UNIVERSITÀ DEGLI STUDI DI PADOVA

Scuola Galileiana di Studi Superiori

Classe di Scienze Naturali

Tesi di Diploma Galileiano

FEYNMAN-KAC MODELS FOR LARGE DEVIATIONS CONDITIONING PROBLEM

Relatore:
Prof. STEFAN GROSSKINSKY

Correlatore:
Prof. ADAM JOHANSEN

Diplomando:
LETIZIA ANGELI

ANNO ACCADEMICO 2016-2017

Contents

Introduction

The study of atypical, rare trajectories of dynamical systems arises in many physical applications such as planetary systems, molecular dynamics, energy transport and chemical reactions. Unfortunately, when a process is complex enough it becomes no longer feasible to simulate repeatedly the true dynamics to observe a large deviation event.

A widely used class of numerical procedures for generating rare events efficiently are the importance sampling methods based on *cloning algorithms*. Such algorithms are based on the evolution of a population of copies of the system which are evolved in parallel and are replicated or killed in such a way as to favour the realisation of the atypical trajectories. One of these algorithms proposed by Giardinà et al. [10, 11] is used to evaluate numerically the *scaled cumulant generating function* (SCGF) of additive observables in Markov processes. The SCGF plays indeed an essential role in the investigation of non-equilibrium systems - a role akin to the free energy in equilibrium ones [19].

This cloning method for estimating the SCGF has been used widely to many physical systems, including chaotic systems, glassy dynamics and non-equilibrium lattice gas models. However, there have been fewer studies on the analytical justification of the algorithm. In particular, even though it is heuristically believed that the SCGF estimator converges to the correct result as the size of the population N increases, there is no proof of this convergence and of how fast the estimator converges. In the last year, Hidalgo et al. [13, 14] proposed a slightly different version of the cloning algorithm, for which they studied the speed of convergence to the SCGF in the case of discrete-time models.

In this work, we propose a variant of the cloning algorithm based on Feynman-Kac models [2, 3, 4, 5, 6, 7, 8]. This novel approach enables us to study analytically the convergence of the algorithm to the scaled cumulant generating function in the case of continuous-time models, adapting already established convergence results for Feynman-Kac models. To our knowledge, the model described in this dissertation has never been covered in the literature on the subject, even though similar results were presented by Del Moral and Miclo [6, 8] in the context of Lyapunov exponents connected to Schrödinger operators.

Feynman-Kac models were originally introduced in the 1940s [15] to express the semigroup of a quantum particle evolving in a potential in terms of a functional-path integral formula. The key idea behind these models is to enter the effects of a potential in the distribution of the paths of a stochastic process. The main advantage of this interpretation is that it is possible to construct explicitly N-particle systems which converge to the associated Feynman-Kac model as the size of the system N tends to infinity.

These considerations have inspired us to find the Feynman-Kac interpretation of the SCGF and apply the theory behind Feynman-Kac models for constructing a suitable cloning algorithm and proving its convergence to the desired quantity.

The dissertation is structured as follows. In Chapter 1, we introduce various mathematical concepts used throughout this work, such as Markov semigroups, Markov generators and elements of large deviation theory needed to define and study our class of rare event conditioning. Part of the Chapter is devoted to defining the class of Pure Jump processes that we consider for conditioning.

In Chapter 2, we consider the problem of studying the atypical trajectories in the long-time limit of a continuous-time pure jump Markov process and show how this is connected to the Feynman-Kac theory. In particular, we rewrite the SCGF in terms of Feynman-Kac measures and provide the time evolution of these measures.

Finally, in Chapter 3, we provide an N-particle interpretation of the Feynman-Kac measures introduced in Chapter 2, so that we are able to construct a numerical procedure to estimate the SCGF. We also provide general conditions, below which we have that the N-particle system converges to the SCGF, as the size of the system N and the time tend to infinity. Under these conditions, we study the L^2-error of the N-particle approximation. A particularly interesting consequence of these results is the convergence of the algorithm to the *SCGF* for any pure jump process defined on a finite state space and with a recurrent transition matrix.

Chapter 1

Preliminaries

1.1 Feller Processes

The purpose of this section is to provide an introduction to Feller processes and to motivate their infinitesimal description. In particular, we define Feller processes, probability semigroups and generators and show that there are one-to-one correspondences among these three probability structures. The definitions and results presented in this section are based on [12] and Chapter 3 of [17].

The *state space* E is a Polish space, i.e. metrizable, separable and complete topological space. The measurable structure on E is given by the Borel σ-algebra $\mathcal{B}(E)$. Moreover, we assume E to be locally compact and we denote $\mathcal{C}(E)$ the space of continuous real-valued functions on E vanishing at infinity, with the uniform norm

$$||f|| := \sup_{x \in E} |f(x)|,$$

which makes $\mathcal{C}(E)$ a Banach space. We denote by $\mathcal{P}(E)$ the set of probability measures on E.

The *path space* of the processes considered is

$$\Omega = D([0,\infty),\, E) := \{\omega : [0,\infty) \to E \text{ càdlàg}\},$$

where *càdlàg* stands for right-continuous functions with left limits. The right-continuous filtration $(\mathcal{F}_t)_{t\geq 0}$ on Ω is such that the σ-algebra $\mathcal{F}_t$ is the smallest such that the mapping $\omega \mapsto \omega(t)$, with $\omega \in \Omega$, is $\mathcal{F}_t$-measurable for each $t \geq 0$.

A continuous-time stochastic process on E is denoted by $(X_t)_{t\geq 0}$, i.e. a family of random variables X_t. The explicit construction of the random variables is $X_t(\omega) = \omega(t)$. In general, the process $(X_t)_{t\geq 0}$ is characterised by a probability measure $\mathbb{P}$ on $(\Omega,\, \mathcal{F})$ with associated expectation denoted by $\mathbb{E}$. With ω drawn from this measure, the function $t \mapsto X_t(\omega)$ is called sample path. In the following we restrict ourselves to Feller processes, which form a particular class of Markov processes. To be able to consider different initial conditions for a given Feller process, we use the following definition.

Definition 1.1.1. A *Feller process* on E consists of a collection of probability measures $(\mathbb{P}^x)_{x\in E}$ on $(\Omega, \mathcal{F})$ and a right-continuous filtration $(\mathcal{F}_t)_{t\geq 0}$ on Ω with respect to which the random variables $X(t)$ are adapted, satisfying

(a) $\mathbb{P}^x(X(0) = x) = 1$,

(b) the mapping $x \mapsto \mathbb{E}^x[f(X(t))]$ is in $\mathcal{C}(E)$ for all $f \in \mathcal{C}(E)$ and $t \geq 0$,

(c) $\mathbb{E}^x[Y \circ \theta_s \,|\, \mathcal{F}_s] = \mathbb{E}^{X(s)}[Y]$, $\mathbb{P}^x$-almost surely, for all $x \in E$ and all bounded measurable Y on Ω. Here $\theta_s\omega(t) := \omega(t + s)$ denotes a time shift.

Remark. The last property in Definition 1.1.1 is the homogeneous Markov property.

The Feller processes can be described using a particular version of semigroup theory.

Definition 1.1.2. A *semigroup* is a family of continuous linear operators $\big(P(t)\big)_{t\geq 0}$ on $\mathcal{C}(E)$ satisfying the following properties:

(i) $P(0)f = f$ for all $f \in \mathcal{C}(E)$,

(ii) $\lim_{t\to 0} P(t)f = f$, for all $f \in \mathcal{C}(E)$,

(iii) $P(s + t)f = P(s)P(t)f$, for all $f \in \mathcal{C}(E)$,

(iv) $P(t)f \geq 0$, for all non-negative $f \in \mathcal{C}(E)$.

Definition 1.1.3. A *probability semigroup* on $\mathcal{C}(E)$ is a semigroup $\big(P(t)\big)_{t\geq 0}$ on $\mathcal{C}(E)$ such that there exist $f_n \in \mathcal{C}(E)$, $n \in \mathbb{N}$, such that $\sup_n ||f_n|| < \infty$ and $P(t)f_n$ converges to 1 pointwise for each $t \geq 0$.

Remark. On compact spaces E, one can simply define a probability semigroup $\big(P(t)\big)_{t\geq 0}$ as a semigroup such that $P(t)1 = 1$, for each $t \geq 0$.

Definition 1.1.4. An *infinitesimal generator* is a linear operator $\mathcal{L}$ on $\mathcal{C}(E)$ (possibly unbounded), with domain $\mathcal{D}_{\mathcal{L}}$ and range $\mathcal{R}(\mathcal{L})$, satisfying the following properties:

1. $\mathcal{D}_{\mathcal{L}}$ is dense in $\mathcal{C}(E)$,
2. if $f \in \mathcal{D}_{\mathcal{L}}$, $\lambda \geq 0$ and $g := f - \lambda\mathcal{L}f$, then

$$\inf_{x\in E} f(x) \geq \inf_{x\in E} g(x),$$

3. $\mathcal{R}(\mathbb{1} - \lambda\mathcal{L}) = \mathcal{C}(E)$, for sufficiently small $\lambda > 0$, where $\mathbb{1}$ denotes the identity function.

Definition 1.1.5. A *probability generator* is an infinitesimal generator $\mathcal{L}$ such that, for small $\lambda > 0$, there exist $f_n \in \mathcal{D}_{\mathcal{L}}$, $n \in \mathbb{N}$, so that $g_n := f_n - \lambda\mathcal{L}f_n$ satisfies $||g_n|| < \infty$ and both f_n and g_n converge to 1 pointwise.

Remark. On compact spaces E, we can simply say that a probability generator is an infinitesimal generator $\mathcal{L}$ such that $\mathcal{L}1 = 0$.

The following two results underline the connection of a Feller process with its generator, through its probability semigroup.

Proposition 1.1.6. *Given a Feller process, define*

$$P(t)f(x) \;=\; \mathbb{E}^x f(X(t)),$$

for $f \in \mathcal{C}(E)$. Then $P(t)$ is a probability semigroup on $\mathcal{C}(E)$.

Proof. See [17], Theorem 3.15.

$\square$

Theorem 1.1.7 (Hille-Yosida)**.** *Suppose that $P(t)$ is a probability semigroup and define $\mathcal{L}$ by*

$$\mathcal{L}f := \lim_{t\to 0} \frac{P(t)f - f}{t} \tag{1.1}$$

on

$$\mathcal{D}_{\mathcal{L}} := \{f \in \mathcal{C}(E) \,|\, \text{the limit (1.1) exists}\}.$$

Then $\mathcal{L}$ is a probability generator. Furthermore, the following statements hold:

- *if $f \in \mathcal{D}_{\mathcal{L}}$, then $P(t)f \in \mathcal{D}_{\mathcal{L}}$ for all $t \geq 0$, and it is a continuously differentiable function on t and satisfies*

$$\frac{d}{dt}P(t)f \;=\; P(t)\mathcal{L}f \;=\; \mathcal{L}P(t)f;$$

- *for $f \in \mathcal{C}(E)$ and $t > 0$,*

$$\lim_{n\to\infty} \left(\mathbb{1} - \frac{t}{n}\mathcal{L}\right)^{-n} f \;=\; P(t)f.$$

Proof. See [17], Theorem 3.16.

$\square$

Conversely, given a probability generator $\mathcal{L}$, we can construct the associated probability semigroup and then the corresponding Feller process, which means constructing the measures $\mathbb{P}^x$ in Definition 1.1.1.

Proposition 1.1.8. *Let $\mathcal{L}$ be a probability generator on $\mathcal{C}(E)$, in the sense of Definition 1.1.5. Then, for sufficiently small $\varepsilon > 0$, the operator*

$$\mathcal{L}_\varepsilon := \mathcal{L}(\mathbb{1} - \varepsilon\mathcal{L})^{-1}$$

is a probability generator and

$$T_\varepsilon(t) := \sum_{n=0}^{\infty} \frac{t^n \mathcal{L}_\varepsilon^n}{n!},$$

for $t \geq 0$, is a well defined probability semigroup. Moreover, for $f \in \mathcal{C}(E)$,

$$P(t)f \;=\; \lim_{\varepsilon\to 0} T_\varepsilon(t)f$$

exists uniformly on bounded t intervals. It defines a probability semigroup whose generator is $\mathcal{L}$ in the sense of Theorem 1.1.7.

Proof. See [17], Theorem 3.24. □

Proposition 1.1.9. *If $P(t)$ is a probability semigroup, then there is a Feller process $X(t)$ satisfying*

$$\mathbb{E}^x[f(X(t))] \;=\; P(t)f(x),$$

for $x \in E$, $t \geq 0$ and $f \in \mathcal{C}(E)$.

Proof. See [17], Theorem 3.26. □

1.2 Pure Jump Markov Processes

In this section, we introduce a particular class of Markov processes, the *Pure Jump Processes*, which will be the main object of our study.

Definition 1.2.1. A **pure jump process** is a right-continuous Markov process such that there exists a sequence of strictly 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 every $n \geq 0$.

To describe these jumps, it is usual to introduce the *escape rate function* $\lambda(x)$, such that $\lambda(x)dt + o(dt)$ is the probability that X_t undergoes a jump during $[t, t+dt]$ starting from the state $X_t = x$. When a jump occurs, X_{t+dt} is then distributed with the kernel $p(x, dy)$, so that the overall transition rate is

$$W(x, dy) := \lambda(x)p(x, dy)$$

for $(x, y) \in E^2$. Over a time interval $[0, T]$, the path of a pure jump process can thus be represented by the sequence of visited states in E, together with the sequence of waiting times in those states, so that the space of paths is equivalent to $[E\times(0,\infty)]^{\mathbb{N}}$.

Assumption 1.2.2. From now on, we assume:

- $\lambda(\cdot) : E \to [0, \infty)$ to be a non-negative, bounded and continuous function;

- $x \mapsto p(x, A)$ is a continuous function for every $A \in \mathcal{B}(E)$.

Remark. Under Assumption 1.2.2, the pure jump process is a Feller process and possesses a probability generator, given by

$$(\mathcal{L}f)(x) = \int_E W(x, dy)[f(y) - f(x)], \quad \forall f \in \mathcal{C}(E),\, x \in E.$$

In particular, the domain of $\mathcal{L}$ is $\mathcal{D}_{\mathcal{L}} = \mathcal{C}(E)$.

1.3 The Large Deviation Principle

In Chapter 2 we will consider the problem of conditioning a pure jump Markov process on a rare event and we will be interested in studying the long-time limit behaviour of the conditioned process. In particular, we will provide a numerical procedure for evaluating the *scaled cumulant generating function* (SCGF), which plays an essential role in the study of large deviations.

The purpose of this section is to define, in a rigorous way, the *large deviation events* we are going to consider and to motivate the importance of the SCGF in the study of the large deviation conditioning problem. The main reference for this section is [9].

Throughout this section, E is a Hilbert space, with inner product denoted by " $\cdot$ ". The large deviation principle (LDP) characterises the limiting behaviour as $n \to \infty$ of a family of probability measures (μ_n) on E in terms of a rate function.

Definition 1.3.1. The function $I : E \to [0, \infty]$ is called a **good rate function** if:

(D1) $I \not\equiv \infty$.

(D2) I is lower semi-continuous.

(D3) I has compact level sets, that is $\{x \in E \,|\, I(x) \le c\}$ is compact for all $c \in [0, \infty)$.

We define the corresponding set function by

$$I(S) = \inf_{x \in S} I(x), \quad S \subset E.$$

Definition 1.3.2. Let a_n be a sequence of positive real numbers such that $\lim_{n\to\infty} a_n = \infty$. The sequence of probability measures $(\mathbb{P}_n)$ on E is said to satisfy the **large deviation principle** (LDP) with speed a_n and with rate function I if:

(D1') I is a rate function in the sense of Definition 1.3.1.

(D2') $\limsup_{n\to\infty} \frac{1}{a_n} \log \mathbb{P}_n(C) \le -I(C) \quad \forall C \subset E$ closed.

(D3') $\liminf_{n\to\infty} \frac{1}{a_n} \log \mathbb{P}_n(O) \ge -I(O) \quad \forall O \subset E$ open.

Definition 1.3.3. Let $\mathbb{P}_n$, $n \in \mathbb{N}$, be a family of probability measures on E. The *cumulant generating functions* $\Lambda_n : E \to (-\infty, +\infty]$ of $\mathbb{P}_n$ are defined to be

$$\Lambda_n(k) = \log \int_E e^{k \cdot x} \mathbb{P}_n(dx), \quad k \in E.$$

Also, let

$$\overline{\Lambda}(k) := \limsup_{n\to\infty} \frac{1}{n} \Lambda_n(nk),$$

using the notation $\Lambda(k)$ whenever the limit exists. In accordance with the statistical mechanics notation, we call $\Lambda(k)$ the *scaled cumulant generating function* (SCGF).

Remark. In statistical mechanics, $-\Lambda(k)$ can be seen as the *canonical free energy* [18].

Theorem 1.3.4. *Let E be a Hilbert space and assume that $\mathbb{P}_n$ satisfies the LDP with a good rate function I and that*

$$\overline{\Lambda}(k) < \infty, \quad k \in E.$$

Then, the following statements hold:

- *for each $k \in E$, the limit $\Lambda(k)$ exists, is finite and satisfies*

$$\Lambda(k) \; = \; \sup_{x \in E} \{k \cdot x - I(x)\};$$

- *if I is convex, then*

$$I(x) \; = \; \sup_{k \in E} \{k \cdot x - \Lambda(k)\} =: \Lambda^*(x);$$

- *if I is not convex, then Λ^* is the affine regularisation (or lower convex hull) of I, i.e. $\Lambda^* \le I$ and $f \le \Lambda^*$, for any convex rate function f such that $f \le I$.*

Proof. The proof can be found in [9], Theorem 4.5.10. The first statement follows directly from Varadhan's Lemma (see [9], Theorem 4.3.1), whereas the last two results are an application of the Duality Lemma (see [9], Lemma 4.5.8).

□

Theorem 1.3.4 allows us to identify convex rate functions as Λ^*, that is as functions of the SCGF Λ.

Chapter 2

Feynamn-Kac Models for Conditioned Jump Processes

2.1 The Large Deviation Conditioning Problem

In this section we provide a brief overview of the large-deviation conditioning problem [1].

We consider a homogeneous continuous-time pure jump Markov process $(X_t)_{t\geq 0}$ with infinitesimal generator $\mathcal{L}$ and taking values on a *compact and separable Hilbert space* E (in particular, E is also Polish), so that all the results stated in Chapter 1 hold. We also recall that $\mathcal{C}(E)$ denotes the space of continuous real-valued functions on E.

Let $\Omega := D([0,\infty), E)$ and the right-continuous filtration $(\mathcal{F}_t)_{t\geq 0}$ on Ω is such that the σ-algebra $\mathcal{F}_t$ is the smallest such that the mapping $\omega \mapsto \omega(t)$, with $\omega \in \Omega$, is $\mathcal{F}_t$-measurable for each $t \geq 0$. Given a generator $\mathcal{L}$, we denote by $\mathbb{P}_{\mu_0,\mathcal{L}}$ the corresponding probability measure on $(\Omega, \mathcal{F})$ with initial distribution μ_0, and by $\mathbb{E}_{\mu_0,\mathcal{L}}$ the associated expectation.

Let $A_T : \Omega \to \mathbb{R}$ be a $\mathcal{F}_T$-measurable random variable or *observable*, taken to be a real measurable function of the paths of X_t over the time interval $[0, T]$, and condition X_t on a general measurable event of the form $\mathcal{A}_T = \{A_T \in B\}$ with $B \subseteq \mathbb{R}$ measurable subset, that is, we condition X_t on the subset

$$\mathcal{A}_T \;=\; \{\omega \in \Omega_T \,|\, A_T(\omega) \in B\} \;=\; A_T^{-1}(B)$$

of sample paths satisfying the constraint that $A_T \in B$.

Assumption 2.1.1. The observable A_T is assumed to satisfy a large deviation principle with respect to $\mathbb{P}_{\mu_0,\mathcal{L}}$ (i.e. the sequence of measures $\mathbb{P}_{\mu_0,\mathcal{L}} \circ A_T^{-1}$ satisfy an LDP in the sense of Definition 1.3.2), and we assume that the rate function I exists and takes values different from 0 or ∞. Moreover, we demand that A_T is of the form

$$A_T(\omega) = \frac{1}{T} \sum_{\substack{s \le T \\ \omega(s_-) \neq \omega(s)}} g\big(\omega(s_-),\, \omega(s)\big) ds, \tag{2.1}$$

where $g \in \mathcal{C}(E^2)$ and $\omega \in \Omega$ realisation of $(X_t)_{t \ge 0}$. Note that A_T is well defined since we have assumed that the pure jump process doesn't explode.

Example. The considered class of observables A_T includes many random variables of mathematical and physical interest, such that:

- the position of the final state X_T with respect to the initial position X_0 on a compact state space $S \subset \mathbb{R}^d$, which can be obtained by considering $g(x, y) = y - x$;
- the particle current on a finite state space $S \subset \mathbb{Z}^d$ across a particular bond $(i,\, j)$, with $||i - j|| = 1$, obtained with $g(x, y) = \mathbb{1}_{\{i\}}(x) \cdot \mathbb{1}_{\{j\}}(y) - \mathbb{1}_{\{j\}}(x) \cdot \mathbb{1}_{\{i\}}(y)$;
- the action functional [16] on finite state spaces with jump rates $W(i, j)$, obtained by setting $g(x, y) = \log \frac{W(x,y)}{W(y,x)}$.

In the study of the long-time limit behaviour of the probability path measure of the conditioned process, i.e.

$$\mathbb{P}_{\mu_0, \mathcal{L}}(d\omega \,|\, A_T \in B), \tag{2.2}$$

a key role is played by the scaled cumulant generating function of the observable A_T

$$\Lambda_{\mu_0}(k) = \lim_{T \to \infty} \frac{1}{T} \log \mathbb{E}_{\mu_0,\, \mathcal{L}}\big[e^{kTA_T}\big], \tag{2.3}$$

with $k \in \mathbb{R}$, as illustrated in Section 1.3. Recalling Theorem 1.3.4, we know that $\Lambda_{\mu_0}(k)$ exists and is finite provided that

$$\limsup_{T \to \infty} \frac{1}{T} \log \mathbb{E}_{\mu_0,\, \mathcal{L}}\big[e^{kTA_T}\big] < \infty.$$

In what follows, we are interested in providing a numerical procedure for evaluating the SCGF. For this purpose, we start by introducing an auxiliary infinitesimal generator.

Proposition 2.1.2. *For pure jump processes, the family of operators*

$$P_k(t) f(x) \,:=\, \mathbb{E}_{x, \mathcal{L}}\big[f\big(X_t\big)\, e^{ktA_t}\big],$$

with $f \in \mathcal{C}(E)$, is well defined and it is a semigroup in the sense of Definition 1.1.2.

Moreover, the infinitesimal generator associated with $P_k(t)$, $t \geq 0$, in the sense of Theorem 1.1.7, can be written in the form

$$(\mathcal{L}_k f)(x) = \int_E W(x, dy)[e^{kg(x,y)} f(y) - f(x)] \,, \tag{2.4}$$

for $f \in \mathcal{C}(E)$ and all $x \in E$, with g defined in (2.1). In particular, the semigroup $P_k(t)$ satisfies the differential equation

$$\frac{d}{dt} P_k(t) f \;=\; P_k(t) \mathcal{L} f, \tag{2.5}$$

for all $f \in \mathcal{C}(E)$ and $t > 0$.

Proof. See [1], Appendix A.1.

□

Remark. The condition $\sup_{x,y \in E} g(x, y) < \infty$ guarantees the domain of $\mathcal{L}_k$ to be all $\mathcal{C}(E)$. Also, observe that $P_k(t)$ is not a probability semigroup and $\mathcal{L}_k$ is not a probability generator, since they do not preserve the probability.

2.2 The Feynman-Kac Interpretation

In this section, we are interested in interpreting the SCGF $\Lambda_{\mu_0}(k)$ through Feynman-Kac models, so that we will be able to apply already established results from Feynman-Kac theory [2, 5, 6].

For this purpose, we introduce the potential function

$$\widehat{G}(x) := \int_{y \in E} p(x, dy) e^{kg(x,y)} > 0, \tag{2.6}$$

for all $x \in E$. Observe that $\widehat{G}$ is bounded since $\sup_{x,y \in E} g(x, y) < \infty$.

Furthermore, we define a new pure jump process given by the escape rate functions $\widehat{\lambda}(x) := \widehat{G}(x) \cdot \lambda(x)$, $x \in E$, and kernel

$$\widehat{p}(x, dy) := \frac{p(x, dy) e^{kg(x,y)}}{\widehat{G}(x)},$$

so that the overall transition rate is

$$\widehat{W}(x, dy) := \widehat{\lambda}(x) \cdot \widehat{p}(x, dy) = W(x, dy) e^{kg(x,y)}, \tag{2.7}$$

and its generator is given by

$$\widehat{\mathcal{L}}(f)(x) := \int_E \widehat{W}(x, dy)[f(y) - f(x)], \tag{2.8}$$

for all $f \in \mathcal{C}(E)$ and $x \in E$.

Lastly, we denote

$$\mathcal{V}(x) := \lambda(x) \cdot \big(\widehat{G}(x) - 1\big). \tag{2.9}$$

Remark. Observe that under Assumption 1.2.2 and assuming $g \in \mathcal{C}(E^2)$, then $\mathcal{V} \in \mathcal{C}(E)$.

Proposition 2.2.1. *The tilted generator $\mathcal{L}_k$ can be written in terms of $(\widehat{\mathcal{L}}, \widehat{G})$ as*

$$\mathcal{L}_k(f)(x) \;=\; \widehat{\mathcal{L}}(f)(x) \;+\; \mathcal{V}(x) \cdot f(x), \tag{2.10}$$

for all $f \in \mathcal{C}(E)$ and $x \in E$.

Proof. The proof is given by using some elementary manipulations

$$\begin{aligned}
\mathcal{L}_k(f)(x) &= \int_E W(x,dy)\big[e^{kg(x,y)} f(y) - f(x)\big] \\
&= \int_E W(x,dy) \cdot e^{kg(x,y)} [f(y) - f(x)] \;+\; \int_E W(x,dy) \left(e^{kg(x,y)} - 1\right) \cdot f(x) \\
&= \widehat{\mathcal{L}}(f)(x) \;+\; \lambda(x) \left(\widehat{G}(x) - 1\right) \cdot f(x),
\end{aligned}$$

for all $f \in \mathcal{C}(E)$ and $x \in E$.

□

It is important to underline the connection between the generator (2.10) and the classical Feynman-Kac theorem.

Theorem 2.2.2 (Feynman-Kac Formula)**.** *Suppose that (X_t) is a Feller process on the state space E with generator $\mathcal{L}$, path measure $\mathbb{P}$ and expectation $\mathbb{E}$, and take $h \in \mathcal{C}(E)$. Define*

$$u_{t,x}(f) := \mathbb{E}_x \left[f(X_t) \exp\left(\int_0^t h(X_s) ds \right) \right], \tag{2.11}$$

for all $f \in \mathcal{D}_{\mathcal{L}}$. Then $u_{t,\cdot}(f) \in \mathcal{D}_{\mathcal{L}}$ for each $t \geq 0$ and $f \in \mathcal{D}_{\mathcal{L}}$. Moreover $u_{t,x}$ solves the equation

$$\frac{d}{dt} u_{t,x}(f) = \mathcal{L} u_{t,x}(f) + h(x) \cdot u_{t,x}(f), \qquad u_{0,x}(f) = f(x), \tag{2.12}$$

for all $f \in \mathcal{D}_{\mathcal{L}}$.

Proof. See [17], Theorem 3.47.

□

Theorem 2.2.2 states that the solution of the dynamical system

$$\frac{d}{dt} u_{t,x}(f) \;=\; \mathcal{L}_k u_{t,x}(f) \;=\; \widehat{\mathcal{L}} u_{t,x}(f) \;+\; \mathcal{V}(x) \cdot u_{t,x}(f)$$

is given by

$$\nu_{t,x}(f) := \mathbb{E}_{x,\widehat{\mathcal{L}}} \left[f(\widehat{X}_t) \exp\left(\int_0^t \mathcal{V}(\widehat{X}_s) ds \right) \right], \tag{2.13}$$

for any $f \in \mathcal{D}_{\widehat{\mathcal{L}}} = \mathcal{C}(E)$. Recalling the differential equation (2.5), we also have

$$\nu_{t,x}(f) = P_k(t)f(x),$$

for all $t > 0$, $x \in E$ and $f \in \mathcal{C}(E)$. We extend the definition of the measure $\nu_{t,x}$ for general initial distributions $\mu_0 \in \mathcal{P}(E)$ via

$$\nu_{t,\mu_0}(f) := \mu_0\big(\nu_{t,\cdot}(f)\big) = \mathbb{E}_{\mu_0,\widehat{\mathcal{L}}}\Big[f(\widehat{X}_t)\exp\Big(\int_0^t \mathcal{V}(\widehat{X}_s)ds\Big)\Big],$$

for all $t > 0$ and $f \in \mathcal{C}(E)$. In the mathematical literature, ν_{t,μ_0} is known as the *unnormalised t-marginal Feynman-Kac measure*. However, it will be more convenient to work with the corresponding *normalised t-marginal Feynman-Kac measure*

$$\mu_{t,\mu_0}(f) := \frac{\mathbb{E}_{\mu_0,\widehat{\mathcal{L}}}\Big[f(\widehat{X}_t)\exp\Big(\int_0^t \mathcal{V}(\widehat{X}_s)ds\Big)\Big]}{\mathbb{E}_{\mu_0,\widehat{\mathcal{L}}}\Big[\exp\Big(\int_0^t \mathcal{V}(\widehat{X}_s)ds\Big)\Big]} = \frac{\nu_{t,\mu_0}(f)}{\nu_{t,\mu_0}(1)}, \tag{2.14}$$

defined for any $t > 0$ and $f \in \mathcal{C}(E)$.

The scaled cumulant generating function $\Lambda_{\mu_0}(k)$ can be rewritten easily in terms of the normalised marginal measure μ_{t,μ_0}, as shown in the following result. In the next sections we will show how this interpretation of the SCGF allows us to construct a numerical procedure for evaluating $\Lambda_{\mu_0}(k)$ and evaluating rigorously the convergence of the algorithm to the correct value.

Proposition 2.2.3. *For every $t > 0$, we have that*

$$\log \mathbb{E}_{\mu_0,\mathcal{L}}\big[e^{ktA_t}\big] = \int_0^t \mu_{s,\mu_0}(\mathcal{V})ds,$$

where $\mathcal{V}$ is defined in (2.9)*. In particular, the limit*

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t \mu_{s,\mu_0}(\mathcal{V})ds$$

exists and is finite if and only if the SCGF (2.3) *is well defined and finite. Moreover, it can be written in the form*

$$\Lambda_{\mu_0}(k) = \lim_{t\to\infty} \frac{1}{t}\int_0^t \mu_{s,\mu_0}(\mathcal{V})ds. \tag{2.15}$$

Proof. First recall that

$$\nu_{t,\mu_0}(f) = \mathbb{E}_{\mu_0,\mathcal{L}}\big[f(X_t)e^{kt\cdot A_t}\big],$$

for all $f \in \mathcal{C}(E)$. In particular,

$$\nu_{t,\mu_0}(1) = \mathbb{E}_{\mu_0,\mathcal{L}}\big[e^{ktA_t}\big]. \tag{2.16}$$

Recalling the Feynman-Kac Formula 2.2.2, we obtain

$$\frac{d}{dt}\log\nu_{t,\mu_0}(1) \;=\; \frac{1}{\nu_{t,\mu_0}(1)}\cdot\frac{d}{dt}\nu_{t,\mu_0}(1) \;=\; \frac{\nu_{t,\mu_0}(\mathcal{L}_k 1)}{\nu_{t,\mu_0}(1)} \;=\; \mu_{t,\mu_0}(\mathcal{L}_k 1).$$

And, thus,

$$\nu_{t,\mu_0}(1) \;=\; \exp\biggl(\int_0^t \mu_{s,\mu_0}(\mathcal{L}_k 1)\,ds\biggr),$$

since $\nu_0(1) = 1$. We can conclude by (2.16), observing that $\mathcal{L}_k(1)(x) = \mathcal{V}(x)$.

□

Remark. Lemma 2.2.3 states in particular that $\Lambda_{\mu_0}(k)$ can be evaluated as the time average of $\mu_{t,\mu_0}(\mathcal{V})$. Therefore, the basic idea in the proposed cloning algorithm is to approximate the measure μ_{t,μ_0} through the evolution of an N-particle system and, thus, obtaining also an estimation of $\Lambda_{\mu_0}(k)$, using Proposition 2.2.3.

2.3 Integro-differential equations

In this section, we are interested in outlining the evolution equation of the time marginal μ_{t,μ_0} in terms of interacting jump-type probability generators and in describing its evolution. The content presented here is based on the works of Del Moral and Miclo [2, 5], since, as we have seen in the previous sections, we can interpret the large-deviation conditioning problem through the Feynman-Kac theory and thus apply already established results.

In all this section the initial distribution μ_0 is fixed and we simplify the notation by writing μ_t instead of μ_{t,μ_0}.

Theorem 2.3.1. *For every $f \in \mathcal{C}(E)$ and $t \geq 0$, the normalised t-marginal μ_t solves the evolution equation*

$$\frac{d}{dt}\mu_t(f) = \mu_t\bigl(\widehat{\mathcal{L}}f \,+\, \widetilde{\mathcal{L}}_{\mu_t} f\bigr), \tag{2.17}$$

with $\widehat{\mathcal{L}}$ given by (2.8) *and*

$$\widetilde{\mathcal{L}}_\mu(f)(x) := \int_E \bigl(f(y) - f(x)\bigr)\bigl(\mathcal{V}(y) - \mathcal{V}(x)\bigr)_+\,\mu(dy), \tag{2.18}$$

depending on $\mu \in \mathcal{P}(E)$, where $a_+ := a \vee 0$.

Proof. Using the fact that $\frac{d}{dt}\nu_t(f) = \nu_t(\mathcal{L}_k f)$, we have

$$\begin{aligned}
\frac{d}{dt}\mu_t(f) &= \frac{d}{dt}\frac{\nu_t(f)}{\nu_t(1)} \\
&= \frac{1}{\nu_t(1)}\cdot\nu_t(\mathcal{L}_k f) - \frac{\nu_t(f)}{\nu_t(1)^2}\,\nu_t(\mathcal{L}_k 1) \\
&= \mu_t(\mathcal{L}_k f) - \mu_t(f)\cdot\mu_t(\mathcal{L}_k 1) \\
&= \mu_t(\widehat{\mathcal{L}}f) + \mu_t(\mathcal{V}f) - \mu_t(f)\cdot\mu_t(\mathcal{V}),
\end{aligned}$$

where the last equality follows by (2.10). Thus, the statement follows by showing

$$\begin{aligned}\mu_t(\mathcal{V}f) \,-\, \mu_t(\mathcal{V})\mu_t(f) &= \mu_t\big(\widetilde{\mathcal{L}}_{\mu_t}(f)\big)\\ &= \int_{E^2} \big(f(y)-f(x)\big)\,\big(\mathcal{V}(y)-\mathcal{V}(x)\big)_+\,\mu_t(dy)\,\mu_t(dx).\end{aligned}$$

Observe that

$$\begin{aligned}\mu_t(\mathcal{V}f) \,-\, \mu_t(\mathcal{V})\mu_t(f) &= \int_E \mathcal{V}(x)f(x)\,\mu_t(dx) \,-\, \int_{E^2} \mathcal{V}(y)f(x)\,\mu_t(dy)\mu_t(dx)\\ &= \int_{E^2} \big(\mathcal{V}(x)-\mathcal{V}(y)\big)\,f(x)\,\mu_t(dy)\mu_t(dx)\\ &= \int_{E^2} \Big[\big(\mathcal{V}(x)-\mathcal{V}(y)\big)_+\,f(x) \,-\, \big(\mathcal{V}(x)-\mathcal{V}(y)\big)_-\,f(x)\Big]\,\mu_t(dy)\mu_t(dx),\end{aligned}$$

where $a_- := -a \vee 0 = (-a)_+$. By a change of variables, we can see that

$$\begin{aligned}\int_{E^2} \big(\mathcal{V}(x)-\mathcal{V}(y)\big)_-\,f(x)\,\mu_t(dy)\mu_t(dx) &= \int_{E^2} \big(\mathcal{V}(y)-\mathcal{V}(x)\big)_-\,f(y)\,\mu_t(dy)\,\mu_t(dx)\\ &= \int_{E^2} \big(\mathcal{V}(x)-\mathcal{V}(y)\big)_+ f(y)\mu_t(dy)\,\mu_t(dx).\end{aligned}$$

Thus,

$$\mu_t(\mathcal{V}f) \,-\, \mu_t(\mathcal{V})\mu_t(f) \qquad = \int_{E^2} \big(f(x)-f(y)\big)\big(\mathcal{V}(x)-\mathcal{V}(y)\big)_+\mu_t(dy)\,\mu_t(dx).$$

This concludes the proof.

□

Remark. Observe that, for every $\mu \in \mathcal{P}(E)$, the generator $\widetilde{\mathcal{L}}_\mu$ (2.18) corresponds to the generator of a jump process with overall transition rates

$$\widetilde{W}_\mu(x,dy) := \big(\mathcal{V}(y)-\mathcal{V}(x)\big)_+ \cdot \mu(dy).$$

The evolution equation (2.17) can be interpreted as the evolution of the $\mathrm{Law}(\overline{X}_t) = \mu_t$ of a time-inhomogeneous process $\overline{X}_t$ with a probability generator $\overline{\mathcal{L}}_{\mu_t} := \widehat{\mathcal{L}} + \widetilde{\mathcal{L}}_{\mu_t}$. In this situation, it is essential to observe that $\overline{\mathcal{L}}_{\mu_t}$ depends on the distribution μ_t of the random state $\overline{X}_t$.

Definition 2.3.2. The stochastic model $\overline{X}_t$, or equivalently the collection of probability generators $(\overline{\mathcal{L}}_\mu)_{\mu\in\mathcal{P}(E)}$, is called a *McKean interpretation* of the non-linear evolution equation in distribution space defined in (2.17).

We conclude the section providing a description of the process $\overline{X}_t$. For simplicity, we call the jumps given by $\widehat{\mathcal{L}}$ *mutation events* and the jumps given by $\widetilde{\mathcal{L}}_{\mu_t}$ *selection events.*

Between the selection events, the process $\overline{X}_t$ evolves as a copy of the reference process $\widehat{X}_t$ with generator $\widehat{\mathcal{L}}$. The rate of the selection events is given by the function

$$\widetilde{\lambda}_{\mu_t}(x) \;:=\; \mu_t\big((\mathcal{V}-\mathcal{V}(x))_+\big).$$

In particular, no selection event occurs when $\mu_t\big((\mathcal{V}-\mathcal{V}(x))_+\big) = 0$, that is when $\mathcal{V}(x) = \mu_t - \operatorname{ess\,sup} \mathcal{V}$.

More formally, the selection times $(T_n)_{n\geq 0}$ are given by the recursive formulae

$$T_{n+1} := \inf\{t \geq T_n \,|\, \int_{T_n}^{t} \widetilde{\lambda}_{\mu_s}(\overline{X}_s)\, ds \geq e_n\},$$

where $T_0 = 0$ and $(e_n)_{n\geq 0}$ stands for a sequence of i.i.d. exponential random variables with unit parameter.

At the jump time T_n, the process $\overline{X}_{T_n-} = x$ jumps to a new site $\overline{X}_{T_n} = y$ randomly chosen with the distribution

$$\widetilde{p}_{T_n-}(x, dy) := \frac{\big(\mathcal{V}(y)-\mathcal{V}(x)\big)_+}{\mu_{T_n-}\big((\mathcal{V}-\mathcal{V}(x))_-\big)}\,\mu_{T_n-}(dy).$$

Definition 2.3.3. Given a non-negative function $\phi \in \mathcal{C}(E)$, the *Boltzmann-Gibbs mapping* $\Psi_\phi : \mathcal{P}_\phi(E) \to \mathcal{P}(E)$, with $\mathcal{P}_\phi(E) := \{\mu \in \mathcal{P}(E) \,|\, \mu(\phi) > 0\}$, is defined for any $\mu \in \mathcal{P}_\phi(E)$ by

$$\Psi_\phi(\mu)(dx) := \frac{1}{\mu(\phi)}\,\phi(x)\,\mu(dx), \tag{2.19}$$

for every $x \in E$.

Observe that the distribution $\widetilde{p}_{t-}(x,\cdot)$ can be rewritten as the Boltzmann-Gibbs mapping $\Psi_\phi(\mu_{t-})$ associated to the non-negative bounded measurable function $\phi(\cdot) := \big(\mathcal{V}(\cdot)-\mathcal{V}(x)\big)_+$. Moreover, by construction $\mu_t(\phi)$ is strictly positive when a selection event occurs, therefore the Boltzmann-Gibbs mapping associated to ϕ is well defined.

We conclude the chapter providing two simple examples that illustrate the construction of the overall transition rates $\widehat{W}$ and $\widetilde{W}_\mu$, with $\mu \in \mathcal{P}(E)$, which provides the construction of the McKean model.

Example 1. Let $S_M = \mathbb{Z}^d \cap [-M, M]^d$, with $M \in \mathbb{N}$, with periodic boundary conditions $M+1 = -M$ and $-M-1 = M$. Consider the continuous-time random walk defined by overall transition rates $W(x, y) = \frac{\lambda}{2^d}$ with $\lambda > 0$, if $||x-y|| = 1$, and $W(x, y) = 0$ otherwise. We are interested in counting the number of times the particle current crosses a particular bond (i, j) with $||j-i|| = 1$. This is obtained by considering

$$g(x,y) := \mathbb{1}_{\{i\}}(x)\cdot \mathbb{1}_{\{j\}}(y) \;+\; \mathbb{1}_{\{j\}}(x)\cdot \mathbb{1}_{\{i\}}(y).$$

Therefore, we can write the potentials $\widehat{G}$ and $\mathcal{V}$ in the form

$$\widehat{G}(x) \;=\; \sum_{||y-x||=1} \frac{e^{kg(x,y)}}{2^d} \;=\; \begin{cases} \frac{e^k}{2^d} + \frac{2^d-1}{2^d} & x \in \{i,\, j\} \\ 1 & \text{otherwise,} \end{cases}$$

and

$$\mathcal{V}(x) \;=\; \lambda\big(\widehat{G}(x) - 1\big) \;=\; \begin{cases} \frac{\lambda}{2^d}\,(e^k - 1) & x \in \{i,\, j\} \\ 0 & \text{otherwise.} \end{cases}$$

We can thus deduce the overall transition rates $\widehat{W}$ and $\widetilde{W}_\mu$, $\mu \in \mathcal{P}(S_M)$:

$$\widehat{W}(x,y) \;=\; W(x,y)\cdot e^{kg(x,y)} \begin{cases} \frac{\lambda e^k}{2^d} & (x,y) \in \{(i,j),\, (j,i)\} \\ \frac{\lambda}{2^d} & ||x-y|| = 1,\; (x,y) \notin \{(i,j),\, (j,i)\} \\ 0 & \text{otherwise,} \end{cases}$$

and, assuming $k > 0$,

$$\widetilde{W}_\mu(x,\, y) \;=\; \begin{cases} 0 & x \in \{i,\, j\} \text{ or } y \notin \{i, j\} \\ \frac{\lambda}{2^d}(e^k - 1)\,\mu(y) & x \notin \{i,\, j\} \text{ and } y \in \{i, j\}, \end{cases}$$

with $\mu \in \mathcal{P}(S_M)$, for any $x, y \in S_M$.

Example 2. As in Example 1, we consider a random walk on $S_M = \mathbb{Z} \cap [-M,\, M]$ (now with dimension $d = 1$) with periodic boundary condition $M + 1 = -M$ and with transition rates $W(x,\, y) = \frac{\lambda}{2}$ with $\lambda > 0$, if $||x - y|| = 1$, and $W(x,\, y) = 0$ otherwise. This time we are interested in conditioning on the average speed of the particle. This is obtained by considering $g(x,\, y) := y - x$. In this case, we have

$$\widehat{G}(x) \;=\; \sum_{||y-x||=1} \frac{e^{kg(x,y)}}{2} \;=\; \frac{e^k + e^{-k}}{2},$$

for all $x \in S_M$. Therefore, $\mathcal{V}(x) \equiv \lambda\big(\frac{e^k+e^{-k}}{2} - 1\big)$ and thus $\widetilde{W}_\mu(x,y) = 0$ for all $x, y \in S_M$ and $\mu \in \mathcal{P}(S_M)$. This means that the McKean interpretation $\overline{\mathcal{L}}_\mu$ is given only by the pure jump process $\widehat{\mathcal{L}}$ and it is independent of $\mu \in \mathcal{P}(S_M)$. In particular, the McKean interpretation is defined by the transition rates

$$\widehat{W}(x,\, y) \;=\; \begin{cases} \frac{\lambda}{2}e^k & y = x + 1 \\ \frac{\lambda}{2}e^{-k} & y = x - 1. \end{cases}$$

Indeed, conditioning on the average speed of the particle simply leads to a random walk on S_M with local drift k, which is described by the overall transition rates $\widehat{W}(x,y)$, with $x, y \in S_M$.

Chapter 3

The Cloning Algorithm

3.1 Mean Field Particle Interpretation

This section is concerned with the design of the mean field N-particle model $\xi_t := (\xi_t^i)_{1 \le i \le N}$ associated with a given collection of generators $\left(\overline{\mathcal{L}}_\mu\right)_{\mu \in \mathcal{P}(E)}$ satisfying the weak evolution equation (2.17). This N-particle model is a Markov process in E^N, with homogeneous infinitesimal generator $\overline{\mathcal{L}}^N_{m(x)}$ defined for $F \in \mathcal{C}(E^N)$ by the formulae

$$\overline{\mathcal{L}}^N_{m(x)}(F)(x^1, \ldots, x^N) := \sum_{i=1}^N \overline{\mathcal{L}}^{(i)}_{m(x)}(F)(x^1, \ldots, x^i, \ldots, x^N),$$

where $m : E^N \to \mathcal{P}(E)$ is defined by $m(x) := \frac{1}{N} \sum_{1 \le i \le N} \delta_{x^i}$, with $x = (x^1, \ldots, x^N) \in E^N$, and $\overline{\mathcal{L}}^{(i)}_{m(x)}$ stands for the operator $\overline{\mathcal{L}}_{m(x)}$ acting on the function $x^i \mapsto F(x^1, \ldots, x^i, \ldots, x^N)$.

By construction, the generator $\overline{\mathcal{L}}^N_{m(x)}$ associated with $\mathcal{L}_k$ is divided into a mutation generator $\overline{\mathcal{L}}^{mut}$ and an interacting selection generator $\overline{\mathcal{L}}^{sel}_{m(x)}$, respectively defined by

$$\overline{\mathcal{L}}^{mut}(F)(x^1, \ldots, x^N) := \sum_{i=1}^N \widehat{\mathcal{L}}^{(i)}(F)(x^1, \ldots, x^i, \ldots, x^N),$$
$$\overline{\mathcal{L}}^{sel}_{m(x)}(F)(x^1, \ldots, x^N) := \sum_{i=1}^N \widetilde{\mathcal{L}}^{(i)}_{m(x)}(F)(x^1, \ldots, x^N),$$

where $\widehat{\mathcal{L}}^{(i)}$ and $\widetilde{\mathcal{L}}^{(i)}_{m(x)}$ stand respectively for the operator $\widehat{\mathcal{L}}$ and the operator $\widetilde{\mathcal{L}}_{m(x)}$ acting on the function $x^i \mapsto F(x^1, \ldots, x^i, \ldots, x^N)$.

Remark. Observe that, for test functions of the form

$$F(x) = m(x)(f) = \frac{1}{N} \sum_{i=1}^N f(x^i),$$

with $f \in \mathcal{C}(E)$, we have that

$$\overline{\mathcal{L}}^{mut}(F)(x) = m(x)\big(\widehat{\mathcal{L}}(f)\big)$$

and

$$\overline{\mathcal{L}}^{sel}_{m(x)}(F)(x) = m(x)\big(\widetilde{\mathcal{L}}_{m(x)}(f)\big).$$

The mutation generator $\overline{\mathcal{L}}^{mut}$ describes the evolution of the particles between the selection events. Between two selection events, the particles evolve independently with $\widehat{\mathcal{L}}$-motions in the sense that they explore the state space as independent copies of the homogeneous process $\widehat{X}_t$ with generator $\widehat{\mathcal{L}}$.

Recalling the definition of $\widetilde{\mathcal{L}}_\mu$ (2.18), we can rewrite the selection generator $\overline{\mathcal{L}}^{sel}_{m(x)}$ as

$$\begin{aligned}\overline{\mathcal{L}}^{sel}_{m(x)}(F)(x) &= \sum_{1\le i\le N} \int_E [F(\theta^i_u(x)) - F(x)] \big(\mathcal{V}(u) - \mathcal{V}(x^i)\big)_+ m(x)(du) \\ &= \sum_{1\le i\le N} m(x)(\phi_{x^i}) \int_E [F(\theta^i_u(x)) - F(x)]\Psi_{\phi_{x^i}}\big(m(x)\big)(du),\end{aligned}$$

where $x = (x^1, \ldots, x^N) \in E^N$, $u \in E$, $\theta^i_u : x \mapsto (x^1, \ldots, x^{i-1}, u, x^{i+1}, \ldots, x^N)$, $\phi_u(\cdot) := \big(\mathcal{V}(\cdot) - \mathcal{V}(u)\big)_+$ and Ψ_ϕ is the Boltzmann-Gibbs mapping (2.19).

In this interpretation, if we denote by T^i_n the n-th jump time of the i-th particle ξ^i_t, we have

$$T^i_{n+1} = \inf\{t \ge T^i_n \mid \int_{T^i_n}^t m(\xi_s)(\phi_{\xi^i_s})ds \ \ge\ e^i_n\},$$

where $(e^i_n)_{n\in\mathbb{N}}$ stands for a sequence of i.i.d. exponential random variables with unit parameter. In other words, the selection rate of the i-particle is given by the average potential variation of the particles with higher values, that is

$$m(x)\big(\phi_{x^i}\big) = \frac{1}{N}\sum_{1\le j\le N} \mathbb{1}_{\mathcal{V}(x^j)>\mathcal{V}(x^i)}\big(\mathcal{V}(x^j) - \mathcal{V}(x^i)\big).$$

In particular, no selection jump occurs in case $\mathcal{V}(x^i) = \max_{1\le j\le N}\mathcal{V}(x^j)$.

At the jump time T^i_n, the process $\xi^i_{T^i_n-} = x^i$ jumps to a new state $\xi^i_{T^i_n} = u$ randomly chosen in the current population with distribution

$$\Psi_{\phi_{x^i}}\big(m(\xi_{T^i_n-})\big) = \sum_{1\le j\le N} \frac{\big(\mathcal{V}(\xi^j_{T^i_n-}) - \mathcal{V}(x^i)\big)_+}{\sum_{1\le j'\le N}\big(\mathcal{V}(\xi^{j'}_{T^i_n-}) - \mathcal{V}(x^i)\big)_+}\,\delta_{\xi^j_{T^i_n-}}.$$

Loosely speaking, it adopts randomly the current state of another particle $\xi^i_{T^i_n} = \xi^j_{T^i_n-}$ among the ones with strictly higher potential value and with a probability proportional to $\mathcal{V}(\xi^j_{T^i_n-}) - \mathcal{V}(\xi^i_{T^i_n-})$.

At the level of the whole population, we can write the selection generator $\overline{\mathcal{L}}^{sel}_{m(x)}$ in the standard form

$$\overline{\mathcal{L}}^{sel}_{m(x)}(F)(x) = \lambda^{sel}(x) \int p^{sel}(x,\, dy)[F(y) - F(x)],$$

with the population selection rate

$$\lambda^{sel}(x) := \sum_{1 \le i \le N} m(x)\big(\phi_{x^i}\big) \; = \frac{1}{N} \sum_{1 \le i,j \le N} \big(\mathcal{V}(x^j) - \mathcal{V}(x^i)\big)_+ \,,$$

for $x \in E^N$, and the population selection transition

$$p^{sel}(x,\, dy) := \sum_{1 \le i,j \le N} \frac{\big(\mathcal{V}(x^j) - \mathcal{V}(x^i)\big)_+}{\sum_{1 \le i',j' \le N} \big(\mathcal{V}(x^{j'}) - \mathcal{V}(x^{i'})\big)_+} \, \delta_{\theta^i_{x^j}(x)}(dy),$$

with $x = (x^1, \ldots, x^N) \in E^N$, $y \in E^N$ and $\theta^i_u : x \mapsto (x^1, \ldots,\, x^{i-1},\, u,\, x^{i+1}, \ldots,\, x^N)$.

3.2 Asymptotic Stability

In this section, we are interested in finding sufficient conditions to have that the limit (2.15) exists. Under these conditions, we will also estimate the quantity

$$\left| \frac{1}{t} \int_0^t \mu_{s,\mu_0}(\mathcal{V}) ds \; - \; \Lambda_{\mu_0}(k) \right|,$$

for $t > 0$ large enough.

For this purpose, we will make use of the asymptotic stability properties of continuous-time Feynman-Kac models [7, 4]. We start introducing the basic notation needed for presenting the asymptotic stability results.

Definition 3.2.1. The *total variation distance* between two probability measures $\eta_1,\, \eta_2 \in \mathbb{P}(E)$ is given by

$$||\eta_1 - \eta_2||_{tv} := \frac{1}{2} \sup_{f \in \mathcal{C}(E)} \frac{|\eta_1(f) - \eta_2(f)|}{||f||}.$$

Definition 3.2.2. Let $\mu_{t,\eta}$ be the normalised t-marginal Feynman-Kac measure (2.14) associated to the initial distribution η. We say that $(\mu_{t,\eta})_{t \ge 0,\, \eta \in \mathcal{P}(E)}$ is *exponentially stable* if there exist $t_0 > 0$ and some constants $\gamma,\, C > 0$ such that

$$||\mu_{t,\eta_1} - \mu_{t,\eta_2}||_{tv} \le C \cdot e^{-\gamma t}, \tag{3.1}$$

for $t \ge t_0$ and any initial distributions $\eta_1,\, \eta_2 \in \mathcal{P}(E)$.

Theorem 3.2.3. *Suppose that $(\mu_{t,\eta})_{t\geq 0,\,\eta\in\mathcal{P}(E)}$ is exponentially stable in the sense of definition 3.2.2, for some $t_0 > 0$ and some parameters γ, $C > 0$. Then for any fixed $\eta \in \mathcal{P}(E)$, the sequence of probability measures $\mu_{t,\eta}$, with $t \geq 0$, converges in total variation, as $t \to \infty$.*

Proof. Since $\mathcal{P}(E)$ is complete with respect to the norm $||\cdot||_{tv}$, it is enough to prove that, for any fixed initial distribution η, the sequence $\mu_{t,\eta}$, with $t \geq 0$, is a Cauchy sequence.

Since $(\mu_{t,\eta})_{t\geq 0,\,\eta\in\mathcal{P}(E)}$ is exponentially stable, we have that

$$||\mu_{t+s,\eta} - \mu_{t,\eta}||_{tv} \;=\; ||\mu_{t,\mu_{s,\eta}} - \mu_{t,\eta}||_{tv} \;\leq\; C \cdot e^{-\gamma\cdot t},$$

for any $t \geq t_0$. Therefore, for any $\delta > 0$ we have that

$$||\mu_{t+s,\eta} - \mu_{t,\eta}||_{tv} \;\leq\; \delta,$$

for any $s > 0$ and $t \;\geq\; \max\{t_0,\, -\frac{\log(\delta/C)}{\gamma}\}$. This concludes the proof. $\square$

Remark. Observe that Theorem 3.2.3 gives us also an estimate of the order of convergence

$$||\eta^* - \mu_{t,\eta}||_{tv} \;\leq C \cdot e^{-\gamma\cdot t},$$

for T large enough and γ given by (3.3), where η^* is the limit measure of the sequence $\{\mu_{t,\eta}\}_{t\geq 0}$.

Corollary 3.2.4. *Suppose that $(\mu_{t,\eta})_{t\geq 0,\,\eta\in\mathcal{P}(E)}$ is exponentially stable in the sense of Definition 3.2.2, for some $t_0 > 0$ and some parameter $\gamma > 0$. Then the sequence $\left(\frac{1}{t}\int_0^t \mu_{s,\mu_0}(\mathcal{V})ds\right)_{t\geq 0}$ converges and the limit corresponds to the SCGF $\Lambda_{\mu_0}(k)$, given by Equation (2.15). Moreover, the following estimate holds*

$$\left|\frac{1}{t}\int_0^t \mu_{s,\mu_0}(\mathcal{V})ds - \Lambda_{\mu_0}(k)\right| \leq \frac{c}{t}, \tag{3.2}$$

for any $t \geq t_0$ and $c = \left(2t_0 + \frac{C\cdot\exp(-t_0\cdot\gamma)}{\gamma}\right)\cdot ||\mathcal{V}||$.

Proof. By Theorem 3.2.3, we know that the sequence $\left(\mu_{t,\mu_0}(\mathcal{V})\right)_{t\geq 0}$ converges to some limit $\nu^*(\mathcal{V}) \in \mathbb{R}$. In particular, for $t > t_0$ we have the estimate

$$\left|\mu_{t,\mu_0}(\mathcal{V}) - \mu^*(\mathcal{V})\right| \leq C\,||\mathcal{V}|| \cdot e^{-\gamma\cdot t}.$$

Thus,

$$\begin{aligned}
\left|\frac{1}{t}\int_0^t \mu_{s,\mu_0}(\mathcal{V})\,ds - \mu^*(\mathcal{V})\right| &= \left|\frac{1}{t}\int_0^{t_0} \mu_{s,\mu_0}(\mathcal{V})ds \,+\, \frac{1}{t}\int_{t_0}^t \mu_{s,\mu_0}(\mathcal{V})\,ds \,-\, \mu^*(\mathcal{V})\right| \\
&\le \left|\frac{1}{t}\int_0^{t_0} \mu_{s,\mu_0}(\mathcal{V})ds\right| \,+\, \left|\frac{1}{t}\int_{t_0}^t \big(\mu_{s,\mu_0}(\mathcal{V}) - \mu^*(\mathcal{V})\big)\,ds \,-\, \frac{t_0}{t}\mu^*(\mathcal{V})\right| \\
&\le \frac{t_0}{t}\,||\mathcal{V}|| \,+\, \frac{1}{t}\int_{t_0}^t \big|\mu_{s,\mu_0}(\mathcal{V}) - \mu^*(\mathcal{V})\big|\,ds \,+\, \frac{t_0}{t}\big|\mu^*(\mathcal{V})\big| \\
&\le \frac{t_0}{t}\,||\mathcal{V}|| \,+\, \frac{C\cdot||\mathcal{V}||}{t}\int_{t_0}^t e^{-s\cdot\gamma}\,ds \,+\, \frac{t_0}{t}\,||\mathcal{V}||.
\end{aligned}$$

since $|\mu^*(\mathcal{V})| \le ||\mathcal{V}||$. Observing that

$$\int_{t_0}^t e^{-s\cdot\gamma}\,ds \;=\; \frac{e^{-t_0\cdot\gamma} - e^{-t\gamma}}{\gamma} \;\le\; \frac{e^{-t_0\cdot\gamma}}{\gamma},$$

we obtain that the limit

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t \mu_{s,\mu_0}(\mathcal{V})ds = \nu^*(\mathcal{V})$$

exists and is finite. Recalling Equation (2.15), we also see that $\nu^*(\mathcal{V}) = \Lambda(k)$. Finally, the computations above imply the estimate (3.2).

□

We conclude the section by providing assumptions that guarantee that $(\mu_{t,\eta})_{t\ge 0,\,\eta\in\mathcal{P}(E)}$ is exponentially stable.

Assumption 3.2.5. We assume that for every $t \ge 0$ there exists a positive constant $\varepsilon_t \in (0,1)$ and a reference probability measure $\mu_t \in \mathcal{P}(E)$ such that

$$\varepsilon_t\,\mu_t(A) \;\le\; \widehat{P}_t(x,\,A) \;\le\; \varepsilon_t^{-1}\,d\mu_t(A),$$

for all $x \in E$ and $A \in \mathcal{B}(E)$, where $\widehat{P}$ is the semigroup associated to the jump generator $\widehat{\mathcal{L}}$ (2.8).

Proposition 3.2.6. *If Assumption 3.2.5 holds for some constant $\varepsilon > 0$, then for any $\tau > 0$ and $p > 1$ we have that $(\mu_{t,\eta})_{t\ge 0,\,\eta\in\mathcal{P}(E)}$ is exponentially stable in the sense of definition 3.2.2, with $t_0 = p\cdot\tau$ and*

$$\gamma = \big(1 - \frac{1}{p}\big)\cdot\frac{z_\tau\cdot\varepsilon}{\tau}, \tag{3.3}$$

with $z_\tau := \exp(-2\tau\cdot||\mathcal{V}||)$.

Proof. See [7], Theorem 3.2. Observe that Assumption 3.2.5, besides the fact that $\mathcal{V}$ is time-homogeneous and bounded, implies the assumption of Theorem 3.2 in [7], by using Proposition 4.3 in [7].

□

Lemma 3.2.7. *Assume the state space E to be finite and the transition matrix $\{\widehat{W}_{x,y}\}_{x,y\in E}$ defined in* (2.7) *to be irreducible. Then Assumption 3.2.5 holds for the semigroup $\widehat{P}$ associated to the transition matrix $\{\widehat{W}_{x,y}\}_{x,y\in E}$ and with constant $\varepsilon = \inf_{x,y,z\in E} \frac{\widehat{P}_t(x,y)}{\widehat{P}_t(z,y)} > 0$.*

Remark. Lemma 3.2.7 implies that, when the state space E is finite, the normalised Feynman-Kac measure $\{\mu_{t,\nu}\}_{t,\nu}$ associate to a irreducible transition matrix $\{\widehat{W}_{x,y}\}_{x,y\in E}$ is exponentially stable and, in particular, we can apply Corollary 3.2.4 so that the limit

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t \mu_{s,\mu_0}(\mathcal{V})ds \;=\; \mathcal{L}(k)$$

exists and is finite. The same result can be proven applying Perron-Frobenius Theorem.

3.3 Estimation of the SCGF

In this section we are interested in approximating the SCGF $\Lambda(k)$ through the particle systems introduced in Section 3.1. In particular, we are interested in evaluating the quantity

$$\mathbb{E}\Big[\Big|\Lambda_{\mu_0}(k) \;-\; \frac{1}{t}\int_0^t \mu^N_{s,\mu_0}(\mathcal{V})ds\Big|^2\Big]^{1/2}.$$

Proposition 3.3.1. *Suppose that $(\mu_{t,\eta})_{t\geq 0,\,\eta\in\mathcal{P}(E)}$ is exponentially stable in the sense of definition 3.2.2, for some $t_0 > 0$ and some parameter $\gamma > 0$. Then, for any $N \geq \exp\big(2t_0(\gamma + 4||\mathcal{V}||)\big)$ and $f \in \mathcal{C}(E)$ we have the uniform estimate*

$$\sup_{t\geq 0} \mathbb{E}\big[\,|\mu^N_{t,\mu_0}(f) - \mu_{t,\mu_0}(f)|^2\big]^{1/2} \;\leq\; \frac{8\,||f||}{N^\beta},$$

with $\beta := \gamma/2(\gamma + 4||\mathcal{V}||)$.

Proof. See [8], Theorem 3.6.

□

Remark. Recall that the requirement of exponential stability in Proposition 3.3.1 is satisfied under assumption 3.2.5, which is in general more easy to verify.

Corollary 3.3.2. *Suppose that $(\mu_{t,\eta})_{t\geq 0,\,\eta\in\mathcal{P}(E)}$ is exponentially stable in the sense of definition 3.2.2, for some $t_0 > 0$ and some parameter $\gamma > 0$. Then, for any $N \geq \exp\big(2t_0(\gamma + 4||\mathcal{V}||)\big)$ and $f \in \mathcal{C}(E)$ we have*

$$\mathbb{E}\Big[\Big|\frac{1}{t}\int_0^t \mu_{s,\mu_0}(\mathcal{V})ds \;-\; \frac{1}{t}\int_0^t \mu^N_{s,\mu_0}(\mathcal{V})ds\Big|^2\Big]^{1/2} \leq \frac{8\,||\mathcal{V}||}{N^\beta},$$

and, in particular,

$$\mathbb{E}\left[\left|\Lambda_{\mu_0}(k) - \frac{1}{t}\int_0^t \mu^N_{s,\mu_0}(\mathcal{V})ds\right|^2\right]^{1/2} \leq \frac{8\,||\mathcal{V}||}{N^\beta} + \frac{c}{t},$$

with $\beta := \gamma/2(\gamma + 4||\mathcal{V}||)$ *and* $c = \left(2t_0 + \frac{C\cdot\exp(-t_0\cdot\gamma)}{\gamma}\right)\cdot||\mathcal{V}||$.

Proof. The first part of the statement follows from Theorem 3.3.1, observing that

$$\begin{aligned}\mathbb{E}\left[\left|\frac{1}{t}\int_0^t \mu_{s,\mu_0}(\mathcal{V})ds - \frac{1}{t}\int_0^t \mu^N_{s,\mu_0}(\mathcal{V})ds\right|^2\right]^{1/2} \\ &\leq \frac{1}{t}\int_0^t \mathbb{E}\left[\left|\mu_{s,\mu_0}(\mathcal{V}) - \mu^N_{s,\mu_0}(\mathcal{V})\right|^2\right]^{1/2} ds \\ &\leq \frac{8\,||\mathcal{V}||}{N^\beta}.\end{aligned}$$

The second part of the statement follows by Corollary 3.2.4, noting that

$$\begin{aligned}\mathbb{E}\left[\left|\Lambda_{\mu_0}(k) - \frac{1}{t}\int_0^t \mu^N_{s,\mu_0}(\mathcal{V})ds\right|^2\right]^{1/2} \\ &\leq \mathbb{E}\left[\left|\Lambda_{\mu_0}(k) - \frac{1}{t}\int_0^t \mu_{s,\mu_0}(\mathcal{V})ds\right|^2\right]^{1/2} + \\ &\quad + \mathbb{E}\left[\left|\frac{1}{t}\int_0^t \mu_{s,\mu_0}(\mathcal{V})ds - \frac{1}{t}\int_0^t \mu^N_{s,\mu_0}(\mathcal{V})ds\right|^2\right]^{1/2}.\end{aligned}$$

□

Observe that, in particular, we are able to estimate the approximation error

$$\left|\mathbb{E}\left[\Lambda_{\mu_0}(k) - \frac{1}{t}\int_0^t \mu^N_{s,\mu_0}(\mathcal{V})ds\right]\right| \leq \frac{8\,||\mathcal{V}||}{N^\beta} + \frac{c}{t},$$

with $\beta := \gamma/2(\gamma + 4||\mathcal{V}||)$ and $c = \left(2t_0 + \frac{C\cdot\exp(-t_0\cdot\gamma)}{\gamma}\right)\cdot||\mathcal{V}||$. From this, we can see that the estimated quantity

$$\int_0^t \mu^N_{s,\mu_0}(\mathcal{V})ds$$

converges to the SCGF as $N, t \to \infty$. Moreover, the order of convergence with respect to N is $1/N^\beta$ and $\beta < 1/2$ is increasing in the variable γ.

We conclude the Chapter by illustrating how to apply the convergence results, through a toy example.

Recall Example 1. We have consider a continuous-time random walk on the state space $S_M = \mathbb{Z}^d \cap [-M, M]^d$, with $M \in \mathbb{N}$, with the boundary assumptions $M+1 = -M$ and $-M-1 = M$. On S_M, we consider a random walk defined by the overall transition rates $W(x, y) = \frac{\lambda}{2^d}$, if $||x-y|| = 1$ and $W(x, t) = 0$ otherwise. We have also introduced the function $g(x,y) := \mathbb{1}_B(x)$, for all $x, y \in S$ and found that

$$\widehat{W}(x,y) = \begin{cases} \frac{\lambda}{2^d} & x \notin B,\ ||y-x|| = 1 \\ \frac{\lambda\, e^k}{2^d} & x \in B,\ ||y-x|| = 1 \\ 0 & \text{otherwise.} \end{cases}$$

It is easy to see $\widehat{W}$ is irreducible, so that Assumption 3.2.5 is satisfied for some parameter $\varepsilon > 0$. We can thus apply Corollary 3.2.4 and obtain that the limit

$$\lim_{t\to\infty} \int_0^t \mu_{s,\mu_0}(\mathcal{V}) ds$$

exists and it represents the SCGF $\Lambda_{\mu_0}(k)$, as stated in 2.2.3. In particular, we are able to estimate the L^2-error given by applying the cloning algorithm described in Section 3.1. Indeed, applying Corollary 3.3.2, we have that, for any $\tau > 0$ and $p > 1$,

$$\mathbb{E}\left[\left|\Lambda_{\mu_0}(k) - \frac{1}{t}\int_0^t \mu^N_{s,\mu_0}(\mathcal{V}) ds\right|^2\right]^{1/2} \leq \frac{8\,||\mathcal{V}||}{N^\beta} + \frac{c}{t},$$

with $\beta := \gamma/2(\gamma + 4||\mathcal{V}||)$ and $c = \left(2t_0 + \frac{C\cdot\exp(-t_0\cdot\gamma)}{\gamma}\right)\cdot||\mathcal{V}||$, where $t_0 = p\cdot\tau$ and

$$\gamma = \left(1 - \frac{1}{p}\right)\cdot\frac{\varepsilon\cdot\exp(-2\tau\cdot||\mathcal{V}||)}{\tau}.$$

Bibliography

[1] R. Chetrite and H. Touchette. Nonequilibrium Markov processes conditioned on large deviations. *Annales Henri Poincaré*, 16:2005–2057, 2015.

[2] P. Del Moral. *Mean field simulation for Monte Carlo integration.* CRC Press, 2013.

[3] P. Del Moral and N. G. Hadjiconstantinou. An introduction to probabilistic methods with applications. *ESAIM: Mathematical Modelling and Numerical Analysis*, 44(5):805–829, 2010.

[4] P. Del Moral and L. Miclo. Asymptotic stability of non linear semigroup of Feynman-Kac type. *Publications du Laboratoire de statistique et probabilités*, 1999.

[5] P. Del Moral and L. Miclo. Branching and interacting particle systems approximations of Feynman-Kac formulae with applications to non-linear filtering. In *Seminaire de probabilites XXXIV*, pages 1–145. Springer, 2000.

[6] P. Del Moral and L. Miclo. A Moran particle system approximation of Feynman-Kac formulae. *Stochastic processes and their applications*, 86(2):193–216, 2000.

[7] P. Del Moral and L. Miclo. On the stability of nonlinear Feynman-Kac semigroups. In *ANNALES-FACULTE DES SCIENCES TOULOUSE MATHEMATIQUES*, volume 11, pages 135–175. Université Paul Sabatier, 2002.

[8] P. Del Moral and L. Miclo. Particle approximations of Lyapunov exponents connected to Schrödinger operators and Feynman-Kac semigroups. *ESAIM: Probability and Statistics*, 7:171–208, 2003.

[9] A. Dembo and O. Zeitouni. *Large deviations techniques and applications*, volume 38. Springer Science & Business Media, 2009.

[10] C. Giardina, J. Kurchan, V. Lecomte, and J. Tailleur. Simulating rare events in dynamical processes. *Journal of statistical physics*, 145(4):787–811, 2011.

[11] C. Giardina, J. Kurchan, and L. Peliti. Direct evaluation of large-deviation functions. *Physical review letters*, 96(12):120603, 2006.

[12] S. Grosskinsky. Large-scale dynamics of stochastic particle systems. Lecture Notes, 2016.

[13] E. G. Hidalgo, T. Nemoto, and V. Lecomte. Finite-time and-size scalings in the evaluation of large deviation functions-part i: Analytical study using a birth-death process. *arXiv preprint arXiv:1607.04752*, 2016.

[14] E. G. Hidalgo, T. Nemoto, and V. Lecomte. Finite-time and-size scalings in the evaluation of large deviation functions-part ii: Numerical approach in continuous time. *arXiv preprint arXiv:1607.08804*, 2016.

[15] M. Kac. On distributions of certain Wiener functionals. *Transactions of the American Mathematical Society*, 65(1):1–13, 1949.

[16] J. L. Lebowitz and H. Spohn. A Gallavotti-Cohen-type symmetry in the large deviation functional for stochastic dynamics. *Journal of Statistical Physics*, 95(1):333–365, 1999.

[17] T. M. Liggett. *Continuous time Markov processes: an introduction*, volume 113. American Mathematical Soc., 2010.

[18] H. Touchette. *Equivalence and nonequivalence of the microcanonical and canonical ensembles: a large deviations study*. PhD thesis, McGill University, 2003.

[19] H. Touchette. The large deviation approach to statistical mechanics. *Physics Reports*, 478(1):1–69, 2009.