqslant & P\left(\sup_{\substack{s,t \in [0,T] \\ 0 < |t-s| < \delta}} \left| \widetilde{X}_t^L - \widetilde{X}_t \right| > \frac{\epsilon}{3} \right) + P\left(\sup_{\substack{s,t \in [0,T] \\ 0 < |t-s| < \delta}} \left| \widetilde{X}_s^L - \widetilde{X}_s \right| > \frac{\epsilon}{3} \right) + P\left(\sup_{\substack{s,t \in [0,T] \\ 0 < |t-s| < \delta}} \left| \widetilde{X}_s - \widetilde{X}_t \right| > \frac{\epsilon}{3} \right) \\ = & P\left(\sup_{t \in [0,T]} \left| \widetilde{X}_t^L - \widetilde{X}_t \right| > \frac{\epsilon}{3} \right) + P\left(\sup_{s \in [0,T]} \left| \widetilde{X}_s^L - \widetilde{X}_s \right| > \frac{\epsilon}{3} \right) + P\left(\sup_{\substack{s,t \in [0,T] \\ 0 < |t-s| < \delta}} \left| \widetilde{X}_s - \widetilde{X}_t \right| > \frac{\epsilon}{3} \right) \end{aligned}$$

Taking $\limsup_{L \to \infty}$ and by Proposition 7.5 , we obtain that

$$\limsup_{L \to \infty} P\left(\sup_{\substack{s,t \in [0,T] \\ 0 < |t-s| < \delta}} \left| \widetilde{X}_t^L - \widetilde{X}_s^L \right| > \epsilon \right) \leqslant P\left(\sup_{\substack{s,t \in [0,T] \\ 0 < |t-s| < \delta}} \left| \widetilde{X}_s - \widetilde{X}_t \right| > \frac{\epsilon}{3} \right). \tag{7.17}$$

Since $(\widetilde{X}_t)$ is uniformly continuous for $0 \leqslant t \leqslant T$, we have

$$\lim_{\delta \to 0_+} P\left(\sup_{\substack{s,t \in [0,T] \\ 0 < |t-s| < \delta}} \left| \widetilde{X}_s - \widetilde{X}_t \right| > \frac{\epsilon}{3} \right) = 0,$$

and it follows that

$$\lim_{\delta \to 0} \limsup_{L \to \infty} P\left(\sup_{\substack{s,t \in [0,T] \\ 0 < |t-s| < \delta}} \left| \widetilde{X}_t^L - \widetilde{X}_s^L \right| > \epsilon \right) = 0.$$

where $\epsilon > 0$ is arbitrary. □

Now we have tightness for $(\widetilde{X}_t^L)_{0\le t\le T}$.

Corollary 7.7. *The family of processes* $(\widetilde{X}_t^L)_{0\le t\le T}$ *as defined in* (7.1) *is tight.*

Proof. We have tightness of $(\widetilde{X}_t^L)_{0\leqslant t\leqslant T}$ if both conditions (X1) and (X2) are satisfied. Condition (X1) is satisfied due to the fact that $(\widetilde{X}_t^L)_{0\leqslant t\leqslant T}$ is bounded and Corollary 7.6 validates Condition (X2). □

Also, convergence in distribution of the process $(\widetilde{X}_t^L)_{0\leqslant t\leqslant T}$ follows from Proposition 7.5 .

Corollary 7.8. *The family of processes* $(\widetilde{X}_t^L)_{0\leqslant t\leqslant T}$ *indexed by L converges in distribution to the process* $(\widetilde{X}_t)_{0\leqslant t\leqslant T}$*, uniformly in time.*

Proof. By Proposition 7.5, we know that

$$\forall \epsilon' > 0,\ \lim_{L\to\infty} P\left(\sup_{0\leqslant t\leqslant T}\left|\widetilde{X}_t^L - \widetilde{X}_t\right| > \epsilon'\right) = 0.$$

Then for any fixed $0 \leqslant t \leqslant T$, we can conclude that $X_t^L \xrightarrow{P} X_t$ and further $X_t^L \xrightarrow{D} X_t$. $\forall\, 0 \leqslant t_1 < t_2 < \cdots < t_m \leqslant T$ and test functions $f_1, f_2, \cdots, f_m \in C_c^\infty$, note that those f_i are uniformly bounded once given, it is easily seen that

$$\lim_{L\mapsto\infty} E\left[f_1(X_{t_1}^L)\cdots f_m(X_{t_m}^L)\right] = E\left[f_1(X_{t_1})\cdots f_m(X_{t_m})\right]. \tag{7.18}$$

This implies $(\widetilde{X}_t^L)_{0\leqslant t\leqslant T} \xrightarrow{D} (\widetilde{X}_t)_{0\leqslant t\leqslant T}$. □

Next we show that the behavior of $(\widetilde{X}_t)_{0\leqslant t\leqslant T}$ is characterized by an ODE which can be obtained from that of $(\widetilde{Z}_t)_{0\leqslant t\leqslant T}$.

Proposition 7.9. *Given $\widetilde{Z}_0 = a_0 = (a_0^1, a_0^2, ..., a_0^k) \in \mathbb{R}_+^k$, in a finite time $0 < T < \infty$, we have*

$$\frac{d\widetilde{X}_t}{dt} = \sum_{w \in J} p_w(\widetilde{X}_t)\left(w - \widetilde{X}_t \alpha(w)\right).$$

Proof. Note that By Proposition 6.11 , we know in a finite time $0 < T < \infty$,

$$\frac{d\widetilde{Z}_t}{dt} = |\widetilde{Z}_t| \sum_{w \in J} p_w \left(\frac{\widetilde{Z}_t}{|\widetilde{Z}_t|}\right) w,$$

or component-wise we have

$$\frac{d\widetilde{Z}_t^i}{dt} = |\widetilde{Z}_t| \sum_{w \in J} p_w \left(\frac{\widetilde{Z}_t}{|\widetilde{Z}_t|}\right) w_i.$$

By Lemma 7.3 , we know $\exists\, \epsilon$ such that $\inf\limits_{0 \leqslant t \leqslant T} |\widetilde{Z}_t| > \epsilon > 0$ for a given initial condition $\widetilde{Z}_0 = a_0 \in \mathbb{R}_+^k$. Applying a similar argument as used in the proof of Lemma 7.3 , we also have that if $\widetilde{Z}_t^i = a_i \geq 0$, then $\inf\limits_{0 \leqslant t \leqslant T} \widetilde{Z}_t^i \geq 0$. This allows us to add up those equations for components and obtain

$$\frac{d|\widetilde{Z}_t|}{dt} = |\widetilde{Z}_t| \sum_{w \in J} p_w \left(\frac{\widetilde{Z}_t}{|\widetilde{Z}_t|}\right) \alpha(w),$$

as appeared in (3.27). Again under the condition of Lemma 7.3 , we can deduce

$$\begin{aligned} \frac{d\widetilde{X}_t^i}{dt} &= \frac{d}{dt}\left(\frac{\widetilde{Z}_t^i}{|\widetilde{Z}|_t}\right) \\ &= \frac{1}{|\widetilde{Z}_t|} \cdot \frac{d\widetilde{Z}_t^i}{dt} - \frac{\widetilde{Z}_t^i}{|\widetilde{Z}_t|^2} \cdot \frac{d|\widetilde{Z}_t^i|}{dt} \\ &= \sum_{w \in J} p_w(\widetilde{X}_t) w_i - \widetilde{X}_t^i \sum_{w \in J} p_w(\widetilde{X}_t)\alpha(w) \\ &= \sum_{w \in J} p_w(\widetilde{X}_t)\left(w_i - \widetilde{X}_t^i \alpha(w)\right), \end{aligned} \tag{7.19}$$

for any given finite time $0 < t < T$ since $\inf_{0 \leqslant t \leqslant T} |\widetilde{Z}_t| > \epsilon > 0$ guarantees that each term makes sense and the calculations can be carried out. Writing in a vector form, we have

$$\frac{d\widetilde{X}_t}{dt} = \sum_{w \in J} p_w(\widetilde{X}_t) \left(w - \widetilde{X}_t \alpha(w) \right).$$

$\square$

With all the relevant results are proven, we proved Theorem 7.1.

Chapter 8

Bound Estimates for Extinction

First, we know from Lemma 5.4 that

$$|\widetilde{Z}_t^L| - |\widetilde{Z}_0^L| - \int_0^t L|\widetilde{Z}_{s-}^L| \sum_{w \in J} \Pi(L\widetilde{Z}_{s-}^L, L\widetilde{Z}_{s-}^L + w) \left(|\widetilde{Z}_{s-}^L + \frac{w}{L}| - |\widetilde{Z}_{s-}^L| \right) ds \triangleq \mathcal{N}_t^L \tag{8.1}$$

is a square integrable martingale with respect to the filtration $(\mathcal{F}_t)_{t \geq 0}$. We argue that we can replace each $\widetilde{Z}_{s-}^L$ with $\widetilde{Z}_s^L$ since a RCLL jump process can have at most countably many jumps implying the set of points at which $\widetilde{Z}_s^L$ differs from $\widetilde{Z}_{s-}^L$ is of measure zero. Therefore, we have

$$|\widetilde{Z}_t^L| - |\widetilde{Z}_0^L| - \int_0^t |\widetilde{Z}_s^L| \sum_{w \in J} \Pi(L\widetilde{Z}_s^L, L\widetilde{Z}_s^L + w) \left(|L\widetilde{Z}_s^L + w| - |L\widetilde{Z}_s^L| \right) ds = \mathcal{N}_t^L. \tag{8.2}$$

Define $Y_s^L = |\widetilde{Z}_s^L|$ and $P_s^L = \sum_{w \in J} \Pi(L\widetilde{Z}_s^L, L\widetilde{Z}_s^L + w) \left(|L\widetilde{Z}_s^L + w| - |L\widetilde{Z}_s^L| \right)$, rewriting (8.2), we have

$$Y_t^L - Y_0^L - \int_0^t Y_s^L P_s^L ds = \mathcal{N}_t^L, \tag{8.3}$$

or in a differential form

$$dY_t^L = Y_t^L P_t^L dt + d\mathcal{N}_t^L \,. \tag{8.4}$$

Let $g(s,x) = xe^{2s}$, applying stochastic differentiation in the sense of (3.8) to $g(s, Y_s^L)$, we get that R_t,

$$R_t = e^{2t}Y_t^L - Y_0^L - U_t \,, \qquad \text{where} \qquad U_t = \int_0^t e^{2s}(2 + P_s^L)Y_s^L ds \tag{8.5}$$

is a martingale. Note that we have $Y_t^L \geq 0$, also $P_t^L \geq -2$ since $\sum_{w \in J} \Pi(L\widetilde{Z}_s^L, L\widetilde{Z}_s^L + w) = 1$, each component w_i of w satisfies $-1 \leq w_i \leq m$ and by construction we have at most two components change at the same time. Then the integration part with respect to s in equation (8.5) is nonnegative and we obtain that

$$R_t + U_t = e^{2t}Y_t^L - Y_0^L \tag{8.6}$$

is a sub-martingale. We know that Y_t^L is nonnegative, but more than that, we are interested in how fast does it approach 0 starting at a positive position. To be more specific, we want to estimate $P(\{\inf_{0 \leq t \leq T} Y_t^L \leq \epsilon\})$ for a small $\epsilon > 0$ in a finite time interval $0 < T < \infty$. Some calculations give

$$
\begin{aligned}
P(\{\inf_{0\le t\le T} Y_t^L \le \epsilon\}) &\le P(\{\inf_{0\le t\le T} Y_t^L e^{2t} \le \epsilon e^{2T}\}) && (8.7)\\
&= P(\{\inf_{0\le t\le T}(Y_t^L e^{2t} - Y_0^L) \le \epsilon e^{2T} - Y_0^L\}) && (8.8)\\
&= P(\{\inf_{0\le t\le T}(R_t + U_t) \le \epsilon e^{2T} - Y_0^L\}) && (8.9)\\
&\le P(\{\inf_{0\le t\le T} R_t \le \epsilon e^{2T} - Y_0^L\}), \qquad \text{since} \quad U_t \ge 0 && (8.10)\\
&= P\left(\{-\sup_{0\le t\le T}(-R_t) \le \epsilon e^{2T} - Y_0^L\}\right) && (8.11)\\
&= P\left(\{\sup_{0\le t\le T}(-R_t) \ge Y_0^L - \epsilon e^{2T}\}\right). && (8.12)
\end{aligned}
$$

Note that $(-R_t)$ is a also a martingale with respect to the filtration $(\mathcal{F}_t)_{t\ge0}$. Assuming $Y_0^L - \epsilon e^{2T} > 0$, which can always be achieved by choosing a small enough ϵ given T and the initial $Y_0^L > 0$, by Doob's sub-martingale L^2 inequality we have that

$$
\begin{aligned}
&P(\{\sup_{0\le t\le T}(-R_t) \ge Y_0^L - \epsilon e^{2T}\})\\
&\le E\left[(R_t)^2\right]\left(Y_0^L - \epsilon e^{2T}\right)^{-2}.
\end{aligned} \tag{8.13}
$$

Let $\langle \mathcal{N}\rangle_s^L$ be the quadratic variation process corresponding to the martingale $\mathcal{N}_s{}^L$, we must have

$$
E\left[(R_t)^2\right] = E\left[\int_0^T (e^{2s})^2 d\langle \mathcal{N}\rangle_s^L\right]. \tag{8.14}
$$

Also note that a quadratic variation process is non-negative and non-decreasing, we have

$$
E\left[(\int_0^T (e^{2s})^2 d\langle \mathcal{N}\rangle_s^L)^2\right] \le e^{4T} E\left[\langle \mathcal{N}\rangle_T^L\right]. \tag{8.15}
$$

Combining equation (8.13), (8.14) and (8.15), we have

$$P\left(\{\sup_{0\le t\le T}(-R_t)\ge Y_0^L-\epsilon e^{2T}\}\right)\le e^{4T}E\left[\langle\mathcal{N}\rangle_T^L\right]\left(Y_0^L-\epsilon e^{2T}\right)^{-2}. \tag{8.16}$$

Clearly we need to estimate $E\left[\langle\mathcal{N}\rangle_T^L\right]$. Note that it is the quadratic variation process corresponding to the martingale $\widetilde{\mathcal{N}_t}^L$, and the latter which satisfies equation (8.2) is a martingale corresponding to the pure jump process Y_s^L, we have

$$E\left[\langle\mathcal{N}\rangle_T^L\right]=E\left[\int_0^T L|\widetilde{Z}_s^L|\sum_{w\in J}\Pi(L\widetilde{Z}_s^L,L\widetilde{Z}_s^L+w)\left(\frac{|L\widetilde{Z}_s^L+w|-|L\widetilde{Z}_s^L|}{L}\right)^2 ds\right]. \tag{8.17}$$

Note that $\sum_{w\in J}\Pi(L\widetilde{Z}_s^L,L\widetilde{Z}_s^L+w)=1$ and $-2\le|L\widetilde{Z}_s^L+w|-|L\widetilde{Z}_s^L|\le 2m$, we obtain from equation (8.17) that

$$E\left[\langle\mathcal{N}\rangle_T^L\right]\le\frac{4m^2}{L}E\left[\int_0^T|\widetilde{Z}_s^L|ds\right]\le\frac{4Tm^2}{L}E\left[\sup_{0\le t\le T}|\widetilde{Z}_s^L|\right] \tag{8.18}$$

Recall in Proposition 5.2 we obtained
$\forall L>1, E\left[\sup_{0\leqslant t\leqslant T}|\widetilde{Z}_{t^-}^L|\right]\leqslant C_1(k,m,T)$, where $C_1(k,m,T)$ is a constant depending on k,m,T but not on L. Note that $\widetilde{Z}_{t^-}^L$ differs from $\widetilde{Z}_t^L$ in t only on a set of at most countably many elements, we know

$$E\left[\sup_{0\leqslant t\leqslant T}|\widetilde{Z}_t^L|\right]=E\left[\sup_{0\leqslant t\leqslant T}|\widetilde{Z}_{t^-}^L|\right]\leqslant C_1(k,m,T). \tag{8.19}$$

Combining (8.7), (8.16) (8.18)and (8.19), we can now conclude that when $Y_0^L-\epsilon e^{2T}>0$

$$P(\{\inf_{0\le t\le T}Y_t^L\le\epsilon\})\le 4Tm^2L^{-1}e^{4T}C_1(k,m,T)\left(Y_0^L-\epsilon e^{2T}\right)^{-2}. \tag{8.20}$$

We now have the main result of this chapter.

Theorem 8.1. *Let* $\epsilon_1 = \dfrac{|a_0|}{2e^{2T}}$. *Then, for any* $0 \le \epsilon \le \epsilon_1$, *the total number of individuals* $|\widetilde{Z}_t^L|$ *on any bounded interval* $[0, T]$ *satisfies the large deviations bound*

$$\limsup_{L\to\infty} L\, P(\{\inf_{0\le t\le T} |\widetilde{Z}_t^L| \le \epsilon\}) \le C_Y(k, m, T, a_0)\,, \tag{8.21}$$

where $C_Y(k, m, T, a_0) = 16Tm^2e^{4T}C_1(k, m, T)(|a_0|)^{-2}$ *is a positive constant given* k, a_0, m, T.

It clear that for arbitrary but fixed $T > 0$, we must have $\lim_{L\to+\infty} P(\{\inf_{0\le t\le T} |\widetilde{Z}_t^L| \le \epsilon\}) = 0$ for small enough ϵ. Note that $\widetilde{Z}_t^L$ converges to $\widetilde{Z}_t$ in probability uniformly on $[0, T]$ by lemma 7.4, we have $P(\{\inf_{0\le t\le T} |\widetilde{Z}_t| \le \epsilon\}) = 0$.

Corollary 8.2. *As a consequence, the limit* $(\widetilde{Z}_t)_{t\ge0}$ *never vanishes, almost surely.*

Remark 8.3. *We claim* $(\widetilde{Z}_t)_{t\ge0}$ *doesn't vanish in finite time. Of course, it is possible that* $z = 0$ *be a stable point of the system, and the solution may go to zero as* $t \to +\infty$.

Chapter 9

Future Research

There could be a further generalization of the approximation bound (3.14). Let

$$A(Z) = \sup_{w} \left| p_w(\frac{Z}{|Z|}) - \Pi(Z, Z + w) \right| \tag{9.1}$$

be the error in the asymptotic formula for the transition probabilities.

Until now we used the bound $A(w, Z) \leq a|Z|^{-1}$, where $a > 0$ does not depend on the Z.

Conjecture 9.1. *The hydrodynamic limit holds under the weaker assumption*

$$\lim_{|Z| \to \infty} A(Z) = 0\,. \tag{9.2}$$

Conjecture 9.2. *Under (9.2), we can prove a fluctuation limit as $L \to \infty$. More precisely, let*

$$\xi_t^L = \sqrt{L}\Big(\frac{Z_t}{L} - z_t\Big) \in \mathbb{R}^k \tag{9.3}$$

be the k-dimensional fluctuation random field for all $t \geq 0$. Assume the initial values scale to a normal random variable as in the Central Limit Theorem. Further assume that A, the *differential operator of the limiting dynamical system, is symmetric (technically we want a complete space of eigenfunctions). Then $(\xi_t^L)_{t\geq 0}$ is a tight k-dimensional Gaussian random field and its limit ξ_t satisfies, for a smooth test function ϕ*

$$d\langle \xi_t, \phi \rangle = \langle \xi_t, A\phi \rangle dt + dU_t \tag{9.4}$$

where

$$U_t^2 - \int_0^t |z_s| \sum_w p_w(\frac{z_s}{|z_s|})|\nabla \phi(z_s) \cdot w|^2 \, ds \tag{9.5}$$

is a martingale.

Formula (9.5) says that if we define the time dependent (in $s \geq 0$) bilinear form

$$\phi \longrightarrow \langle C_s\phi, \phi \rangle = |z_s| \sum_w p_w(\frac{z_s}{|z_s|})|\nabla \phi(z_s) \cdot w|^2 \,, \qquad \phi \in C_c^\infty(\mathbb{R}^k, \mathbb{R})\,, \tag{9.6}$$

then

$$d\xi_t = A^*\xi_t dt + \sqrt{C_t}\, dW_t \tag{9.7}$$

where W_t is a k - dimensional Brownian motion.

Assume that the increments w are no longer $w \sim O(1)$ but $w \sim O(|Z|)$. This is the case when the birth rate is not a microscopic function of type, but of individual. In that case, heuristically, each individual may produce an offspring at certain rate, generating a population change macroscopically proportional to the total population Z. This is, by analogy, similar to how branching occurs in a Galton-Watson process.

Conjecture 9.3. *Determine the exact condition, incorporating $w \sim O(|Z|)$, such that the hydrodynamic limit holds. Research if the extension allows more realistic applications.*

Bibliography

[1] David Aldous. Stopping times and tightness. *Ann. Probability*, 6(2):335–340, 1978.

[2] Krishna B. Athreya. On a characteristic property of Polya's urn. *Studia Sci. Math. Hungar.*, 4:31–35, 1969.

[3] Michel Benaïm, Sebastian J. Schreiber, and Pierre Tarrès. Generalized urn models of evolutionary processes. *Ann. Appl. Probab.*, 14(3):1455–1478, 2004.

[4] Patrick Billingsley. *Convergence of probability measures*. John Wiley & Sons, Inc., New York-London-Sydney, 1968.

[5] Patrick Billingsley. *Probability and measure*. Third edition. John Wiley & Sons, Inc., New York, 1995.

[6] David Blackwell and James B. MacQueen. Ferguson distributions via Pólya urn schemes. *Ann. Statist.*, 1:353–355, 1973.

[7] Leo Breiman. *Probability*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.

[8] Donald A. Dawson. *Introductory Lectures on Stochastic Population Systems*. Technical Report Series No. 451. Laboratory for Research in Statistics and Probability, Carleton University, 2010.

[9] Rick Durrett. *Probability: Theory and examples*. Fourth edition. Cambridge University Press, Cambridge, 2010.

[10] Florian Eggenberger and George Pólya. Über die statistik verketteter vorgänge. *ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik*, 3(4):279–289, 1923.

[11] Alison M. Etheridge. *An introduction to superprocesses*. American Mathematical Society, Providence, RI, 2000.

[12] William Feller. Diffusion processes in genetics. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pages 227–246. University of California Press, Berkeley and Los Angeles, 1951.

[13] William Feller. *An introduction to probability theory and its applications. Vol. I.* Third edition. John Wiley & Sons, Inc., New York-London-Sydney, 1968.

[14] Bernard Friedman. A simple urn model. *Comm. Pure Appl. Math.*, 2:59–70, 1949.

[15] Francis Galton. Problem 4001. *Educational Times*, 1:17, 1873.

[16] Josef Hofbauer and Karl Sigmund. *Evolutionary games and population dynamics*. Cambridge University Press, Cambridge, 1998.

[17] Christiaan Huygens. *De ratiociniis in ludo aleae*. Ex officinia J. Elsevirii, 1980.

[18] Norman L. Johnson and Samuel Kotz. *Urn models and their application*. John Wiley & Sons, New York-London-Sydney, 1977.

[19] Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*. Second edition. Springer-Verlag, New York, 1991.

[20] David G. Kendall. Branching processes since 1873. *J. London Math. Soc.*, 41:385–406, 1966.

[21] David G. Kendall. The genealogy of genealogy: Branching processes before (and after) 1873. *Bull. London Math. Soc.*, 7(3):225–253, 1975.

[22] Claude Kipnis and Claudio Landim. *Scaling limits of interacting particle systems*. Springer-Verlag, Berlin, 1999.

[23] Samuel Kotz, Hosam Mahmoud, and Philippe Robert. On generalized Pólya urn models. *Statist. Probab. Lett.*, 49(2):163–173, 2000.

[24] Thomas M. Liggett. *Interacting particle systems*. Springer-Verlag, New York, 1985.

[25] Thomas M. Liggett. *Continuous time Markov processes: An introduction*. American Mathematical Society, Providence, RI, 2010.

[26] Hosam Mahmoud. Pólya urn models and connections to random trees: A review. *Journal of the Iranian Statistical Society*, 2(1):53–114, 2003.

[27] Hosam M. Mahmoud. *Pólya urn models*. CRC Press, 2008.

[28] Andrey A. Markov. Generalization of a problem on a sequential exchange of balls (in Russian). In *Collected Works, Meeting of Physico-Mathematical Society of the Academy of Sciences*, 1917.

[29] Andrey A. Markov. Extension of the law of large numbers to quantities, depending on each other (1906). Reprint. *Journal Électronique d'Histoire des Probabilités et de la Statistique [electronic only]*, 2(1b):Article–10, 2006.

[30] J. R. Norris. *Markov chains*. Cambridge University Press, Cambridge, 1998.

[31] David Pollard. *Convergence of stochastic processes*. Springer, 1984.

[32] Philip Protter. *Stochastic integration and differential equations*. Springer, 1990.

[33] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*. Third edition. Springer-Verlag, Berlin, 1999.

[34] L. C. G. Rogers and David Williams. *Diffusions, Markov processes, and martingales. Vol. 1*. Second edition. John Wiley & Sons, Ltd., Chichester, 1994.

[35] Sebastian J. Schreiber. Urn models, replicator processes, and random genetic drift. *SIAM Journal on Applied Mathematics*, 61(6):2148–2167 (electronic), 2001.

[36] O. B. Sheynin. D. Bernoulli's work on probability. *RETE*, 1(3/4):273–300, 1971/72.

[37] Henry William Watson and Francis Galton. On the probability of the extinction of families. *The Journal of the Anthropological Institute of Great Britain and Ireland*, 4:138–144, 1875.