Feynman-Kac Models for Large Deviation Conditioning Problem

by

Letizia Angeli

Thesis

Submitted for the degree of

Master of Science

Mathematics Institute

The University of Warwick

August 2017

Contents

Abstract

Rare trajectories of stochastic systems are important to understand but difficult to sample directly. In this work, we introduce a new importance sampling method based on cloning for evaluating directly large deviation functions (LDFs) associated to the distribution of additive observables. These large deviation functions are given in terms of the typical properties of a modified dynamics which rises from the Feynman-Kac theory and, since the LDFs no longer involve rare events, can be evaluated efficiently. The method we propose also allows one to study analitically the order of convergence of the estimator to the correct value of the quantity of interest, adapting already established results for Feynman-Kac models.

Introduction

The study of atypical, rare trajectories of dynamical systems arises in many physical applications such as 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, or for sufficiently many times to observe a large deviation event.

A widely used class of numerical procedures for generating rare events efficiently are importance sampling methods based on *cloning*. 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. [1; 2] 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 [3; 4].

This cloning method for estimating the SCGF has been used widely in 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. [5; 6] proposed a slightly different version of the cloning algorithm, for which they studied the speed of convergence to the SCGF not on a fully rigorous level.

In this work, we propose a different cloning algorithm based on Feynman-Kac models [7]. This new approach enables us to study analytically the convergence of the algorithm to the scaled cumulant generating function, 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 [8] in the context of Lyapunov exponents connected to Schrödinger operators.

Feynman-Kac models were originally introduced in the 1940s [9] 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 many-particle systems which converge to the associated Feynman-Kac model as the size of the system 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 present the general theory of Feynman-Kac models used throughout this work, providing three different interpretations of these models and illustrating two structural stability properties essential to the construction of the cloning algorithm.

In Chapter 2 we consider the problem of studying atypical trajectories in the long-time limit of a discrete-time Markov process, and show how to apply the theory of Feynman-Kac models for estimating the associated SCGF. We start the chapter presenting general results of statistical mechanics to motivate the importance of the SCGF in the study of rare events. In particular we will see that, under certain assumptions, the conditioned Markov process can be represented by a conditioning-free model obtained by replacing this conditioning with an exponential factor involving a Lagrange parameter dual to the constraint [3]. From the point of view of statistical mechanics [4], this new process can be seen as a nonequilibrium generalisation of the *canonical ensemble* associated with the Markov process X_n, whereas the conditioned process corresponds to a nonequilibrium generalisation of the *microcanonical ensemble*. In this context, the SCGF can be seen as a generalisation of the free energy. In Section 2.2, we represent the canonical path ensemble and the associated SCGF via time-homogeneous Feynman-Kac models defined on the state space [7]. We conclude the chapter constructing the cloning algorithm which rises from the N-particle system interpretation of Feynman-Kac models.

Finally, in Chapter 3, we present new results on the order of convergence of the considered cloning algorithm to the correct value of the SCGF, under different conditions.

Chapter 1

Feynman-Kac Theory

In this chapter, we present the general theory of Feynman-Kac models, in particular providing three main different descriptions of these models, i.e.

- path-space probability distributions,
- time-marginal flows (seen as solutions of nonlinear equations),
- limits of N-particle system models.

In Chapter 2 we will see how to apply the Feynman-Kac theory presented here to the study of the large-deviation conditioning problem.

The main reference for this chapter is [7].

1.1 Basic Notation and definitions

In this Section, we provide the basic notation we will use in the presentation of the Feynman-Kac Theory.

We denote respectively by $\mathcal{M}(E)$ and $\mathcal{P}(E)$ the set of bounded and signed measures and the set of probability measures on a given measurable space $(E, \mathcal{E})$. Also, $\mathcal{B}_b(E)$ denotes the set of bounded measurable functions on $(E, \mathcal{E})$.

Definition 1.1.1. Let $(E_0, \mathcal{E}_0)$, $(E_1, \mathcal{E}_1)$ be measurable spaces. A map $M : E_0 \times \mathcal{E}_1 \to [0, +\infty]$ is called a *transition kernel* from $(E_0, \mathcal{E}_0)$ to $(E_1, \mathcal{E}_1)$ if:

1. $x_0 \mapsto M(x_0, A_1)$ is $\mathcal{E}_0$-measurable for any $A_1 \in \mathcal{E}_1$;
2. $M(x_0, \cdot) \in \mathcal{M}(E_1)$ for any $x_0 \in E_0$.

If $M(x_0, \cdot) \in \mathcal{P}(E_1)$ for any $x_0 \in E_0$, we say that M is a *Markov kernel*.

Any transition kernel $M(x_0, dx_1)$ from a measure space $(E_0, \mathcal{E}_0)$ to another measure space $(E_1, \mathcal{E}_1)$ generates two operators, one acting on bounded $\mathcal{E}_1$-measurable functions $f_1 \in \mathcal{B}_b(E_1)$ and taking values in $\mathcal{B}_b(E_0)$

$$(Mf_1)(x_0) = \int_{E_1} M(x_0, dx_1) f_1(x_1), \quad \forall (x_0, f_1) \in E_0 \times \mathcal{B}_b(E_1),$$

and the other one acting on measures $\mu_0 \in \mathcal{M}(E_0)$ and taking values in $\mathcal{M}(E_1)$

$$(\mu_0 M)(A_1) = \int_{E_0} \mu_0(dx_0) M(x_0, A_1), \quad \forall (\mu_0, A_1) \in \mathcal{P}(E_0) \times \mathcal{E}_1. \tag{1.1}$$

We finally define the composite operator of two transition kernels M_1 and M_2, respectively from $(E_0, \mathcal{E}_0)$ to $(E_1, \mathcal{E}_1)$ and from $(E_1, \mathcal{E}_1)$ to $(E_2, \mathcal{E}_2)$, by

$$(M_1 M_2)(x_0, dx_2) = \int_{E_1} M_1(x_0, dx_1) M_2(x_1, dx_2).$$

1.2 The Feynman-Kac Models

The main purpose of this section is to present the Feynman-Kac models, providing first the traditional path-space definition and, then, the associated time-marginal flows. We will show in Section 1.3 that these two different representations of Feynman-Kac models actually possess the same algebraic structure. The flow interpretation of the Feynman-Kac models will be particularly helpful in the construction of the corresponding N particle models.

Let $(E_n, \mathcal{E}_n)$, $n \in \mathbb{N}$, be a collection of measurable spaces. The Feynman-Kac models are built with two main ingredients: a sequence of bounded and $\mathcal{E}_n$-measurable potential functions $G_n : E_n \to [0, \infty)$ and a (non-homogeneous) Markov chain

$$\left(\Omega = \prod_{n \geq 0} E_n,\, \mathcal{F} = (\mathcal{F}_n)_{n \in \mathbb{N}},\, X = (X_n)_{n \in \mathbb{N}},\, \mathbb{P}_{\nu_0} \right)$$

associated to a collection of Markov kernels M_n from E_{n-1} to E_n with initial distribution $\nu_0 \in \mathcal{P}(E_0)$ and where $\mathcal{F}$ is a filtration with respect to which X is adapted.

We use the notation $\mathbb{E}_{\nu_0}$ for the expectations with respect to $\mathbb{P}_{\nu_0}$ and $\mathbb{P}_{\nu_0,n}$ for the distribution on $E_{[0,n]} := \prod_{p=0}^{n} E_p$ given by

$$\mathbb{P}_{\nu_0,n}(d(x_0, \ldots, x_n)) = \nu_0(x_0) M(x_0, dx_1) \ldots M(x_{n-1}, dx_n).$$

Thus, for any $F_n \in \mathcal{B}_b(E_{[0,n]})$ we have

$$\mathbb{E}_{\nu_0}\big[F_n(X_0,\dots,X_n)\big] = \int_{E_{[0,n]}} F_n(x_0,\dots,x_n)\mathbb{P}_{\nu_0,n}(d(x_0,\dots,x_n)).$$

Assumption 1.2.1. We limit ourselves to consider strictly positive potential functions.

The Feynman-Kac models associated with the sequence of pairs $(G_n,\, M_n)_{n\in\mathbb{N}}$ are traditionally defined as follows.

Definition 1.2.2. The *Feynman-Kac prediction* and *updated path models* associated with the sequence of pairs $(G_n,\, M_n)_{n\in\mathbb{N}}$, and with initial distribution ν_0, are the sequence of path measures defined respectively by

$$\mathbb{Q}_{\nu_0,\,n}(d\omega) := \frac{1}{Z_n}\left[\prod_{p=0}^{n-1} G_p(\omega_p)\right]\mathbb{P}_{\nu_0,\,n}(d\omega),$$
$$\widehat{\mathbb{Q}}_{\nu_0,\,n}(d\omega) := \frac{1}{\widehat{Z}_n}\left[\prod_{p=0}^{n} G_p(\omega_p)\right]\mathbb{P}_{\nu_0,\,n}(d\omega),$$

for any $\omega = (\omega_0,\dots,\omega_n)$, $n \in \mathbb{N}$, with $\omega_p \in E_p$, $p = 0,\dots,n$, and where Z_n and $\widehat{Z}_n$ are the normalising constants

$$Z_n := \mathbb{E}_{\nu_0}\left[\prod_{p=0}^{n-1} G_p(\omega_p)\right] \quad \text{and} \quad \widehat{Z}_n := Z_{n+1} = \mathbb{E}_{\nu_0}\left[\prod_{p=0}^{n} G_p(\omega_p)\right].$$

Sometimes, we will make use of a weaker definition of the measures $\mathbb{Q}_{\nu_0,n}$ and $\widehat{\mathbb{Q}}_{\nu_0,n}$, given for any test function $f_n \in \mathcal{B}_b(E_{[0,n})$ by the formulae

$$\mathbb{Q}_{\nu_0,n}(f_n) = \frac{1}{Z_n}\mathbb{E}_{\nu_0}\left[f_n(X_0,\dots,X_n)\prod_{p=0}^{n-1} G_p(X_p)\right],$$
$$\widehat{\mathbb{Q}}_{\nu_0,n}(f_n) = \frac{1}{\widehat{Z}_n}\mathbb{E}_{\nu_0}\left[f_n(X_0,\dots,X_n)\prod_{p=0}^{n} G_p(X_p)\right].$$

The main benefit in using the Feynman-Kac interpretation of the canonical ensemble is that these models can be seen as the limit of N-particle systems, as N tends to infinity. To show that, it is convenient to introduce, besides the traditional path F-K distributions, the flow of the terminal time marginals.

Definition 1.2.3. The *unnormalised prediction* γ_n and *updated Feynman-Kac model* $\widehat{\gamma}_n$ associated with the sequence of pairs $(G_n,\, M_n)_{n\in\mathbb{N}}$ are the sequences of bounded nonnegative measures on E_n defined weakly for any $f_n \in \mathcal{B}_b(E_n)$ by

$$\gamma_n(f_n) := \mathbb{E}_{\nu_0}\left[f_n(X_n)\prod_{p=0}^{n-1} G_p(X_p)\right],$$
$$\widehat{\gamma}_n(f_n) := \mathbb{E}_{\nu_0}\left[f_n(X_n)\prod_{p=0}^{n} G_p(X_p)\right].$$

Definition 1.2.4. The *normalised prediction* η_n and *updated Feynman-Kac model* $\widehat{\eta}_n$ associated with the sequence of pairs $(G_n, M_n)_{n\in\mathbb{N}}$ are sequences of probability measures on E_n defined for any $f_n \in \mathcal{B}_b(E_n)$ by

$$\eta_n(f_n) := \frac{\gamma_n(f_n)}{\gamma_n(1)}, \qquad \widehat{\eta}_n(f_n) := \frac{\widehat{\gamma}_n(f_n)}{\widehat{\gamma}_n(1)}.$$

Remark. Observe that $\gamma_n(1) = Z_n$ and $\widehat{\gamma}_n(1) = \widehat{Z}_n$. Moreover, for any bounded function $f_n \in \mathcal{B}_b(E_{[0,n]})$ which depends only on the terminal point of the path, i.e. $f_n(\omega_0, \ldots, \omega_n) = f(\omega_n)$, $f \in \mathcal{B}_b(E_n)$, we have

$$\mathbb{Q}_{\nu_0,n}(f_n) \;=\; \eta_n(f) \quad \text{and} \quad \widehat{\mathbb{Q}}_{\nu_0,n}(f_n) \;=\; \widehat{\eta}_n(f). \tag{1.2}$$

From the definitions, we can see that the unnormalised flow can be computed in terms of the normalised distributions. More precisely, we have that

$$\gamma_n(f_n) = \eta_n(f_n)\prod_{p=0}^{n-1}\eta_p(G_p). \tag{1.3}$$

In particular, we have $\gamma_0 = \eta_0 = \nu_0$, and also $\widehat{\gamma}_0 = \widehat{\eta}_0 = \nu_0$.

We conclude the section showing that the flow of the normalised time marginal can be interpreted as the solution of nonlinear equations.

Definition 1.2.5. The *Boltzmann-Gibbs transformation* associated with a potential function G_n on $(E_n, \mathcal{E}_n)$ is the mapping from $\mathcal{P}(E_n)$ into itself defined for any $\eta \in \mathcal{P}(E_n)$ by the Boltzmann-Gibbs measure

$$\Psi_n(\eta)(dx_n) \;:=\; \frac{G_n(x_n)\eta(dx_n)}{\eta(G_n)}$$

Proposition 1.2.6. *The normalised prediction and updated Feynman-Kac distributions η_n and $\widehat{\eta}_n$ associated with the pairs $(G_n,\, M_n)_{n\in\mathbb{N}}$ satisfy the nonlinear recursive* *equations*

$$\eta_n = \Phi_n(\eta_{n-1}), \tag{1.4}$$

$$\widehat{\eta}_n = \widehat{\Phi}_n(\widehat{\eta}_{n-1}), \tag{1.5}$$

with the mappings Φ_n *and* $\widehat{\Phi}_n$ *from* $\mathcal{P}(E_{n-1})$ *into* $\mathcal{P}(E_n)$ *defined for any* $\eta \in \mathcal{P}(E_{n-1})$ *by*

$$\Phi_n(\eta) := \Psi_{n-1}(\eta)M_n \quad \text{and} \quad \widehat{\Phi}_n(\eta) = \widehat{\Psi}_n(\eta M_n). \tag{1.6}$$

Proof. First, note that the Boltzmann-Gibbs transformation is well-defined for G_n strictly positive and bounded.

Recall that we can write $\widehat{\gamma}_n(f_n) = \gamma_n(f_n G_n)$. By the definition of the normalised Feynman-Kac models, we have

$$\widehat{\eta}_n(f_n) \;=\; \frac{\gamma_n(f_n G_n)/\gamma_n(1)}{\gamma_n(G_n)/\gamma_n(1)} \;=\; \frac{\eta_n(f_n G_n)}{\eta_n(G_n)}$$

Thus, we can see that

$$\widehat{\eta}_n = \Psi_n(\eta_n). \tag{1.7}$$

Moreover,

$$\begin{aligned}\gamma_n(f_n) \;&=\; \mathbb{E}_{\nu_0}\left[f_n(X_n)\prod_{p=0}^{n-1} G_p(X_p)\right] \\ &=\; \mathbb{E}_{\nu_0}\left[M_n(f_n)(X_{n-1})\prod_{p=0}^{n-1} G_p(X_p)\right] = \widehat{\gamma}_{n-1}(M_n(f_n)).\end{aligned}$$

From this we find that

$$\eta_n(f_n) = \frac{\widehat{\gamma}_{n-1}(M_n(f_n))}{\widehat{\gamma}_{n-1}(1)} = \widehat{\eta}_{n-1}M_n(f_n). \tag{1.8}$$

Combining (1.7) with (1.8), we obtain

$$\eta_n(f_n) \underset{(1.8)}{=} \widehat{\gamma}_{n-1}M_n(f_n) \underset{(1.7)}{=} \Psi_{n-1}(\eta_{n-1})M_n(f_n) \;=\; \Phi_n(\eta_{n-1})(f_n),$$

and

$$\widehat{\eta}_n(f_n) \underset{(1.7)}{=} \Psi_n(\eta_n)(f_n) \underset{(1.8)}{=} \Psi_n(\widehat{\eta}_{n-1}M_n)(f_n) \;=\; \widehat{\Phi}_n(\widehat{\eta}_{n-1})(f_n).$$

This concludes the proof. □

Remark. Note that Φ_n and Ψ_n, are well-defined on the whole set of distributions

$\mathcal{P}(E_n)$, since $\eta(G_n) > 0$. It is also important to stress that (1.6) is not the only possible decomposition, for instance a more general class of solutions $(\Phi_n, \widehat{\Phi}_n)$ for the nonlinear recursive equations (1.4)-(1.5) is given by the McKean models (see [7], Section 2.5.3), but also other interpretations are possible.

Decomposition (1.6) shows in particular that the Feynman-Kac flow is a well-defined two-step updating/prediction model

$$\eta_n \xrightarrow{\Psi_n} \widehat{\eta}_n := \Psi_n(\eta_n) \xrightarrow{M_{n+1}} \eta_{n+1} := \widehat{\eta}_n M_{n+1} \tag{1.9}$$

with $\eta_n, \widehat{\eta}_n \in \mathcal{P}(E_n)$ and $\eta_{n+1} \in \mathcal{P}(E_{n+1})$.

This consideration is the basic idea behind the construction of the corresponding N-particle model, as we will see in Section 1.4.

1.3 Structural Stability Properties

Before providing the N-particle interpretation of the Feynman-Kac models, it is important to underline the interplay between the Feynman-Kac path models and the time-marginal flows, as also the connection between prediction and updated models. These two basic structural stability properties of the Feynman-Kac models will allow us to extend the results for the prediction marginals to the corresponding prediction and updated path measures.

In what follows, in order to avoid confusion, we make explicit the associated sequence of potential/transition pairs $(G_n, M_n)_{n\in\mathbb{N}}$ in the notation of the Feynman-Kac distributions and flows considered. For instance, we will write $\mathbb{Q}_{\nu_0,n}^{(G_n, M_n)}$ instead of simply $\mathbb{Q}_{\nu_0,n}$.

Consider a Feynman-Kac path model $\mathbb{Q}_{\nu_0,n}^{(G_n, M_n)}$ and the corresponding time-marginal flow $(\eta_n^{(G_n, M_n)}, \widehat{\eta}_n^{(G_n, M_n)})$ associated to the sequence of pairs $(G_n, M_n)_{n\in\mathbb{N}}$ and with initial distribution ν_0. We are interested in proving that there exist sequences of potential/kernel pairs, respectively $(\mathcal{G}_n, \mathcal{M}_n)_{n\in\mathbb{N}}$ and $(\widehat{G}_n, \widehat{M}_n)_{n\in\mathbb{N}}$, which can be construct explicitly and such that

$$\eta_n^{(\mathcal{G}_n, \mathcal{M}_n)} = \mathbb{Q}_{\nu_0,n}^{(G_n, M_n)} \quad \text{and} \quad \eta_n^{(\widehat{G}_n, \widehat{M}_n)} = \widehat{\eta}_n^{(G_n, M_n)}.$$

We call these two equalities structural stability properties, since they means that

(1) Feynman-Kac path models have the same structure and can be interpreted as Feynman-Kac time-marginal flows;

(2) updated Feynman-Kac flows have the same structure and can be interpreted as prediction Feynmnan-Kac flows.

These properties will be fundamental in Chapter 2, for interpreting the canonical path ensemble and the SCGF through Feynman-Kac models associated to convenient pairs of potential/kernels.

Proposition 1.3.1 (**1**st Structural Stability Property)**.** *Let* $\mathbb{Q}^{(G_n,\, M_n)}_{\nu_0,n}$ *be a Feynman-Kac path model associated to the sequence of potential/kernel pairs* $(G_n,\, M_n)_{n\in\mathbb{N}}$ *defined on* $\Omega := \prod_{n\in\mathbb{N}} E_n$ *and with initial distribution* ν_0*.*

Consider the sequence of potential/kernel pairs $(\mathcal{G}_n,\, \mathcal{M}_n)_{n\in\mathbb{N}}$ *where*

$$\mathcal{G}_n(\omega^{(n)}) := G_n(\omega_n)$$

is defined for any $\omega^{(n)} := (\omega_0, \ldots, \omega_n) \in \prod_{p=0}^n E_p$ *and* $\mathcal{M}_n$*,* $n \in \mathbb{N}$*, are the Markov transitions associated to the non-homogeneous path process* $\mathcal{X}_n := X_{0:n} \in \prod_{p=0}^n E_p$ *associated to the Markov chain* X_n *and with initial distribution* ν_0*.*

Then, we have that

$$\eta_n^{(\mathcal{G}_n,\, \mathcal{M}_n)} \;=\; \mathbb{Q}^{(G_n,\, M_n)}_{\nu_0,n}\,,$$

where $\eta_n^{(\mathcal{G}_n,\, \mathcal{M}_n)}$ *is the prediction Feynman-Kac flow associated to the sequence of pairs* $(\mathcal{G}_n,\, \mathcal{M}_n)_{n\in\mathbb{N}}$*.*

Proof. By the definition of normalised prediction flow 1.2.4, $\eta_n^{(\mathcal{G}_n,\, \mathcal{M}_n)}$ is given by

$$\eta_n^{(\mathcal{G}_n,\, \mathcal{M}_n)}(d\omega^{(n)}) = \frac{1}{\gamma_n^{(\mathcal{G}_n,\, \mathcal{M}_n)}(1)} \prod_{p=0}^{n-1} \mathcal{G}_p(\omega^{(p)}) \,\cdot\, \mathbb{P}_{\nu_0,n}(d\omega^{(n)}),$$

where

$$\gamma_n^{(\mathcal{G}_n,\, \mathcal{M}_n)}(f_n) = \mathbb{E}_{\nu_0}\Big[f_n(\mathcal{X}_n) \prod_{p=0}^{n-1} \mathcal{G}_p(\mathcal{X}_p)\Big].$$

Therefore, the n-time marginal distribution $\eta_n^{(\mathcal{G}_n,\, \mathcal{M}_n)}$ associated with the pairs $(\mathcal{G}_n,\, \mathcal{M}_n)_{n\in\mathbb{N}}$ coincides with the Feynman-Kac path measure $\mathbb{Q}^{(G_n,\, M_n)}_{\nu_0,n}$ associated with the pairs $(G_n, M_n)_{n\in\mathbb{N}}$,

$$\eta_n^{(\mathcal{G}_n,\, \mathcal{M}_n)}(d(\omega_0, \ldots, \omega_n)) = \mathbb{Q}^{(G_n,\, M_n)}_{\nu_0,n}(d(\omega_0, \ldots, \omega_n)).$$

□

Corollary 1.3.2. *Let* $\eta_n^{(\mathcal{G}_n,\mathcal{M}_n)}$ *be the prediction Feynman-Kac flow associated to the pairs* $(\mathcal{G}_n, \mathcal{M}_n)_{n\in\mathbb{N}}$*, where* $\mathcal{M}_n$*,* $n \in \mathbb{N}$*, are the Markov transitions of the path process associated to a Markov chain with transitions* M_n *from* E_{n-1} *to* E_n *and initial distribution* ν_0 *and where* $\mathcal{G}_n$*,* $n \in \mathbb{N}$*, depend only on the terminal point of the path, i.e.* $\mathcal{G}_n(\omega^{(n)}) = G_n(\omega_n)$*, for some bounded function* $G_n \in \mathcal{B}_b(E_n)$*, where* $\omega^{(n)} := (\omega_0, \ldots, \omega_n) \in \prod_{p=0}^n E_p$*.*

Then, we have that

$$\eta_n^{(\mathcal{G}_n,\mathcal{M}_n)} = \mathbb{Q}_{\nu_0,n}^{(G_n,\, M_n)},$$

where $\mathbb{Q}_{\nu_0,n}^{(G_n,\, M_n)}$ *is the Feynman-Kac path distribution associated to the sequence of pairs* $(G_n,\, M_n)_{n\in\mathbb{N}}$*.*

The second important structural stability property of Feynman-Kac models is given by the connection of the updated measures with the prediction ones. This will allow us to transfer the results for prediction models to the updated flows.

Proposition 1.3.3 ($\mathbf{2}^{\text{nd}}$ Structural Stability Property)**.** *Let* $\widehat{\eta}_n^{(G_n,\, M_n)}$ *be the updated Feynman-Kac flow associated to the sequence of potential/kernel pairs* $(G_n,\, M_n)_{n\in\mathbb{N}}$ *defined on* $\Omega := \prod_{n\in\mathbb{N}} E_n$ *and with initial distribution* ν_0*.*

Consider the sequence of potential/kernel pairs $(\widehat{G}_n,\, \widehat{M}_n)_{n\in\mathbb{N}}$ *defined by*

$$\widehat{G}_n(x_n) := \int_{E_{n+1}} G_{n+1}(x_{n+1}) M_{n+1}(x_n,\, dx_{n+1}), \tag{1.10}$$

$$\widehat{M}_n(x_{n-1}, dx_n) := \frac{M_n(x_{n-1}, dx_n)\, G_n(x_n)}{\widehat{G}_{n-1}(x_{n-1})}, \tag{1.11}$$

for any $x_{n-1} \in E_{n-1}$ *and* $x_n \in E_n$*,* $n \in \mathbb{N}$*.*

Then, we have

$$\eta_n^{(\widehat{G}_n,\, \widehat{M}_n)} = \widehat{\eta}_n^{(G_n,\, M_n)},$$

where $\eta_n^{(\widehat{G}_n,\, \widehat{M}_n)}$ *is the prediction Feynman-Kac flow associated to the pairs* $(\widehat{G}_n,\, \widehat{M}_n)_{n\in\mathbb{N}}$ *and with initial distribution given by* $\widehat{\nu}_0(dx_0) := \frac{G_0(x_0)}{\nu_0(G_0)}\nu_0(dx_0)$*, for all* $x_0 \in E_0$*.*

Proof. First observe that $\widehat{M}_n$ is a well defined Markov kernel from E_{n-1} to E_n, since $\widehat{G}_n(x) > 0$.

By definition,

$$\widehat{\eta}_n^{(G_n,\, M_n)}(f_n) = \frac{\mathbb{E}_{\nu_0}\big[f_n(X_n) \prod_{p=0}^n G_p(X_p)\big]}{\mathbb{E}_{\nu_0}\big[\prod_{p=0}^n G_p(X_p)\big]},$$

for any $f_n \in \mathcal{B}_b(E_n)$. The numerator can be expanded as follows.

$$\mathbb{E}_{\nu_0}\left[f_n(X_n)\prod_{p=0}^{n} G_p(X_p)\right] =$$

$$\begin{aligned}
&= \int_{E_{[0,n]}} f_n(x_n) \prod_{p=1}^{n} M_p(x_{p-1}, dx_p) G_p(x_p) \cdot G_0(x_0)\, \nu_0(dx_0) \\
&= \int_{E_{[0,n]}} f_n(x_n) \prod_{p=1}^{n} \widehat{M}_p(x_{p-1}, dx_p) \widehat{G}_{p-1}(x_{p-1}) \cdot \nu_0(G_0)\, \widehat{\nu}_0(dx_0) \\
&= \nu_0(G_0) \int_{E_{[0,n]}} f_n(x_n) \prod_{p=0}^{n-1} \widehat{M}_{p+1}(x_p, dx_{p+1}) \widehat{G}_p(x_p)\, \widehat{\nu}_0(dx_0) \\
&= \nu_0(G_0) \mathbb{E}_{\widehat{\nu}_0}\left[f_n(\widehat{X}_n) \prod_{p=0}^{n-1} \widehat{G}_p(\widehat{X}_p)\right],
\end{aligned}$$

where $\widehat{X}_n$ is the Markov chain associated to the transition kernels $\widehat{M}_n$.

Analogously, we have that

$$\mathbb{E}_{\nu_0}\left[\prod_{p=0}^{n} G_p(X_p)\right] = \nu_0(G_0) \mathbb{E}_{\widehat{\nu}_0}\left[\prod_{p=0}^{n-1} \widehat{G}_p(\widehat{X}_p)\right].$$

Thus,

$$\widehat{\eta}_n^{(G_n,\, M_n)}(f_n) = \eta_n^{(\widehat{G}_n,\, \widehat{M}_n)}(f_n).$$

□

We conclude the section explaining briefly how these two Structural Stability Properties will be used in practice in Chapter 2 to characterise the canonical path ensemble:

- First, we will see that the canonical path ensemble can be interpreted as the updated Feynman-Kac flow $\widehat{\eta}_n^{(\mathcal{G}_n, \mathcal{M}_n)}$ associated to a sequence of pairs $(\mathcal{G}_n, \mathcal{M}_n)_{n \in \mathbb{N}}$ and with initial distribution μ_0, where $\mathcal{M}_n$, $n \in \mathbb{N}$, represent the transitions of a Markov path process (see Lemma 2.2.1);

- Using the second Structural Stability Property (Proposition 1.3.3), we can interpret the updated time marginal $\widehat{\eta}_n^{(\mathcal{G}_n, \mathcal{M}_n)}$ as the prediction time marginal associated to $(\widehat{\mathcal{G}}_n,\, \widehat{\mathcal{M}}_n)_{n \in \mathbb{N}}$ with initial distribution $\widehat{\mu}_0(dx_0) := \frac{\mathcal{G}_0(x_0)}{\mu_0(\mathcal{G}_0)} \mu_0(dx_0)$, $x_0 \in E_0$, where the pairs $(\widehat{\mathcal{G}}_n,\, \widehat{\mathcal{M}}_n)$ are related to $(\mathcal{G}_n,\, \mathcal{M}_n)$ through the equations (2.8)-(2.9) (see Lemma 2.2.2);

- We will see that $\widehat{\mathcal{M}}_n$, $n \in \mathbb{N}$, are the transitions of a Markov path process associated to a Markov chain with transitions $\widehat{M}_n$, $n \in \mathbb{N}$. Moreover, the potentials $\widehat{\mathcal{G}}_n$ depend only on the terminal state of the paths, i.e. $\widehat{\mathcal{G}}_n(\omega_0, \ldots, \omega_n) = \widehat{G}_n(\omega_n)$, for some bounded function $\widehat{G}_n \in \mathcal{B}_b(E_n)$. Using the first Structural Stability Property (Corollary 1.3.2), we can interpret the prediction time marginal $\eta_n^{(\widehat{\mathcal{G}}_n, \widehat{\mathcal{M}}_n)}$ with initial distribution $\widehat{\mu}_0$ as the Feynman-Kac path distribution $\mathbb{Q}_{\widehat{\mu}_0, n}^{(\widehat{G}_n, \widehat{M}_n)}$ associated to the pairs $(\widehat{G}_n, \widehat{M}_n)_{n\in\mathbb{N}}$ and with initial distribution $\widehat{\mu}_0$ (see Proposition 2.2.3).

These results will be explained in more detail in Section 2.2 and they will be applied also for characterising the scaled cumulant generating function.

1.4 The N-particle Interpretation

The main purpose of this section is to construct the N-particle system corresponding to the Feynman-Kac model $(\gamma_n, \eta_n)_{n\in\mathbb{N}}$ associated to the pairs potential/kernel $(G_n,\, M_n)_{n\in\mathbb{N}}$ on state spaces E_n and with initial distribution ν_0.

For the moment we limit ourselves to providing the heuristic construction of these particle systems, while rigorous convergence results to η_n will be discussed in Section 3.1.

Definition 1.4.1 (Feynman-Kac particle model)**.** Let $\Phi_n : \mathcal{P}(E_{n-1}) \to \mathcal{P}(E_n)$ be the collection of mappings defining the Feynman-Kac model associated to the sequence of pairs $(G_n,\, M_n)_{n\in\mathbb{N}}$ and let $\nu_0 \in \mathcal{P}(E)$ be the initial distribution. The corresponding *Feynman-Kac interacting particle model* is a sequence of nonhomogeneous Markov chains

$$\left(E_n^N,\, \mathcal{F}_n^N,\, \xi^{(N)} = (\xi_n^{(N)})_{n\in\mathbb{N}},\, \mathbb{P}_{\nu_0}^N \right)$$

taking values at each time $n \in \mathbb{N}$ in the product space $E_n^N := E_n \times \cdots \times E_n$. The initial configuration ξ_0 consists of N i.i.d. random variables with common law ν_0. The transitions are given by

$$\mathbb{P}_{\nu_0}^N(\xi_n^{(N)} \in dx_n \,|\, \xi_{n-1}^{(N)}) = \prod_{p=1}^{N} \Phi_n(m(\xi_{n-1}^{(N)}))(dx_n^p), \tag{1.12}$$

where $m(\xi_{n-1}^{(N)})(\,\cdot\,) := \frac{1}{N}\sum_{i=1}^N \delta_{\xi_{n-1}^i}(\,\cdot\,) \in \mathcal{P}(E_{n-1})$ and $dx_n = (dx_n^1, \ldots, dx_n^N)$ is an infinitesimal neighbourhood of a point $x_n \in E_n^N$.

In what follows, when there is no possible confusion, we write ξ_n instead of $\xi_n^{(N)}$.

In case the N-particle system is related to an operator Φ_n of the form (1.6), the elementary transitions (1.12) are given by

$$\begin{aligned}\mathbb{P}^N_{\nu_0}(\xi_{n+1} \in dx_{n+1} \,|\, \xi_n) &= \prod_{p=0}^{N} \big(\Psi_n(m(\xi_n))\, M_n\big)(dx^p_{n+1}) \\ &= \int_{E_n^N} \Psi_n^N(m(\xi_n))(\,dy_n) \,\cdot\, M^N_{n+1}(y_n,\, dx_{n+1}), \qquad (1.13)\end{aligned}$$

where the Boltzmann-Gibbs transformations $\Psi_n^N : \mathcal{P}(E_n) \to \mathcal{P}(E_n^N)$ are defined by

$$\Psi_n^N(\eta)(dy_n) := \prod_{p=1}^{N} \Psi_n(\eta)(dy_n^p),$$

for every $\eta \in \mathcal{P}(E_n)$, and the mutation transitions $M_n^N : E_{n-1}^N \to E_n^N$ are defined by

$$M_n^N(y_{n-1},\, dx_n) = \prod_{p=1}^{N} M_n(y^p_{n-1}, dx^p_n).$$

The integral decomposition (1.13) shows that this particle model has the same updating/prediction nature as that of the limiting Feynman-Kac model. More precisely, introduce the empirical measures $(\eta_n^N,\, \widehat{\eta}_n^N) \in \mathcal{P}(E_n) \times \mathcal{P}(E_n)$ given by

$$\eta_n^N := \frac{1}{N}\sum_{i=1}^{N} \delta_{\xi_n^i} \quad \text{and} \quad \widehat{\eta}_n^N := \frac{1}{N}\sum_{i=1}^{N} \delta_{\widehat{\xi}_n^i}\,.$$

Replacing the normalised Feynman-Kac measures $(\eta_n,\, \widehat{\eta}_n)$ by the empirical measures $(\eta_n^N,\, \widehat{\eta}_n^N)$ in the two-step updating/prediction transitions in the distribution spaces $\mathcal{P}(E_n)$ described in (1.9), we obtain a two-step selection/mutation transitions on $\mathcal{P}(E_n)$

$$\eta_n^N \in \mathcal{P}(E_n) \xrightarrow{\text{selection}} \widehat{\eta}_n^N \in \mathcal{P}(E_n) \xrightarrow{\text{mutation}} \eta_{n+1}^N \in \mathcal{P}(E_{n+1})\,,$$

which is equivalent to a two-step selection/mutation transitions in product spaces

$$\xi_n \in E_n^N \xrightarrow{\text{selection}} \widehat{\xi}_n \in E_n^N \xrightarrow{\text{mutation}} \xi_{n+1} \in E_{n+1}^N\,. \qquad (1.14)$$

The initial configuration ξ_0 consists of N random variables $\xi_0^i \in E_0$ i.i.d. with common law ν_0. The two steps involved in the process can be described as follows:

Selection: Given a configuration $\xi_n \in E_n^N$ of the system at time n, the selection transition consists in selecting randomly N states $\widehat{\xi}_n^i$ with respective distributions

$$\widehat{\xi}_n^i \sim \Psi_{m(\xi_n),\,n}.$$

In other words, we select randomly an index $j \in \{1, \dots, n\}$ with distribution

$$\Psi_n(m(\xi_n)) = \sum_{i=1}^{N} \frac{G_n(\xi_n^i)}{\sum_{k=1}^{N} G_n(\xi_n^k)} \delta_{\xi_n^i} \tag{1.15}$$

and we set $\widehat{\xi}_n^i = \xi_n^j$.

Mutation: During the mutation stage, each selected particle $\widehat{\xi}_n^i$ evolves randomly according to the Markov transition M_{n+1}. In other words, given the selected configuration $\widehat{\xi}_n \in E_n^N$, the mutation transition consists in sampling randomly N independent random states ξ_{n+1}^i with respective distributions $M_{n+1}(\widehat{\xi}_n^i, \cdot)$.

Chapter 2

The Large-Deviation Conditioning Problem

The main purpose of the chapter is to provide a numerical procedure based on Feynman-Kac models to study the long-time limit behaviour of a discrete-time Markov process conditioned on rare events. In Section 2.1 we introduce the basic notation for defining the large-deviation conditioning problem and we also provide the main statistical mechanics results for the study of the problem, introducing the canonical path ensemble and the scaled cumulant generating function (SCGF).

Then, in Section 2.2, we will apply the theory of Feynman-Kac models to our problem of large deviation conditioning and show that the canonical path ensemble can be seen as a (homogeneous) Feynman-Kac path model.

We conclude the chapter providing the construction of the cloning algorithm associated to the Feynman-Kac N-particle model that estimates the SCGF. In Chapter 3, we will give rigorous results for estimating the approximation errors and the order of convergence as the size of the system N tends to ∞.

2.1 Basic Notation and Definitions

In this section we provide a brief overview of the large-deviation conditioning problem [3]. The results presented here allow us to study the long-time limit behaviour of a discrete-time Markov process conditioned on rare events introducing the associated canonical path ensemble and the scaled cumulant generating function. In the following sections, we will see how these quantities can be interpreted using Feynman-Kac models.

Let $(S, \mathcal{S})$ be $\mathbb{R}^d$ or a counting space, equipped respectively with the Lebesgue or counting topology. Let $\Omega_n := S^{n+1}$ be the path space of the trajectories until time n, for every $n \in \mathbb{N}_0$. We consider a discrete-time Markov process

$$\left(\Omega := S^{\mathbb{N}_0},\, \mathcal{F} = (\mathcal{F}_n)_{n\in\mathbb{N}},\, X = (X_n)_{n\in\mathbb{N}_0},\, \mathbb{P}_{\mu_0}\right),$$

with homogeneous Markov kernel

$$\mathbb{P}(X_n \in dx_n \,|\, X_{n-1} = x_{n-1}) =: M(x_{n-1},\, dx_n),$$

and with initial distribution $\mu_0 \in \mathcal{P}(S)$. As usual, we use the natural filtration generated by the process, i.e. $\mathcal{F}_n$ is the smallest σ-algebra on Ω such that all X_p, $p = 0, \ldots, n$, are measurable.

We consider observables $A_n : \Omega_n \to \mathbb{R}$ over the time interval $[0,\, n]$ of the form

$$A_n := \frac{1}{n}\sum_{k=1}^{n} g(X_{k-1},\, X_k) + f(X_{k-1}), \tag{2.1}$$

where $f \in \mathcal{B}_b(S)$ and $g \in \mathcal{B}_b(S^2)$ are bounded measurable functions.

Remark. We require f and g to be bounded so that the potential functions we will construct from A_n in Section 2.2 will be bounded. However, this requirement is not necessary for obtaining bounded potential functions (see, for instance, Example 3).

Example. The class of observables given by (2.1) includes many random variables of mathematical and physical interest, including:

- the occupation time in some set B, obtained with $f = \mathbb{1}_B$ and $g = 0$;
- the particle current across a particular bond $(i,\, i+1)$, obtained with $f = 0$ and $g(x, y) = \mathbb{1}_{\{i\}}(x) \cdot \mathbb{1}_{\{i+1\}}(y) \,-\, \mathbb{1}_{\{i+1\}}(x) \cdot \mathbb{1}_{\{i\}}(y)$;
- the action functional [10] obtained by setting $f = 0$ and $g(x, y) = \log \frac{M(x,y)}{M(y,x)}$, provided $\inf_{x,y} M(x, y) > 0$.

Assumption 2.1.1. The observable A_n is assumed to satisfy a *large deviation principle* (LDP) with respect to $\mathbb{P}_{\mu_0,n}$ with rate function I, that is the function $I : S \to [0, \infty]$ is lower semi-continuous with compact level sets such that $I \not\equiv \infty$ and the following bounds hold:

- $\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}_{\mu_0,n}(A_n \in C) \leq -\inf_{x\in C} I(x) \quad \forall C \subset \mathbb{R}$ closed.
- $\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}_{\mu_0,n}(A_n \in O) \geq -\inf_{x\in O} I(x) \quad \forall O \subset \mathbb{R}$ open.

We are interested in studying the long-time limit behaviour of $\mathbb{P}_{\mu_0,n}(\mathcal{A}_n)$ and of the probability path measure of the conditioned process $X_n|_{\mathcal{A}_n}$, where $\mathcal{A}_n$ is a general measurable event

$$\mathcal{A}_n := \{\omega \in \Omega_n \,|\, A_n(\omega) \in da\}$$

of sample paths satisfying the constraint that $A_n \in da$, $a \in \mathbb{R}$.

The probability path measure of the conditioned process can be seen as the microcanonical path ensemble

$$\mathbb{P}^a_{\mu_0,n}(d\omega) := \mathbb{P}_{\mu_0,n}(d\omega | A_n \in da).$$

Applying general statistical mechanics results [4], we have that, under certain hypotheses, $\mathbb{P}^a_{\mu_0,n}$ and the canonical path ensemble

$$\mathbb{P}_{\mu_0,n,k}(d\omega) := \frac{e^{knA_n(\omega)}\mathbb{P}_{\mu_0,n}(d\omega)}{\mathbb{E}_{\mu_0}[e^{knA_n}]}, \quad k \in \mathbb{R}, \tag{2.2}$$

are equivalent at three different levels (namely thermodynamical, observable and measure). In order to state these equivalence results rigorously, we need to introduce the *scaled cumulant generating function* (SCGF) of the observable A_n for every $k \in \mathbb{R}$:

$$\Lambda(k) := \lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}_{\mu_0}[e^{knA_n}], \tag{2.3}$$

which is defined provided the limit exists.

Theorem 2.1.2 (Thermodynamical Equivalence)**.** *Assume that, for every $k \in \mathbb{R}$,*

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{E}_{\mu_0}[e^{knA_n}] < \infty.$$

Then, for each $k \in \mathbb{R}$, the limit $\Lambda(k)$ exists, is finite and satisfies

$$\Lambda(k) = \sup_{a\in\mathbb{R}}\{k \cdot a - I(a)\}.$$

Furthermore, if I is strictly convex at a, then

$$I(a) = \sup_{k\in\mathbb{R}}\{k \cdot a - \Lambda(k)\}.$$

Proof. See [11], Theorem 4.5.10. □

Theorem 2.1.3 (Observable Equivalence)**.** *Let B_n be an observable which satisfies* *an LDP with respect to the canonical path measure $\mathbb{P}_{\mu_0,n,k}$ with rate function I_k and an LDP with respect to the microcanonical path measure $\mathbb{P}^a_{\mu_0,n}$ with rate function I^a.*

Under the assumptions of Theorem 2.1.2, if I is strictly convex at a, then $\mathcal{E}^a = \mathcal{E}_k$ for all $k \in \partial I(a)$, where

$$\mathcal{E}^a := \{b \in \mathbb{R} \,|\, I^a(b) = 0\}, \qquad \mathcal{E}_k := \{b \in \mathbb{R} \,|\, I_k(b) = 0\},$$

and $\partial I(a)$ is the subdifferential set of I at a, that is

$$\partial I(a) := \{x \in \mathbb{R} \,|\, x \cdot a - I(a) \leq x \cdot y - I(y), \quad y \in \mathbb{R}\}.$$

Proof. See [4], Theorem 7. □

Theorem 2.1.4 (Measure Equivalence)**.** *Under the assumptions of Theorem 2.1.2, if I is strictly convex at a, then*

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}^a_{\mu_0,n} = \lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}_{\mu_0,n,k}\,,$$

almost everywhere with respect to both $\mathbb{P}^a_{\mu_0,n}$ and $\mathbb{P}_{\mu_0,n,k}$.

Proof. See [4], Theorem 12. □

In the following sections, we will provide the Feynman-Kac interpretation of the canonical path ensemble and, in particular, we construct the N-particle system which allows us to approximate the SCGF.

2.2 The Feynman-Kac Interpretation

In this section, we apply the theory of Feynman-Kac models developed in Chapter 1 to the problem of large deviation conditioning. In particular, using the Structural Stability Properties discussed in Section 1.3, we provide a Feynman-Kac interpretation of the canonical math measure and of the SCGF which is associated to a time-homogeneous potential/kernel pair $(\widehat{G}, \widehat{M})$, defined on the state space S.

Lemma 2.2.1. *Let $\mathcal{X}_n := X_{0:n} \in \Omega_n$, $n \in \mathbb{N}$, be the Markov path process associated to the Markov chain X_n with initial distribution μ_0 and transition kernel M. We denote by $\mathcal{M}_n$ the corresponding Markov kernel and introduce the sequence of (bounded) potential functions*

$$\mathcal{G}_n(\mathcal{X}_n) := \exp(kg(X_{n-1},\, X_n) + kf(X_{n-1})),$$

for $n \geq 1$, where $\mathcal{X}_n = (X_0, \dots, X_n) \in \Omega_n$, assuming $\mathcal{G}_0(\mathcal{X}_0) \equiv 1$.

Then, the canonical path ensemble (2.2) *and the scaled cumulant generating function* (2.3) *can be written respectively as*

$$\mathbb{P}_{\mu_0,n,k} = \widehat{\eta}_n^{(\mathcal{G}_n,\mathcal{M}_n)} \tag{2.4}$$

and

$$\Lambda(k) = \lim_{n\to\infty} \frac{1}{n} \log \widehat{\gamma}_n^{(\mathcal{G}_n,\mathcal{M}_n)}(1)\,, \tag{2.5}$$

where $\widehat{\eta}_n^{(\mathcal{G}_n,\mathcal{M}_n)}$ and $\widehat{\gamma}_n^{(\mathcal{G}_n,\mathcal{M}_n)}$ are the normalised and unnormalised updated flows associated to the pairs $(\mathcal{G}_n, \mathcal{M}_n)_{n\in\mathbb{N}}$ and with initial distribution μ_0.

Proof. The statement follows easily from the definitions (2.2)-(2.3). Indeed, we have that

$$\mathbb{P}_{\mu_0,n,k}(d\omega^{(n)}) = \frac{\prod_{p=0}^n \mathcal{G}_p(\omega^{(p)}) \,\cdot\, \mathbb{P}_{\mu_0,n}(d\omega^{(n)})}{\mathbb{E}_{\mu_0}\big[\prod_{p=0}^n \mathcal{G}_p(\mathcal{X}_p)\big]} = \widehat{\eta}_n^{(\mathcal{G}_n,\mathcal{M}_n)}(d\omega^{(n)}),$$

for any $\omega^{(n)} := (\omega_1, \dots, \omega_n) \in \Omega_n$, and

$$\mathbb{E}_{\mu_0}\big[e^{knA_n}\big] = \mathbb{E}_{\mu_0}\Big[\prod_{p=0}^n \mathcal{G}_p(\mathcal{X}_p)\Big] = \widehat{\gamma}_n^{(\mathcal{G}_n,\mathcal{M}_n)}(1).$$

□

Lemma 2.2.2. *Consider the sequence of pairs $(\widehat{\mathcal{G}}_n,\, \widehat{\mathcal{M}}_n)_{n\in\mathbb{N}}$ defined by*

$$\widehat{\mathcal{G}}_n(\omega^{(n)}) := \int_{\Omega_{n+1}} \mathcal{M}_{n+1}(\omega^{(n)},\, d\omega^{(n+1)})\, \mathcal{G}_{n+1}(\omega^{(n+1)})\,,$$
$$\widehat{\mathcal{M}}_n(\omega^{(n-1)},\, d\omega^{(n)}) := \frac{\mathcal{M}_n(\omega^{(n-1)},\, d\omega^{(n)})\, \mathcal{G}_n(\omega^{(n)})}{\widehat{\mathcal{G}}_{n-1}(\omega^{(n-1)})}\,,$$

for any $\omega^{(n-1)} \in \Omega_{n-1}$ and $\omega^{(n)} \in \Omega_n$, $n \in \mathbb{N}$, with $(\mathcal{G}_n,\, \mathcal{M}_n)_{n\in\mathbb{N}}$ given by Lemma 2.2.1.

Then, the canonical path ensemble (2.2) *and the scaled cumulant generating function* (2.3) *can be written respectively as*

$$\mathbb{P}_{\mu_0,n,k} = \eta_n^{(\widehat{\mathcal{G}}_n,\, \widehat{\mathcal{M}}_n)} \tag{2.6}$$

and

$$\Lambda(k) = \lim_{n\to\infty} \frac{1}{n} \log \gamma_n^{(\widehat{\mathcal{G}}_n,\, \widehat{\mathcal{M}}_n)}(1), \tag{2.7}$$

where $\eta_n^{(\widehat{\mathcal{G}}_n,\, \widehat{\mathcal{M}}_n)}$ and $\gamma_n^{(\widehat{\mathcal{G}}_n,\, \widehat{\mathcal{M}}_n)}$ are the normalised and unnormalised prediction flows *associated with the sequence of pairs* $(\widehat{\mathcal{G}}_n,\, \widehat{\mathcal{M}}_n)$ *and with initial distribution* $\widehat{\mu}_0(dx_0) := \frac{\mathcal{G}_0(x_0)}{\mu_0(\mathcal{G}_0)}\mu_0(dx_0)$.

Proof. Using the Feynman-Kac interpretation of the canonical path ensemble and SCGF given by Lemma 2.2.1, we have that the statement follows straightforward by applying the second Structural Stability Property (Proposition 1.3.3) to the normalised updated flow $\widehat{\eta}_n^{(\mathcal{G}_n,\mathcal{M}_n)}$ associated to the pairs $(\mathcal{G}_n, \mathcal{M}_n)_{n\in\mathbb{N}}$ and with initial distribution μ_0.

□

This result tells us that the canonical path ensemble and the SCGF can be estimated through the Feynman-Kac N-particle system described in Section 1.4 associated to $(\widehat{\mathcal{G}}_n,\, \widehat{\mathcal{M}}_n)_{n\in\mathbb{N}}$. One of the main advantages of considering the prediction time marginal $\eta_n^{(\widehat{\mathcal{G}}_n,\, \widehat{\mathcal{M}}_n)}$ associated to the pairs $(\widehat{\mathcal{G}}_n,\, \widehat{\mathcal{M}}_n)_{n\in\mathbb{N}}$ instead of the updated time marginal $\widehat{\eta}_n^{(\mathcal{G}_n,\mathcal{M}_n)}$ associated to $(\mathcal{G}_n, \mathcal{M}_n)_{n\in\mathbb{N}}$ is that the non-homogeneous pairs potential/kernel $(\widehat{\mathcal{G}}_n,\, \widehat{\mathcal{M}}_n)_{n\in\mathbb{N}}$ on the path spaces Ω_n can be rewritten more practically as homogeneous pairs potential/kernel $(\widehat{G},\, \widehat{M})$ on the state space S.

Proposition 2.2.3. *Consider the time-homogeneous potential/kernel pair* $(\widehat{G}, \widehat{M})$ *on* S *given by*

$$\widehat{G}(x_n) := \int_S F(x_n,\, z)M(x_n,\, dz)\,, \tag{2.8}$$

$$\widehat{M}(x_n,\, dy_{n+1}) := \frac{F(x_n,\, y_{n+1})M(x_n,\, dy_{n+1})}{\widehat{G}(x_n)}., \tag{2.9}$$

where

$$F(X_{n-1},\, X_n) := \exp(kg(X_{n-1},\, X_n) + kf(X_{n-1})). \tag{2.10}$$

The canonical path ensemble (2.2) *and the SCGF* (2.3) *can be written respectively as*

$$\mathbb{P}_{\mu_0,n,k} \;=\; \mathbb{Q}_{\mu_0,n}^{(\widehat{G},\,\widehat{M})} \tag{2.11}$$

and

$$\Lambda(k) = \lim_{n\to\infty} \frac{1}{n} \log \gamma_n^{(\widehat{G},\,\widehat{M})}(1)\,, \tag{2.12}$$

where $\mathbb{Q}_{\mu_0,n}^{(\widehat{G},\,\widehat{M})}$ *and* $\gamma_n^{(\widehat{G},\,\widehat{M})}$ *are the Feynman-Kac prediction path model and the corresponding unnormalised prediction flow associated to the sequence of pairs* $(\widehat{G},\, \widehat{M})_{n\in\mathbb{N}}$ *and with initial distribution* μ_0.

Proof. First of all, observe that F (and thus also $\widehat{G}$) is bounded. Moreover, we can see from the definitions that $F(X_{n-1}, X_n) = \mathcal{G}_n(\mathcal{X}_n)$. Therefore, we can write

$$\begin{aligned}\widehat{\mathcal{G}}_n(\theta_n) &= \int_{\Omega_{n+1}} \mathcal{G}_{n+1}(\theta_{n+1})\mathcal{M}_{n+1}(\theta_n,\, d\theta_{n+1}) \\ &= \int_S F(x_n,\, z)M(x_n,\, dz) \;=\; \widehat{G}(x_n),\end{aligned}$$

for any $\theta_n = (x_0, \ldots, x_n) \in \Omega_n$. In particular, $\widehat{\mathcal{G}}_n$ is a potential function of the path space Ω_n depending only on the terminal state of the trajectories.

Moreover, for any $\theta_n = (x_0, \ldots, x_n) \in \Omega_n$ and $\theta_{n+1} = (y_0, \ldots, y_{n+1}) \in \Omega_{n+1}$ we have that

$$\begin{aligned}\widehat{\mathcal{M}}_{n+1}(\theta_n,\, d\theta_{n+1}) \;&=\; \frac{F(x_n,\, y_{n+1})M(x_n,\, dy_{n+1})}{\widehat{G}(x_n)} \cdot \delta_{\theta_n}(dy_0, \ldots, dy_n) \\ &=\; \widehat{M}(x_n,\, dy_{n+1}) \cdot \delta_{\theta_n}(dy_0, \ldots, dy_n).\end{aligned}$$

This means that $\widehat{\mathcal{M}}_n$, $n \in \mathbb{N}$, are the transitions of the path process associated to the homogeneous Markov chain on S given by the kernel $\widehat{M}$ and with initial distribution $\widehat{\mu}_0$. Recalling the definition of $\widehat{\mu}_0$ given in Lemma 2.2.2, we can see that

$$\widehat{\mu}_0(dx_0) = \frac{\mathcal{G}_0(x_0)}{\mu_0(\mathcal{G}_0)}\mu_0(dx_0) = \mu_0(dx_0),$$

for all $x_0 \in S$, since $\mathcal{G}_0 \equiv 1$.

Therefore we can apply Corollary 1.3.2 (first Structural Stability Property of Feynman-Kac models) to the sequence of pairs $(\widehat{\mathcal{G}}_n, \widehat{\mathcal{M}}_n)_{n\in\mathbb{N}}$. Combining this result with Lemma 2.2.2, we obtain Equation (2.11).

Finally, recalling observation (1.2), we can see that for every bounded function $f_n \in \mathcal{B}_b(\Omega_n)$ depending only on the terminal point of the paths, i.e. $f_n(x_0, \ldots, x_n) = f(x_n)$, $f \in \mathcal{B}_b(S)$, we can write

$$\eta_n^{(\widehat{\mathcal{G}}_n,\, \widehat{\mathcal{M}}_n)}(f_n) = \eta_n^{(\widehat{G},\, \widehat{M})}(f). \tag{2.13}$$

In particular, (2.13) holds for $f_n = \widehat{\mathcal{G}}_n$. Recalling the relation (1.3), we can see that

$$\gamma_n^{(\widehat{\mathcal{G}}_n,\, \widehat{\mathcal{M}}_n)}(1) = \gamma_n^{(\widehat{G},\, \widehat{M})}(1).$$

This, applied with Lemma 2.2.2, concludes the proof. □

Proposition 2.2.3 tells us that the canonical ensemble (2.2) can be seen as the Feynmnan-Kac prediction path model associated to a time-homogeneous potential/kernel pair $(\widehat{G},\, \widehat{M})$ on the state space S. Moreover, thanks to equality (2.12), we will see that we can estimate the SCGF through simulations of Feynman-Kac particle systems associated to the pair $(\widehat{G},\, \widehat{M})$.

2.3 The Cloning Algorithm

Cloning algorithms are numerical procedures aimed at simulating rare events efficiently, using a population dynamic scheme. In such algorithms, copies of the system are evolved in parallel and the ones showing the rare behaviour of interest are multiplied iteratively. This class of algorithms is also used to evaluate numerically the scaled cumulant generating function (2.3) in several physical applications, including chaotic systems, glassy dynamics and non-equilibrium lattice models. Even if it is heuristically believed that the estimator originated by the cloning algorithm converges to the correct value of the SCGF, there are only few studies focusing on the analytical justification of the algorithm [5; 6].

We propose here a different cloning algorithm which makes use of the N-particle interpretation of Feynman-Kac models. This approach enables us to study analytically the convergence of the algorithm to the quantity of interest. Novel rigorous convergence results will be presented in Chapter 3.

Recalling the construction of the process in Section 1.4, the Feynman-Kac interacting particle model associated to $(\widehat{G},\, \widehat{M})$ is a sequence of homogeneous Markov chains

$$\Big(S^N,\, \mathcal{F}^N,\, \xi^{(N)} = \big(\xi_n^{(N)}\big)_{n\in\mathbb{N}},\, \mathbb{P}^N_{\mu_0}\Big)$$

taking values in the product space $S^N := S \times \cdots \times S$. The initial configuration ξ_0 consists of N i.i.d. random variables with common law μ_0. The transitions are given by

$$\mathbb{P}^N_{\mu_0}(\xi_{n+1} \in dx_{n+1} \,|\, \xi_n) \;=\; \int_{S^N} \Psi^N(m(\xi_n))(dy_n) \cdot \widehat{M}^N(y_n,\, dx_{n+1}),$$

where $\Psi^N : \mathcal{P}(S) \to \mathcal{P}(S^N)$ is defined by

$$\Psi^N(\eta)(dy) \;=\; \prod_{p=1}^{N} \Psi(\eta)(dy^p) \;=\; \prod_{p=1}^{N} \frac{\widehat{G}(y^p)}{\eta(\widehat{G})}\, \eta(dy^p),$$

for every $\eta \in \mathcal{P}(S)$ and $y \in S^N$, and the Markov transitions $\widehat{M}^N : S^N \to S^N$ are defined by

$$\widehat{M}^N(y_{n-1},\, dx_n) \;=\; \prod_{p=1}^{N} \widehat{M}(y^p_{n-1},\, dx^p_n). \tag{2.14}$$

As explained in Section 1.4, the underlying two-step process $(\xi_n, \widehat{\xi}_n)_{n\in\mathbb{N}}$ can be described as follows:

Selection: Given a configuration $\xi_n \in S^N$ of the system at time n, the selection transition consists in randomly selecting, for every $i = 1\dots, N$, an index $j \in \{1, \dots, N\}$ with distribution

$$\Psi(m(\xi_n)) := \sum_{i=1}^{N} \frac{\widehat{G}(\xi_n^i)}{\sum_{k=1}^{N} \widehat{G}(\xi_n^k)} \delta_{\xi_n^i},$$

and we set $\widehat{\xi}_n^i = \xi_n^j$.

Mutation: Given a configuration $\widehat{\xi}_n \in \Omega_n^N$, the mutation transition consists in sampling randomly N independent random states ξ_{n+1}^i with time-homogeneous kernels $\widehat{M}(\widehat{\xi}_n^i, \cdot)$, $i = 1, \dots, N$.

The model described above can be thought as the evolution of a population whose individuals reproduce and die subject to a natural evolution interaction mechanism. During the selection stage, the elements with high potential $\widehat{G}$ are more likely to be multiplied, while elements with low potential are eliminated. Then, each individual evolves and mutates according to transition $\widehat{M}$.

In order to estimate the scaled cumulant generating function $\Lambda(k)$ which satisfies relation (2.12), we need to introduce the approximation measures associated to the Feynman-Kac flow $(\gamma_n, \eta_n) := (\gamma_n^{(\widehat{G}, \widehat{M})}, \eta_n^{(\widehat{G}, \widehat{M})})$.

Definition 2.3.1. The N-particle *approximation measures* associated with the Feynman-Kac flows (γ_n, η_n) are given by:

$$\eta_n^N := \frac{1}{N} \sum_{i=1}^{N} \delta_{\xi_n^i} \in \mathcal{P}(S), \tag{2.15}$$

$$\gamma_n^N(\cdot) := \eta_n^N(\cdot) \prod_{p=1}^{n-1} \eta_p^N(\widehat{G}) \in \mathcal{M}_+(S), \tag{2.16}$$

where $\mathcal{M}_+(S)$ denotes the set of positive measures on S.

Remark. The following identities hold

$$\gamma_n^N(f) = \gamma_n^N(1) \cdot \eta_n^N(f), \tag{2.17}$$

$$\gamma_{n+1}^N(1) = \gamma_n^N(1) \cdot \eta_n^N(\widehat{G}), \tag{2.18}$$

where $f \in \mathcal{B}_b(S)$.

Algorithm 1 represents the cloning algorithm for the SCGF estimation and it is exactly the mean field particle interpretation of the model within Del Moral's framework [7].

Algorithm 1 Cloning Algorithm for simulating the SCGF

for $i \leftarrow 1, N$ **do**
  Sample $\xi_0^i \sim \mu_0$;
  Compute $\widehat{G}(\xi_0^i)$;
Compute $\eta_0^N(\widehat{G}) = \frac{1}{N}\sum_{i=1}^{N} \widehat{G}(\xi_0^i)$;
Set $\gamma_1^N(1) = \eta_0^N(\widehat{G})$;
for $p \leftarrow 1, n-1$ **do**
  for $i \leftarrow 1, N$ **do**
    Sample $\widehat{\xi}_{p-1}^i \sim \Psi(m(\xi_{p-1}))$;
    Sample $\xi_p^i \sim \widehat{M}\big(\widehat{\xi}_{p-1}^i, \cdot\big)$;
    Compute $\widehat{G}(\xi_p^i)$;
  Compute $\eta_p^N(\widehat{G}) = \frac{1}{N}\sum_{i=1}^{N} \widehat{G}(\xi_p^i)$;
  Update $\gamma_{p+1}^N(1) = \gamma_p^N(1) \cdot \eta_p^N(\widehat{G})$;
Return $\frac{1}{n} \log \gamma_n^N(1)$.

In Chapter 3 we will show some rigorous convergence results of the approximation measures (γ_n^N, η_n^N) with respect to the Feynman-Kac flow (γ_n, η_n). In particular, we will discuss the order of convergence of the approximation error associated to the quantity $\frac{1}{n}\log \gamma_n^N(1) - \Lambda(k)$.

2.4 Toy Examples

We conclude the chapter providing some simple examples that illustrate the construction of the pair $(\widehat{G}, \widehat{M})$ which provides the implementation of the Algorithm 1. We will reconsider these toy examples also at the end of Chapter 3, for illustrating how to apply the convergence results we will achieve.

Example 1. Let $S = \mathbb{Z}$ and consider the random walk defined by the transition probabilities $M(x, x+1) = p \in (0,1)$ and $M(x, x-1) = 1-p$ for all $x \in S$. We also introduce the observable $A_n = (X_n - X_0)/n$, so that we have the operator F from (2.10) is given by $F(x,y) = \exp(ky - kx)$. Note that the operator F thus defined is not bounded, but we can overcome the problem considering $F(x,y) = \exp((ky-kx)\mathbb{1}_{\{x+1,x-1\}}(y))$ instead, since the allowed jumps from x are only on $x-1$ and $x+1$, for all $x \in S$.

By construction, we can write $\widehat{G}$ (2.8) as

$$\widehat{G}(x) = p \cdot e^{k(x+1)-kx} \;+\; (1-p) \cdot e^{k(x-1)-kx} = p \cdot e^{k} + (1-p) \cdot e^{-k} =: K,$$

and the mutation transitions $\widehat{M}$ (2.9) as

$$\widehat{M}(x,\, x+1) = \frac{p \cdot e^{k}}{K}\,, \qquad \widehat{M}(x,\, x-1) = \frac{(1-p) \cdot e^{-k}}{K}\,,$$

for all $x \in \mathbb{Z}$.

Observe that the mutation transition $\widehat{M}$ corresponds to the original process transition M with an added drift $e^{\pm k}/K$. Moreover, the potential function $\widehat{G}$ is independent of $x \in \mathbb{Z}$, therefore the selection stage performed in the cloning algorithm consists in selecting uniformly an index j from the set $\{1, \ldots, N\}$ and set $\widehat{\xi}^{i}_{n} = \xi^{j}_{n}$, for every $i = 1, \ldots, N$.

Therefore, in this simple application we can see that $\eta^{N}_{n}(\widehat{G}) \equiv K$, for all $n \in \mathbb{N}$, and thus $\gamma^{N}_{n}(1) = K^{n}$. This implies

$$\lim_{n\to\infty} \frac{1}{n} \log \gamma^{N}_{n}(1) \;=\; \log K.$$

We want to show that $\log K = \Lambda(k)$. For every $n \in \mathbb{N}$, we have that

$$\begin{aligned} \mathbb{E}_{\mu_0}[e^{k(X_{n+1}-X_0)}] \;&=\; p \cdot \mathbb{E}_{\mu_0}[e^{k(X_n - X_0 + 1)}] + (1-p) \cdot \mathbb{E}_{\mu_0}[e^{k(X_n - X_0 - 1)}] \\ &=\; K \cdot \mathbb{E}_{\mu_0}[e^{k(X_n - X_0)}]. \end{aligned}$$

By induction, we obtain

$$\mathbb{E}_{\mu_0}[e^{k(X_n - X_0)}] \;=\; K^{n} \cdot \mathbb{E}_{\mu_0}[1] \;=\; K^{n}\,.$$

Thus,

$$\Lambda(K) \;=\; \lim_{n\to\infty} \frac{1}{n} \log\left(\mathbb{E}_{\mu_0}[e^{k(X_n - X_0)}]\right) \;=\; \log K.$$

Example 2. Let $S = \{1, \ldots, m\}$ finite with periodic boundary condition $m + 1 = 1$. We consider, as above, the random walk defined by the transition probabilities $M(x,\, x{+}1) = p \in (0, 1)$ and $M(x,\, x{-}1) = 1{-}p$ for all $x \in S$. We are interested in the occupation time in $\{1\}$, so that the considered observable is $A_n = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\{1\}}(X_i)$. In this case, $F(x, y) = \exp(k \cdot \mathbb{1}_{\{1\}}(y))$.

Therefore,

$$\widehat{G}(m) = p \cdot e^{k} + (1-p), \qquad \widehat{G}(2) = (1-p) \cdot e^{k} + p,$$

and $\widehat{G}(x) = 1$ for $x \neq m$, 2. Moreover,

$$\widehat{M}(m,1) = \frac{p \cdot e^k}{\widehat{G}(m)}, \qquad \widehat{M}(m,m-1) = \frac{(1-p)\cdot e^k}{\widehat{G}(m)},$$
$$\widehat{M}(2,3) = \frac{p \cdot e^k}{\widehat{G}(2)}, \qquad \widehat{M}(2,1) = \frac{(1-p)e^k}{\widehat{G}(2)},$$

and $\widehat{M}(x,x+1) = p$ or $\widehat{M}(x,x-1) = 1-p$, otherwise.

Example 3. We consider the same observable of Example 1, i.e. $A_n = (X_n - X_0)/n$, but this time on state space $S = \mathbb{R}$ and with Markov transitions $M(x,dy) = \frac{1}{\sqrt{2\pi}} e^{-(y-x)^2/2} dy$, given by the gaussian distribution centred in x. In this case $F(x,y) := \exp(ky - kx)$ is not bounded (so it doesn't fall within the cases considered) however we can see that

$$\begin{aligned} \widehat{G}(x) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{k(y-x)} \cdot e^{-(y-x)^2/2} dy \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{\frac{-(y-(x+k))^2+k^2}{2}} = e^{k^2/2}, \end{aligned}$$

in particular $\widehat{G}$ is a bounded potential function, so the convergence results from Feynman-Kac theory presented in Chapter 3 still apply (provided the hypothesis are satisfied).

Note also that the mutation transitions are given by

$$\widehat{M}(x,dy) = e^{-k^2/2} \cdot \frac{1}{\sqrt{2\pi}} e^{\frac{-(y-(x+k))^2+k^2}{2}} dy = \frac{1}{\sqrt{2\pi}} e^{\frac{-(y-(x+k))^2}{2}} dy,$$

for all $x, y \in \mathbb{R}$.

As in Example 1, since $\widehat{G}$ is constant on $S = \mathbb{R}$, we can see that

$$\lim_{n\to\infty} \frac{1}{n} \log \gamma_n^N(1) = \log e^{k^2/2} = \frac{k^2}{2}.$$

We want to show that $\Lambda(k) = \frac{k^2}{2}$. For every $n \in \mathbb{N}$, we have that

$$\begin{aligned}
\mathbb{E}_{\mu_0}[e^{k(X_{n+1}-X_0)}] &= \int_{\mathbb{R}} \mathbb{E}_{\mu_0}\big[e^{k(X_{n+1}-X_0+y)} M(X_n,\, X_n+y)\big] dy \\
&= \int_{\mathbb{R}} e^{ky} \cdot \mathbb{E}_{\mu_0}\left[e^{k(X_{n+1}-X_0)} \cdot \frac{e^{-y^2/2}}{\sqrt{2\pi}}\right] dy \\
&= \mathbb{E}_{\mu_0}[e^{k(X_n-X_0)}] \cdot \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} \cdot e^{\frac{-(y-k)^2+k^2}{2}} dy \\
&= \mathbb{E}_{\mu_0}[e^{k(X_n-X_0)}] \cdot \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} \cdot e^{\frac{-(y-k)^2+k^2}{2}} dy \\
&= e^{k^2/2} \cdot \mathbb{E}_{\mu_0}[e^{k(X_n-X_0)}].
\end{aligned}$$

By induction, we obtain $\mathbb{E}_{\mu_0}[e^{k(X_n-X_0)}] = e^{n\cdot k^2/2}$. Thus,

$$\Lambda(K) = \lim_{n\to\infty} \frac{1}{n} \log\left(\mathbb{E}_{\mu_0}[e^{k(X_n-X_0)}]\right) = \frac{k^2}{2}.$$

Chapter 3

Convergence Results

We are interested in studying the approximation error and the order of convergence of the Cloning Algorithm 1 presented in Section 2.3 for estimating the SCGF (2.3), as the size of the system N goes to infinity. More precisely, recalling the definition of the approximation measures 2.3.1, we are interested in estimating the L^p-error

$$\mathbb{E}_{\mu_0}\left[\,\left|\frac{1}{n}\log\gamma_n^N(1)\;-\;\Lambda(k)\right|^p\,\right]^{1/p}, \tag{3.1}$$

for n, $N \in \mathbb{N}$ and $p \geq 0$, as well as the probability

$$\mathbb{P}\left(\left|\frac{1}{n}\log\gamma_n^N(1) - \Lambda(k)\right| > \delta\right), \tag{3.2}$$

for $\delta > 0$ and n, $N \in \mathbb{N}$.

We start presenting some of the main convergence results for general Feynman-Kac models, as introduced in [7], Section 7.4.

In Section 3.2 we apply these results for estimating the quantities (3.1) and (3.2) in the case of finite state space S and, in Section 3.3 we provide sufficient conditions for having time-homogeneous estimates in case of finite state space S.

3.1 Some Preliminary Results

To simplify the presentation, we first recall the basic notation used in the previous chapters. As seen in Section 2.2, the scaled cumulant generating function can be written as

$$\Lambda(k) = \lim_{n\to\infty}\frac{1}{n}\log\gamma_n(1),$$

where γ_n is the unnormalised prediction Feynman-Kac flow on S associated to the pair $(\widehat{G}, \widehat{M})$, where $\widehat{G} : S \to (0, \infty)$ is the potential function defined in (2.8) and $\widehat{M}$ is defined in (2.9) and it is the transition kernel of a (homogeneous) Markov process on S.

Let η_n be the corresponding normalised Feynman-Kac flow and recall the the definition of the approximation measures 2.3.1 (η_n^N, γ_n^N).

In this section we present the main relevant convergence results of the approximation measures (η_n^N, γ_n^N) to the time-marginal prediction measures (η_n, γ_n), that will allow us to estimate the quantities (3.1) and (3.2), in the following sections. The main reference for this section is [7], Section 7.4.

Proposition 3.1.1. *For each $p \geq 1$ and $n, N \in \mathbb{N}$, we have*

$$\mathbb{E}[\gamma_n^N(f)] = \gamma_n(f) \tag{3.3}$$

and

$$\mathbb{E}\Big[|\gamma_n^N(1) - \gamma_n(1)|^p \Big]^{1/p} \leq \frac{\lfloor p/2 \rfloor ! \cdot \alpha \cdot (n+1)}{\sqrt{N}},$$

where $\alpha > 0$ constant.

Proof. See [7], first part of Theorem 7.4.2, p. 239. □

Proposition 3.1.2. *Let $\delta > 0$ and $\sqrt{N} \geq 4/\delta$. For each $n \in \mathbb{N}$ we have the estimate*

$$\mathbb{P}\Big(|\gamma_n^N(1) - \gamma_n(1)| > \delta \Big) \leq 8(n+1)e^{-N\delta_n^2/2},$$

with $\delta_n := \delta/(n+1)$.

Proof. See [7], second part of Theorem 7.4.2, p. 239. □

Remark. Observe that Proposition 3.1.2 implies (using Borel-Cantelli) the strong law of large numbers for $\gamma_n^N(1)$, that is $|\gamma_n^N(1) - \gamma_n(1)|$ converges almost surely to 0 as $N \to \infty$, for any fixed $n \in \mathbb{N}$.

Under certain regularity conditions of the pair $(\widehat{G}, \widehat{M})$, the Feynman-Kac model has several regularity and asymptotic stability properties which guarantee the existence and uniqueness of an invariant measure η_∞ and also allow us to obtain time-uniform estimates of the normalised prediction measure η_n. A detailed discussion for normalised time marginals η_n^N can be found in [7], Section 7.4.3. We present here only the two main statements, which we will use in section 3.2 for studying the convergence of the cloning algorithm.

Assumption 3.1.3. We assume that:

- $\sup_{x\in S}\widehat{G}(x)<\infty$ and $\inf_{x\in S}\widehat{G}(x)>0$. In this case, we define $\varepsilon := \frac{\inf_x \widehat{G}(x)}{\sup_x \widehat{G}(x)}$;
- there exist $m\geq 1$ and $\epsilon\in(0,\,1)$ such that $\widehat{M}^m(x,\,\cdot\,)\geq \epsilon\cdot\widehat{M}^m(y,\,\cdot\,)$, for every $x,\,y\in S$.

Remark. Observe that, if S is finite, $\widehat{G}$ is strictly positive and there exists $m\geq 1$ such that $\widehat{M}^m(x,z)>0$ for all $x,\,z\in S$, then Assumption 3.1.3 holds. It is useful also to recall that, in case S is finite, there exists $m\geq 1$ such that $\widehat{M}^m(x,z)>0$ for every $x,z\in S$ if and only if $\widehat{M}$ is irreducible and aperiodic.

Remark. The potential function $\widehat{G}$ used in the definition of the cloning algorithm and given by Equation (2.8) is always strictly positive, by construction.

Assumption 3.1.3 guarantees the existence and uniqueness of an invariant measure η_∞. More formally, recall the one-step mapping $\Phi:\mathcal{P}(S)\to\mathcal{P}(S)$ defined in Proposition 1.2.6 by

$$\Phi(\mu)=\Psi(\mu)\widehat{M},$$

where $\Psi:\mathcal{P}(S)\to\mathcal{P}(S)$ is the Boltzmann-Gibbs transformation given by

$$\Psi(\mu)(dx):=\frac{1}{\mu(\widehat{G})}\widehat{G}(x)\mu(dx).$$

Definition 3.1.4. Given a mapping $\Theta:\mathcal{P}(S)\to\mathcal{P}(S)$, a measure $\mu\in\mathcal{P}(S)$ is said to be Θ-invariant if $\mu=\Theta(\mu)$.

Proposition 3.1.5. *Suppose Assumption 3.1.3 holds for some integer parameter $m\geq 1$ and some real numbers $\varepsilon,\,\epsilon\in(0,1)$. Then there exists a unique invariant measure $\eta=\Phi(\eta)\in\mathcal{P}(S)$ and for any $n\in\mathbb{N}$ and $f\in\mathcal{B}_b(S)$ we have*

$$\mathbb{E}_\eta\left[f(X_n)\prod_{p=0}^{n-1}\widehat{G}(X_p)\right]=\eta(f)\eta(\widehat{G})^n,$$

where $\mathbb{E}_\eta$ is the expectation with respect to the law of a homogeneous Markov chain X_n with transitions $\widehat{M}$ and initial distribution η.

Proof. See [7], Theorem 5.2.1, p.161. □

Assumption 3.1.3 also allows us to have time-uniform estimates of the quantities

$$\mathbb{E}\left[|\eta_n^N(f)-\eta_n(f)|^p\right]^{1/p}\qquad\text{and}\qquad\mathbb{P}\Big(|\eta_n^N(f)-\eta_n(f)|>\delta\Big),$$

for every bounded function $f : S \to \mathbb{R}$, $p \geq 1$ and $\delta > 0$.

Theorem 3.1.6. *Let $n \geq 0$, $p \geq 1$, $N \in \mathbb{N}$ and $f : S \to \mathbb{R}$ be a bounded function, $||f||_\infty := C < \infty$.*

Under Assumption 3.1.3, we have the uniform estimate

$$\mathbb{E}_{\mu_0}\left[\left|\eta_n^N(f) - \eta_n(f)\right|^p\right]^{1/p} \leq \frac{2\,C \cdot d(p)^{1/p} \cdot m}{\sqrt{N} \cdot \varepsilon^{2m+1} \cdot \epsilon^3},$$

where ε, ϵ and m are given by Assumption 3.1.3 and $d(p)$ is defined for any $p \geq 1$ by

$$d(p) = \begin{cases} (p)_k 2^{-k} & p = 2k, \\ \frac{(p)_k}{\sqrt{p/2}}\, 2^{-p/2} & p = 2k-1, \end{cases} \tag{3.4}$$

where $(p)_k := p!/(p-k)!$ with $k \in \mathbb{N}$.

Proof. See [7], Theorem 7.4.4, p.246. □

Corollary 3.1.7. *Let $n \geq 0$, $\delta > 0$, $N \in \mathbb{N}$ and $f : S \to \mathbb{R}$ be a bounded function, $||f||_\infty := C < \infty$. Under Assumption 3.1.3, we have*

$$\mathbb{P}(|\eta_n^N(f) - \eta_n(f)| > \delta) \leq (1 + \frac{\delta}{C}\sqrt{N/2})\, e^{-Nb(\delta)^2/2},$$

where $b(\delta) := \delta \cdot \varepsilon^{2m+1} \cdot \epsilon^3/(C \cdot 2m)$, with m, ε, ϵ given in Assumption 3.1.3.

Proof. See [7], Corollary 7.4.3, p.247. □

3.2 Estimation of the Scaled Cumulant Generating Function

In this section, we are interested in estimating the quantities (3.1) and (3.2) in case S finite and $\widehat{M}$ irreducible. In particular, the results presented here hold also for non aperiodic Markov chains. We will discuss the case in which Assumption 3.1.3 holds in Section 3.3.

In order to estimate (3.1) and (3.2), it is useful to evaluate first the quantities

$$\mathbb{E}_{\mu_0}\left[\,|\frac{1}{n}\log\gamma_n^N(1) \;-\; \frac{1}{n}\log\gamma_n(1)|^p\,\right]^{1/p}, \tag{3.5}$$

for $n,\ N \in \mathbb{N}$ and $p \geq 0$, and

$$\mathbb{P}\Big(|\frac{1}{n} \log \gamma_n^N(1) - \frac{1}{n} \log \gamma_n(1)| > \delta \Big), \tag{3.6}$$

for $\delta \in (0,\ 1)$ and $n,\ N \in \mathbb{N}$.

Proposition 3.2.1. *Let the state space S be $\mathbb{R}^d$ or a counting space. Consider the potential/kernel pair $(\widehat{G},\ \widehat{M})$ given by (2.8)-(2.9) and such that $c := \inf_{x \in S} \widehat{G}(x) > 0$. For each $p \geq 1$ and $n, N \in \mathbb{N}$ we have*

$$\mathbb{E}_{\mu_0}\big[\log \gamma_n^N(1)\big] \leq \log \gamma_n(1),$$

and the equality holds if and only if $\widehat{G}$ constant. Moreover,

$$\mathbb{E}_{\mu_0}\Big[|\frac{1}{n} \log \gamma_n^N(1) - \frac{1}{n} \log \gamma_n(1)|^p \Big]^{\frac{1}{p}} \leq \frac{\alpha \lfloor p/2 \rfloor !}{\sqrt{N}} \cdot \frac{n+1}{n \cdot c^n},$$

for some finite constant α.

Remark. Observe that we don't make any particular assumption on $\widehat{M}$. In particular, it is not irreducible in general. Observe also that, in the case S is finite, the assumption $c > 0$ in Proposition 3.2.1 always holds, since $\widehat{G}$ is strictly positive, by construction.

Proof. The proof is a simple application of Proposition 3.1.1. The first part of the statement follows by Jensen, indeed

$$\mathbb{E}\Big[\log \frac{\gamma_n^N(1)}{\gamma_n(1)} \Big] \leq \log \mathbb{E}\Big[\frac{\gamma_n^N(1)}{\gamma_n(1)} \Big] = 0,$$

and the equality holds if and only if $\gamma_n^N(1)$ is deterministic. By definition of γ_n^N (2.16), it holds if and only if $\eta_n^N(\widehat{G})$ (2.15) is deterministic and, thus, if and only if $\widehat{G}$ is constant on S.

For the second part of the statement, we use the fact that $|\log x - \log y| \leq |x-y| / \min\{x, y\}$, so that

$$|\log \gamma_n^N(1) - \log \gamma_n(1)| \leq \frac{|\gamma_n^N(1) - \gamma_n(1)|}{\min\{\gamma_n^N(1),\ \gamma_n(1)\}} \leq \frac{|\gamma_n^N(1) - \gamma_n(1)|}{c^n}. \tag{3.7}$$

Indeed, by definition of η_n (1.4), we have $\eta_n(\widehat{G}) \geq c$, thus, using relation (1.3), we obtain $\gamma_n(1) \geq c^n$. Analogously, we have $\gamma_n^N(1) \geq c^n$.

Using estimate (3.7) in the estimation of the L^p-error and applying Proposition 3.1.1, we can conclude

$$\begin{aligned}\frac{1}{n}\mathbb{E}_{\mu_0}\Big[|\log\gamma_n^N(1)-\log\gamma_n(1)|^p\Big]^{\frac{1}{p}} &\le \frac{1}{n\cdot c^n}\cdot\mathbb{E}\Big[|\gamma_n^N(1)-\gamma_n(1)|^p\Big]^{1/p}\\ &\le \frac{\alpha\,\lfloor p/2\rfloor!}{\sqrt{N}}\cdot\frac{n+1}{n\cdot c^n}.\end{aligned}$$

□

In particular, Proposition 3.2.1 implies that, when $\widehat{G}$ is not constant on S, $\frac{1}{n}\log\gamma_n^N(1)$ is not an unbiased estimator for $\frac{1}{n}\log\gamma_n(1)$, but its expected value converges to $\frac{1}{n}\log\gamma_n(1)$ as the size of the system N goes to infinity, with speed of convergence $\frac{1}{\sqrt{N}}$.

Proposition 3.2.2. *Let the state space S be $\mathbb{R}^d$ or a counting space. Consider the potential/kernel pair $(\widehat{G},\,\widehat{M})$ given by (2.8)-(2.9). For each $n\in\mathbb{N}$ we have the estimate*

$$\mathbb{P}\bigg(|\frac{1}{n}\log\gamma_n^N(1)-\frac{1}{n}\log\gamma_n(1)|>\delta\bigg)\;\le\;8(n+1)e^{-N\delta_n^2/2},$$

where $\delta_n := \gamma_n(1)\cdot(1-e^{-n\delta})/(n+1)$.

Remark. Observe that we don't make any particular assumption neither on $\widehat{G}$ nor $\widehat{M}$. In particular, $\inf_{x\in S}\widehat{G}(x)$ is not necessarily strictly positive and $\widehat{M}$ is not irreducible, in general.

Proof. Fix $\delta>0$ and $n,\,N\in\mathbb{N}$. We have that

$$\begin{aligned}\mathbb{P}\bigg(|\frac{1}{n}\log\gamma_n^N(1)-\frac{1}{n}\log\gamma_n(1)|>\delta\bigg) &= \mathbb{P}\bigg(\Big|\log\frac{\gamma_n^N(1)}{\gamma_n(1)}\Big|>n\delta\bigg)\\ &\le \mathbb{P}\bigg(\frac{\gamma_n^N(1)}{\gamma_n(1)}\notin[e^{-n\delta},\,e^{n\delta}]\bigg)\\ &= \mathbb{P}\bigg(\frac{\gamma_n^N(1)}{\gamma_n(1)}-1\notin[e^{-n\delta}-1,\,e^{n\delta}-1]\bigg)\\ &= \mathbb{P}\bigg(\gamma_n^N(1)-\gamma_n(1)\;\notin[\gamma_n(1)\cdot(e^{-n\delta}-1),\,\gamma_n(1)\cdot(e^{n\delta}-1)]\bigg)\\ &\le \mathbb{P}\bigg(|\gamma_n^N(1)-\gamma_n(1)|>\gamma_n(1)\cdot(1-e^{-n\delta})\bigg),\end{aligned}$$

since $e^{n\delta}-1>1-e^{-n\delta}$ for every $\delta>0$. We conclude applying Proposition 3.1.2, considering $\gamma_n(1)\cdot(1-e^{-n\delta})$ instead of δ. □

Remark. Applying Borel-Cantelli, we can see that Proposition 3.2.2 implies the strong law of large numbers for $\frac{1}{n}\log\gamma_n^N(1)$ as $N\to\infty$, that is

$$\lim_{N\to\infty}\frac{1}{n}\log\gamma_n^N(1) \;=\; \frac{1}{n}\log\gamma_n(1),$$

$\mathbb{P}$-almost surely.

Lemma 3.2.3. *Let S be finite and $\widehat{M}$ irreducible, then we have*

$$\left|\frac{1}{n}\log\gamma_n(1)-\Lambda(k)\right|\leq\frac{d}{n},$$

for every $n\in\mathbb{N}$, where d is constant and $\Lambda(k)$ is the SCGF (2.3)*.*

Proof. Consider the non-conservative homogeneous Markov chain on S defined by the Markov transitions $\widetilde{M}(x,\,y):=\widehat{G}(x)\widehat{M}(x,\,y)$ and with initial distribution μ_0. Note that we can write $\gamma_n(1)$ in terms of the kernel $\widetilde{M}$, namely

$$\begin{aligned}\gamma_n(1) &= \mathbb{E}_{\mu_0}\left[\prod_{p=0}^{n-1}\widehat{G}(X_p)\right]\\ &= \sum_{x_0\in S}\cdots\sum_{x_n\in S}\mu_0(dx_0)\prod_{p=0}^{n-1}\widehat{G}(x_p)\widehat{M}(x_p,\,dx_{p+1})\\ &= \sum_{(x,y)\in S^2}\mu_0(x)\widetilde{M}^n(x,\,y). \end{aligned} \tag{3.8}$$

Like $\widehat{M}$, also $\widetilde{M}$ is irreducible, so we can apply the Perron-Frobenius Theorem (see [11], Theorem 3.1.1) and we have that $\widetilde{M}$ possesses an eigenvalue ρ such that:

(i) ρ real, strictly positive and $|\lambda|\leq\rho$ for any eigenvalue λ of $\widetilde{M}$,

(ii) there exist left $l(\cdot)$ and right $r(\cdot)$ eigenvectors corresponding to the eigenvalue ρ that have strictly positive coordinates and are unique up to a constant multiple.

In particular, denoting $\alpha:=\sup_{x\in S}r(x)$ and $\beta:=\inf_{x\in S}r(x)$, we have that

$$\frac{1}{\alpha}\widetilde{M}^n(x,y)r(y)\leq\widetilde{M}^n(x,y)1\;\leq\;\frac{1}{\beta}\widetilde{M}^n(x,y)r(y),$$

for every $x,y\in S$.

Using the fact that $\sum_{x\in S}\widetilde{M}(x,y)r(y)=\rho\cdot r(y)$, we obtain

$$\frac{1}{\alpha}\rho^n r(y)\;\leq\;\sum_{x\in S}\widetilde{M}^n(x,y)1\;\leq\;\frac{1}{\beta}\rho^n r(y),$$

for every $y \in S$.

Therefore, combining the last result with (3.8), we obtain

$$\frac{\rho^n}{\alpha} \le \gamma_n(1) \le \frac{\rho^n}{\beta}.$$

In particular,

$$\Lambda(k) := \lim_{n\to\infty} \frac{1}{n} \log \gamma_n(1) = \rho,$$

and

$$\frac{1}{n} \log \frac{1}{\alpha} \le \frac{1}{n} \log \gamma_n(1) - \Lambda(k) \le \frac{1}{n} \log \frac{1}{\alpha}.$$

We can conclude by taking $d := \max\{|\log \frac{1}{\alpha}|, |\log \frac{1}{\beta}|\}$.

□

Combining Lemma 3.2.3 respectively with Proposition 3.2.1 and Proposition 3.2.2, we obtain the following convergence results.

Corollary 3.2.4. *Let the state space S be finite and consider the potential/kernel pair $(\widehat{G}, \widehat{M})$ given by (2.8)-(2.9) and such that $\widehat{M}$ is irreducible. For every n, $N \in \mathbb{N}$ and $p \ge 1$, we have*

$$\mathbb{E}_{\mu_0}\left[\left|\frac{1}{n} \log \gamma_n^N(1) - \Lambda(k)\right|^p \right]^{1/p} \le \frac{\alpha \lfloor p/2 \rfloor!}{\sqrt{N}} \cdot \frac{n+1}{n \cdot c^n} + \frac{d}{n},$$

for some finite constant α and with $c := \min_{x\in S} \widehat{G}(x) > 0$.

Corollary 3.2.5. *Let the state space S be finite and consider the potential/kernel pair $(\widehat{G}, \widehat{M})$ given by (2.8)-(2.9) and such that $\widehat{M}$ is irreducible. Let $\delta > 0$, $N \ge 16/\delta^2$ and $n > d/\delta$, where d is the constant given by Lemma 3.2.3. Then, we have the estimate*

$$\mathbb{P}\left(\left|\frac{1}{n} \log \gamma_n^N(1) - \Lambda(k)\right| > \delta \right) \le 8(n+1)\, e^{-N\delta_n^2/2},$$

where $\delta_n := \gamma_n(1) \cdot (1 - e^{d-n\delta})/(n+1)$.

Proof. Using Lemma 3.2.3, we can see that

$$\begin{aligned}\left|\frac{1}{n} \log \gamma_n^N(1) - \Lambda(k)\right| &\le \left|\frac{1}{n} \log \gamma_n^N(1) - \frac{1}{n} \log \gamma_n(1)\right| + \left|\frac{1}{n} \log \gamma_n(1) - \Lambda(k)\right| \\ &\le \left|\frac{1}{n} \log \gamma_n^N(1) - \frac{1}{n} \log \gamma_n(1)\right| + \frac{d}{n}.\end{aligned}$$

Therefore,

$$\mathbb{P}\Big(\big|\frac{1}{n}\log\gamma_n^N(1) - \Lambda(k)\big| > \delta\Big) \leq \mathbb{P}\Big(\big|\frac{1}{n}\log\gamma_n^N(1) - \frac{1}{n}\log\gamma_n(1)\big| > \delta - \frac{d}{n}\Big).$$

When $\delta - \frac{d}{n} > 0$, the conclusion follows applying Proposition 3.2.2 and taking $\delta - \frac{d}{n}$ instead of δ.

□

3.3 Time-Uniform Estimates

We present in this section time-uniform estimations for the quantities (3.1) and (3.2), under Assumption 3.1.3. The proofs presented here are based on the time-uniform estimations for the normalised time marginals η_n, as stated in Proposition 3.1.6 and in Corollary 3.1.7.

Proposition 3.3.1. *Under Assumption 3.1.3, for every $p, n, N \in \mathbb{N}$ we have that*

$$\frac{1}{n}\mathbb{E}_{\mu_0}\Big[\,\big|\log\gamma_n^N(1) \,-\, \log\gamma_n(1)\big|^p\Big]^{1/p} \leq \frac{2\cdot d(p)^{1/p}\cdot m}{\sqrt{N}\cdot\varepsilon^{2m+2}\cdot\epsilon^3},$$

with m, ε, ϵ given by Assumption 3.1.3 and $d(p)$ defined in (3.4).

Proof. First, recall that

$$\gamma_n(1) = \prod_{p=0}^{n-1}\eta_p(\widehat{G}), \quad \gamma_n^N(1) = \prod_{p=0}^{n-1}\eta_p^N(\widehat{G}). \tag{3.9}$$

Let $c := \inf_{x\in S}\widehat{G}(x)$ and $C := \sup_x \widehat{G}(x)$. Note that, for all $p \in \mathbb{N}$, we have

$$|\log\eta_p^N(\widehat{G}) - \log\eta_p(\widehat{G})| \leq \frac{|\eta_p^N(\widehat{G} - \eta_p(\widehat{G})|}{\min\{\eta_p^N(\widehat{G}),\, \eta_p(\widehat{G})\}} \leq \frac{|\eta_p^N(\widehat{G} - \eta_p(\widehat{G})|}{c},$$

since $0 < c \leq \eta_p(\widehat{G})$ and also $c \leq \eta_p^N(\widehat{G})$. Thus,

$$\begin{aligned}
\frac{1}{n}\,\mathbb{E}_{\mu_0}\left[\,\big|\log\gamma_n^N(1) \;-\; \log\gamma_n(1)\big|^p\right]^{1/p} &= \frac{1}{n}\,\mathbb{E}\left[\left|\sum_{p=0}^{n-1}(\log\eta_p^N(\widehat{G}) - \log\eta_p(\widehat{G}))\right|^p\right]^{1/p} \\
&\leq \frac{1}{n}\sum_{p=0}^{n-1}\mathbb{E}\left[\left|\log\eta_p^N(\widehat{G}) - \log\eta_p(\widehat{G})\right|^p\right]^{1/p} \\
&\leq \frac{1}{n}\sum_{p=0}^{n-1}\frac{1}{c}\,\mathbb{E}\left[\left|\eta_p^N(\widehat{G} - \eta_p(\widehat{G})\right|^p\right]^{1/p}.
\end{aligned}$$

Using Theorem 3.1.6 and recalling that $\varepsilon = c/C$, we obtain

$$\begin{aligned}
\frac{1}{n}\sum_{p=0}^{n-1}\frac{1}{c}\,\mathbb{E}\left[\left|\eta_p^N(\widehat{G} - \eta_p(\widehat{G})\right|^p\right]^{1/p} &\leq \frac{1}{n}\sum_{p=0}^{n-1}\frac{2\cdot d(p)^{1/p}\cdot m}{\sqrt{N}\cdot\varepsilon^{2m+2}\cdot\epsilon^3} \\
&= \frac{2\cdot d(p)^{1/p}\cdot m}{\sqrt{N}\cdot\varepsilon^{2m+2}\cdot\epsilon^3}.
\end{aligned}$$

□

In particular, Proposition 3.3.1 implies

$$\left|\mathbb{E}_{\mu_0}\big[\frac{1}{n}\log\gamma_n^N(1)\big] - \frac{1}{n}\log\gamma_n(1)\right| \leq \frac{c(1)}{\sqrt{N}} + O(\frac{1}{\sqrt{N}}).$$

Therefore, $\frac{1}{n}\log\gamma_n^N(1)$ is not an unbiased estimator for $\frac{1}{n}\log\gamma_n(1)$, but its expected value converges to $\frac{1}{n}\log\gamma_n(1)$ as the size of the system N goes to infinity, with speed of convergence $\frac{1}{\sqrt{N}}$.

Remark. The main difference between Proposition 3.3.1 and Proposition 3.2.1, is that under Assumption 3.1.3 we can obtain a time-independent estimation of the L^p-error (3.5). In particular, we are now able to swap the limits

$$\begin{aligned}
0 &= \lim_{N\to\infty}\lim_{n\to\infty}\mathbb{E}_{\mu_0}\left[|\frac{1}{n}\log\gamma_n^N(1) - \frac{1}{n}\log\gamma_n(1)|\right]^{1/p} \\
&= \lim_{n\to\infty}\lim_{N\to\infty}\mathbb{E}_{\mu_0}\left[|\frac{1}{n}\log\gamma_n^N(1) - \frac{1}{n}\log\gamma_n(1)|\right]^{1/p}.
\end{aligned}$$

Under the weaker assumptions of Proposition 3.2.1, this is possible only when $c :=$

$\inf_{x\in S}\widehat{G}(x) \geq 1$. In particular,

$$\lim_{n\to\infty} \mathbb{E}_{\mu_0}\left[|\frac{1}{n}\log\gamma_n^N(1) - \frac{1}{n}\log\gamma_n(1)|\right]^{1/p} = 0,$$

if $c > 1$, and

$$\lim_{n\to\infty} \mathbb{E}_{\mu_0}\left[|\frac{1}{n}\log\gamma_n^N(1) - \frac{1}{n}\log\gamma_n(1)|\right]^{1/p} \leq \frac{\alpha\lfloor p/2\rfloor!}{\sqrt{N}},$$

with α constant, for $c = 1$. While, when $c < 1$, the estimation given in Proposition 3.2.1 of the L^p-error diverges as n tends to ∞.

Proposition 3.3.2. *Under Assumption 3.1.3, for every $\delta > 0$ and every $n, N \in \mathbb{N}$, we have that*

$$\mathbb{P}\left(|\frac{1}{n}\log\gamma_n^N(1) - \frac{1}{n}\log\gamma_n(1)| > \delta\right) \leq \left(1 + \varepsilon(1-e^{-n\delta})\sqrt{N/2}\right)e^{-N\cdot b(\delta)^2/2},$$

where $b(\delta) := (1-e^{-n\delta})\cdot\varepsilon^{2m+2}\cdot\epsilon^3/2m$, with m, ε and ϵ given in Assumption 3.1.3.

Proof. Fix $\delta > 0$ and $n, N \in \mathbb{N}$. Using the decomposition (3.9), we have

$$\begin{aligned}
&\mathbb{P}\left(|\frac{1}{n}\log\gamma_n^N(1) - \frac{1}{n}\log\gamma_n(1)| > \delta\right) \\
&\qquad = \mathbb{P}\left(\frac{1}{n}\left|\sum_{p=0}^{n-1}\log\eta_p^N(\widehat{G}) - \log\,\eta_p(\widehat{G})\right| > \delta\right) \\
&\qquad \leq \mathbb{P}\left(\sum_{p=0}^{n-1}|\log\eta_p^N(\widehat{G}) - \log\eta_p(\widehat{G})| > n\delta\right) \\
&\qquad \leq \mathbb{P}\left(|\log\eta_n^N(\widehat{G}) - \log\eta_n(\widehat{G})| > n\delta\right). \qquad (3.10)
\end{aligned}$$

Moreover, we have that

$$\begin{aligned}
\mathbb{P}\left(|\log\eta_n^N(\widehat{G}) - \log\eta_n(\widehat{G})| \geq n\delta\right) &= \mathbb{P}\left(\frac{\eta_n^N(\widehat{G})}{\eta_n(\widehat{G})} \notin (e^{-n\delta},\, e^{n\delta})\right) \\
&= \mathbb{P}\left(\frac{\eta_n^N(\widehat{G})}{\eta_n(\widehat{G})} - 1 \notin (e^{-n\delta}-1,\, e^{n\delta}-1)\right) \\
&= \mathbb{P}\left(\eta_n^N(\widehat{G}) - \eta_n(\widehat{G}) \notin \left(\eta_n(\widehat{G})(e^{-n\delta}-1),\, \eta_n(\widehat{G})(e^{n\delta}-1)\right)\right).
\end{aligned}$$

Denoting $c := \min_x \widehat{G}(x)$ and recalling that $0 < c < \eta_n(\widehat{G})$, we can see that

$$\begin{aligned}\mathbb{P}\Big(|\log\eta_n^N(\widehat{G}) - \log\,\eta_n(\widehat{G})| \geq n\delta\Big) &\leq \mathbb{P}\Big(\eta_n^N(\widehat{G}) - \eta_n(\widehat{G}) \notin \big(c(e^{-n\delta}-1),\, c(e^{n\delta}-1)\big)\Big)\\ &\leq \mathbb{P}\Big(|\eta_n^N(\widehat{G}) - \eta_n(\widehat{G})| > c(1-e^{-n\delta})\Big),\end{aligned}$$

since $e^{n\delta} - 1 > 1 - e^{-n\delta}$ for every $\delta > 0$.

Recalling Corollary 3.1.7 and that $\varepsilon := c/C$, with $C := \sup_x \widehat{G}(x)$, we obtain

$$\mathbb{P}\Big(|\log\eta_n^N(\widehat{G}) - \log\,\eta_n(\widehat{G})| \geq n\delta\Big) \leq \Big(1 + \varepsilon(1-e^{-n\delta})\sqrt{N/2}\Big)\,e^{-Nb(\delta)^2/2}, \quad (3.11)$$

for every $\delta > 0$ and $n \in \mathbb{N}$. Combining (3.11) with (3.10), we obtain the statement.

□

Remark. Proposition 3.3.2 doesn't give us a time-homogeneous estimate of the quantity (3.6), but thanks to Assumption 3.1.3, we can now get rid of the term $\gamma_n(1)$, which appears in the estimation given by Proposition 3.2.2. Moreover, we can see that, under the assumptions of Proposition 3.3.2, we have

$$\lim_{n\to\infty} \mathbb{P}\Big(|\frac{1}{n}\log\gamma_n^N(1) - \frac{1}{n}\log\gamma_n(1)| > \delta\Big) \leq (1+\varepsilon\sqrt{N/2})\cdot e^{-N\cdot b(N)^2/2},$$

with $b(N) := (1+\varepsilon\sqrt{N/2})\cdot\varepsilon^{2m+2}\cdot\epsilon^3/2m$, independent of δ. In particular,

$$\begin{aligned}0 &= \lim_{N\to\infty}\lim_{n\to\infty} \mathbb{P}\Big(|\frac{1}{n}\log\gamma_n^N(1) - \frac{1}{n}\log\gamma_n(1)| > \delta\Big)\\ &= \lim_{n\to\infty}\lim_{N\to\infty} \mathbb{P}\Big(|\frac{1}{n}\log\gamma_n^N(1) - \frac{1}{n}\log\gamma_n(1)| > \delta\Big),\end{aligned}$$

for all $\delta > 0$. Whereas, Proposition 3.2.2 doesn't allow us to estimate the quantity 3.6 when n tends to infinity without knowing γ_n^1 in general, except for some particular cases. Let $c := \inf_{x\in S}\widehat{G}(x)$ and $C := \sup_{x\in S}\widehat{G}(x)$. Since $c^n \leq \gamma_n^1 \leq C^n$, we can say, for instance, that

$$\lim_{n\to\infty} \mathbb{P}\Big(|\frac{1}{n}\log\gamma_n^N(1) - \frac{1}{n}\log\gamma_n(1)| > \delta\Big) = 0,$$

if $c > 1$, for all $\delta > 0$, and

$$\lim_{n\to\infty} \mathbb{P}\Big(|\frac{1}{n}\log\gamma_n^N(1) - \frac{1}{n}\log\gamma_n(1)| > \delta\Big) \leq 1,$$

if $C < 1$, for every $\delta > 0$. Unfortunately, we cannot say anything when $c < 1$ and $C > 1$, without estimating $\gamma_n(1)$ first or without Assumption 3.1.3 to hold.

For passing to the estimations of the approximation errors (3.1) and (3.2), we make use of the following result, which holds for generic state spaces S, contrary to Lemma 3.2.3 used in the non-uniform case.

Proposition 3.3.3. *Under Assumption 3.1.3, the SCGF is given by*

$$\Lambda(k) = \log \eta_\infty(\widehat{G}),$$

where η_∞ is the unique Φ-invariant measure. Moreover, we have the estimate

$$\left| \frac{1}{n} \log \gamma_n(1) \; - \; \log \eta_\infty(\widehat{G}) \right| \; \leq \; \frac{4m}{n \cdot \epsilon^3 \cdot \varepsilon^{2m}}, \tag{3.12}$$

where m, ε and ϵ are defined as in Assumption 3.1.3.

Proof. The proof of the first part of the statement is given in [7], Proposition 12.4.1, p.473 (its proof makes use of Proposition 3.1.5).

For proving the estimate (3.12) we consider the estimate

$$\left| \frac{1}{n} \log \gamma_n(1) \; - \; \log \eta_\infty(\widehat{G}) \right| \; \leq \; \frac{2}{n} r \cdot \beta(\Phi)$$

proved in [7], Proposition 12.4.1, where $r := \sup_{x,y \in S} \frac{\widehat{G}(x)}{\widehat{G}(y)} = \frac{1}{\varepsilon}$ and

$$\beta(\Phi) = \sum_{n \geq 0} \beta(\Phi^n),$$

with Φ^n the nonlinear evolution semigroup defined iteratively by $\Phi^n = \Phi^{n-1} \circ \Phi$ and $\Phi^0 = \mathrm{Id}$ and with $\beta(\Phi^n)$ the Lipschitz coefficient of Φ^n, so that

$$||\Phi^n(\mu) - \Phi^n(\nu)||_{\text{tv}} \leq \beta(\Phi^n) ||\mu - \nu||_{\text{tv}},$$

where $|| \cdot ||_{\text{tv}}$ is the total variation distance in $\mathcal{P}(S)$.

Under Assumption 3.1.3, we are able to apply the estimate

$$\beta(\Phi^n) \leq \frac{2}{\epsilon \cdot \varepsilon^m} \left(1 - \epsilon^2 \cdot \varepsilon^{m-1}\right)^{\lfloor n/m \rfloor},$$

which is proved in [7], Proposition 4.3.5.

Therefore, writing each $k \in \mathbb{N}$ as $k = q \cdot m + r$, with $q := \lfloor k/m \rfloor$ (euclidean) quotient and r (euclidean) remainder, we obtain

$$\begin{aligned}
\frac{2}{n} r \cdot \beta(\Phi) &\leq \frac{2}{n \cdot \varepsilon} \sum_{k \geq 0} \frac{2}{\epsilon \cdot \varepsilon^m} \left(1 - \epsilon^2 \cdot \varepsilon^{m-1}\right)^{\lfloor k/m \rfloor} \\
&= \frac{4}{n \cdot \epsilon \cdot \varepsilon^{m+1}} \cdot \sum_{q \geq 0} \sum_{r=0}^{m-1} \left(1 - \epsilon^2 \cdot \varepsilon^{m-1}\right)^q \\
&= \frac{4}{n \cdot \epsilon \cdot \varepsilon^{m+1}} \cdot m \cdot \sum_{q \geq 0} \left(1 - \epsilon^2 \cdot \varepsilon^{m-1}\right)^q \\
&= \frac{4m}{n \cdot \epsilon \cdot \varepsilon^{m+1}} \cdot \frac{1}{\epsilon^2 \cdot \varepsilon^{m-1}} \\
&= \frac{4m}{n \cdot \epsilon^3 \cdot \varepsilon^{2m}}.
\end{aligned}$$

This concludes the proof. □

Combining Proposition 3.3.3 respectively with Proposition 3.3.1 and Proposition 3.3.2, we obtain the following convergence results.

Corollary 3.3.4. *Under Assumption 3.1.3, we have*

$$\mathbb{E}_{\mu_0}\left[\, \left|\frac{1}{n} \log \gamma_n^N(1) \;-\; \Lambda(k)\right|^p \,\right]^{1/p} \;\leq\; \frac{c(p)}{\sqrt{N}} + \frac{4m}{n \cdot \epsilon^3 \cdot \varepsilon^{2m}} + O\left(\frac{1}{\sqrt{N}}\right),$$

for every $n, N \geq 0$, with $c(p)$ defined as in Proposition 3.3.1.

Corollary 3.3.5. *Let $\delta > 0$, $n > \frac{4m}{\delta \cdot \epsilon^3 \cdot \varepsilon^{2m}}$ and $N \in \mathbb{N}$. Under Assumption 3.1.3, we have*

$$\mathbb{P}\left(\left|\frac{1}{n} \log \gamma_n^N(1) - \Lambda(k)\right| > \delta \right) \;\leq\; \left(1 + \varepsilon\left(1 - e^{\frac{4m}{\epsilon^3 \cdot \varepsilon^{2m}} - n\delta}\right)\sqrt{N/2}\right) e^{-Nb(\delta)^2/2},$$

with $b(\delta) := \left(1 - e^{\frac{4m}{\epsilon^3 \cdot \varepsilon^{2m}} - n\delta}\right) \cdot \varepsilon^{2m+2} \cdot \epsilon^3 / 2m$, with m, ε and ϵ given in Assumption 3.1.3.

To sum up the results presented in the chapter, we have provided estimations for the approximation errors (3.1) and (3.2) in the case S is finite and $\widehat{M}$ irreducible (Corollary 3.2.4 and Corollary 3.2.5) and in the case $(\widehat{G}, \widehat{M})$ satisfies Assumption 3.1.3 (Corollary 3.3.4 and Corollary 3.3.5). In both cases, the order of convergence of the L^p-error (3.1) is $O(\frac{1}{\sqrt{N}})$, while the order of convergence of the probability measure (3.2) is exponential.

We conclude recalling the examples presented in Section 2.4 for illustrating how to apply the results discussed above.

Example 4. Recall Example 1. In Section 2.4, we have shown that the potential $\widehat{G}$ is given by

$$\widehat{G}(x) \;=\; p\cdot e^{k} + (1-p)\cdot e^{-k} := K,$$

and the mutation transitions $\widehat{M}$ (2.9) are given by

$$\widehat{M}(x,x+1) = \frac{p\cdot e^{k}}{K}, \qquad \widehat{M}(x,x-1) = \frac{(1-p)\cdot e^{-k}}{K}.$$

Since $\widehat{G}$ is constant over S, by Proposition 3.2.1 we have that

$$\mathbb{E}_{\mu_0}\big[\log \gamma_n^N(1)\big] \;=\; \log \gamma_n(1).$$

This agrees with what we showed in Example 1. The same observation holds for Example 2.

Example 5. Recall Example 2 on finite state space $S = \{1,\dots,m\}$. In Section 2.4, we have shown that

$$\widehat{G}(m) = p\cdot e^{k} + (1-p), \qquad \widehat{G}(2) = (1-p)\cdot e^{k} + p,$$

and $\widehat{G}(x) = 1$ for $x \neq m$, 2. Moreover,

$$\widehat{M}(m,1) = \frac{p\cdot e^{k}}{\widehat{G}(m)}, \qquad \widehat{M}(m,m-1) = \frac{(1-p)\cdot e^{k}}{\widehat{G}(m)},$$
$$\widehat{M}(2,3) = \frac{p\cdot e^{k}}{\widehat{G}(2)}, \qquad \widehat{M}(2,1) = \frac{(1-p)e^{k}}{\widehat{G}(2)},$$

and $\widehat{M}(x,x+1) = p$ or $\widehat{M}(x,x-1) = 1-p$, otherwise. In particular, the mutation transition $\widehat{M}$ is irreducible.

If m is an odd number, then $\widehat{M}$ is aperiodic and thus Assumption 3.1.3 is satisfied and, in particular, Corollary 3.3.4 and Corollary 3.3.5 hold. Otherwise, if m is even, Assumption 3.1.3 doesn't hold anymore but we can still apply the results presented in 3.2. In particular, Corollary 3.2.4 and Corollary 3.2.5 provide (non-uniform) estimations for the approximation errors (3.1) and (3.2).

Note also that, if $k \geq 0$, $c := \inf_{x\in S} \widehat{G}(x) = 1$, whereas $c < 1$ when $k < 0$.

Thus, recalling the remarks in Section 3.3, we are able to swap the limits

$$
\begin{aligned}
0 &= \lim_{N\to\infty} \lim_{n\to\infty} \mathbb{E}_{\mu_0}\left[|\frac{1}{n}\log\gamma_n^N(1) - \frac{1}{n}\log\gamma_n(1)| \right]^{1/p} \\
&= \lim_{n\to\infty} \lim_{N\to\infty} \mathbb{E}_{\mu_0}\left[|\frac{1}{n}\log\gamma_n^N(1) - \frac{1}{n}\log\gamma_n(1)| \right]^{1/p},
\end{aligned}
$$

and

$$
\begin{aligned}
0 &= \lim_{N\to\infty} \lim_{n\to\infty} \mathbb{P}\left(|\frac{1}{n}\log\gamma_n^N(1) - \frac{1}{n}\log\gamma_n(1)| > \delta \right) \\
&= \lim_{n\to\infty} \lim_{N\to\infty} \mathbb{P}\left(|\frac{1}{n}\log\gamma_n^N(1) - \frac{1}{n}\log\gamma_n(1)| > \delta \right),
\end{aligned}
$$

for all $\delta > 0$, only if m odd or $k \geq 0$.

Bibliography

[1] Cristian Giardina, Jorge Kurchan, and Luca Peliti. Direct evaluation of large-deviation functions. *Physical review letters*, 96(12):120603, 2006.

[2] Cristian Giardina, Jorge Kurchan, Vivien Lecomte, and Julien Tailleur. Simulating rare events in dynamical processes. *Journal of statistical physics*, 145(4):787–811, 2011.

[3] Raphaël Chetrite and Hugo Touchette. Nonequilibrium markov processes conditioned on large deviations. *Annales Henri Poincaré*, 16:2005–2057, 2015.

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

[5] Esteban G Hidalgo, Takahiro Nemoto, and Vivien 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.

[6] Esteban G Hidalgo, Takahiro Nemoto, and Vivien 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.

[7] Pierre Del Moral. Feynman-kac formulae. In *Feynman-Kac Formulae*, pages 47–93. Springer, 2004.

[8] Pierre Del Moral and Laurent Miclo. Particle approximations of lyapunov exponents connected to schrödinger operators and feynman–kac semigroups. *ESAIM: Probability and Statistics*, 7:171–208, 2003.

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

[10] Joel L Lebowitz and Herbert Spohn. A gallavotti–cohen-type symmetry in the large deviation functional for stochastic dynamics. *Journal of Statistical Physics*, 95(1):333–365, 1999.

[11] Amir Dembo and Ofer Zeitouni. *Large deviations techniques and applications*, volume 38. Springer Science & Business Media, 2009.