

A central limit theorem for Fleming–Viot particle systems

Frédéric Cérou, Bernard Delyon, Arnaud Guyader, Mathias Rousset

▶ **To cite this version:**

Frédéric Cérou, Bernard Delyon, Arnaud Guyader, Mathias Rousset. A central limit theorem for Fleming–Viot particle systems. Annales de l'Institut Henri Poincaré (B) Probabilités et Statistiques, 2020, 56 (1), pp.637-666. 10.1214/19-AIHP976 . hal-02484209

HAL Id: hal-02484209
https://hal.science/hal-02484209

Submitted on 18 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Central Limit Theorem for Fleming-Viot Particle Systems

Frédéric Cérou[1]
INRIA Rennes & IRMAR, France
frederic.cerou@inria.fr

Bernard Delyon
Université Rennes 1 & IRMAR, France
bernard.delyon@univ-rennes1.fr

Arnaud Guyader
LPSM, Sorbonne Université & CERMICS, France
arnaud.guyader@upmc.fr

Mathias Rousset
INRIA Rennes & IRMAR, France
mathias.rousset@inria.fr

Abstract

Fleming-Viot type particle systems represent a classical way to approximate the distribution of a Markov process with killing, given that it is still alive at a final deterministic time. In this context, each particle evolves independently according to the law of the underlying Markov process until its killing, and then branches instantaneously from the state of another randomly chosen particle. While the consistency of this algorithm in the large population limit has been recently studied in several articles, our purpose here is to prove Central Limit Theorems under very general assumptions. For this, the key suppositions are that the particle system does not explode in finite time, and that the jump and killing times have atomless distributions. In particular, this includes the case of elliptic diffusions with hard killing.

Index Terms — Sequential Monte Carlo, Interacting particle systems, Process with killing

2010 Mathematics Subject Classification: 82C22, 82C80, 65C05, 60J25, 60K35, 60K37

[1]Corresponding author.

Contents

1 Introduction **2**

2 Main result **5**
2.1 Notation and assumptions 5
2.2 Main result 7
2.3 Some comments on the asymptotic variance 9
2.4 Connection with sequential Monte Carlo methods 11
2.5 Example: Diffusion process with hard obstacle 14
2.6 Other examples 15

3 Proofs **15**
3.1 Overview 15
3.2 Well-posedness and non-simultaneity of jumps 16
3.3 Martingale decomposition 18
3.4 Quadratic variation estimates for $\mathbb{M}$ and $\mathcal{M}$ 21
3.5 $\mathbb{L}^2$-estimate for $\gamma_T(\varphi)$ 27
3.6 Time uniform estimate for p_t 28
3.7 Approximation of the quadratic variation of $\gamma^N(Q)$ 29
3.8 Convergence of i_t^N 32
3.9 Another formulation of $\sigma_T^2(\varphi)$ 35
3.10 Martingale Central Limit Theorem 35

4 Appendix **40**
4.1 Preliminary on Feller processes 40
4.2 Stopping times and martingales 43
4.3 Proof of Lemma 3.1: (A) $\Rightarrow$ (A') 43
4.4 About Soft Killing and Carré-du-Champ operator 44
4.5 Integration rules 45

1 Introduction

Let $X = (X_t)_{t \geqslant 0}$ denote a Markov process evolving in a state space of the form $F \cup \{\partial\}$, where $\partial \notin F$ is an absorbing state. X evolves in F until it reaches ∂ and then remains trapped there forever. Let us also denote τ_∂ the associated killing time, meaning that

$$\tau_\partial := \inf\{t \geqslant 0, X_t = \partial\}.$$

Given a deterministic final time $T > 0$, we are interested both in the distribution of X_T given that it has still not been killed at time T, denoted

$$\eta_T := \mathcal{L}(X_T | \tau_\partial > T),$$

and in the probability of this event, that is

$$p_T := \mathbb{P}(\tau_\partial > T),$$

with the assumption that $p_T > 0$. We also define η_t and p_t for $0 \leqslant t < T$ accordingly. Without loss of generality, we will assume for simplicity that $\mathbb{P}(X_0 = \partial) = 0$ and $p_0 = 1$ so that $\eta_0 = \mathcal{L}(X_0)$. Let us stress that in all this paper, T is held fixed and finite.

A crude Monte Carlo method approximating these quantities consists in:

- simulating N i.i.d. random variables, also called particles in the present work,
$$X_0^1, \dots, X_0^N \overset{\text{i.i.d.}}{\sim} \eta_0,$$
- letting them evolve independently according to the dynamic of the underlying process X up to time T,
- and eventually considering the estimators
$$\hat{\eta}_T^N := \frac{\sum_{i=1}^N \mathbf{1}_{X_T^i \in F}\ \delta_{X_T^i}}{\sum_{i=1}^N \mathbf{1}_{X_T^i \in F}} \qquad \text{and} \qquad \hat{p}_T^N := \frac{\sum_{i=1}^N \mathbf{1}_{X_T^i \in F}}{N},$$
with the convention that $0/0 = 0$.

It is readily seen that these estimators are not suitable for large T, typically when $T \gg \mathbb{E}[\tau_\partial]$, since they lead to a rare event estimation problem. The typical situation of interest corresponds to the case where p_T is positive but very close to zero. In this case, one has to simulate a sample with size N of order $1/p_T$, which might be intractable in practice.

To be more specific, since the random variables X_T^i are i.i.d., we know that $N\hat{p}_T^N$ has a binomial distribution with parameters N and p_T, so that the relative variance of $\hat{p}_T^N$ is equal to $(1 - p_T)/(Np_T) \approx 1/(Np_T)$. By the Central Limit Theorem, we also have

$$\sqrt{N}(\hat{p}_T^N - p_T) \xrightarrow[N\to\infty]{d} (0, p_T(1 - p_T)). \tag{1.1}$$

Let us emphasize that, even if the probability p_T is very low, one might be interested in a precise estimation of this quantity, as well as in the estimation

of the law of the process given that it is still alive, that is η_T. For example, this is the case in sequential Monte Carlo methods for simulating and estimating rare events, in particular concerning the so-called Adaptive Multilevel Splitting algorithm which is at the origin of the work presented here. In [4], we make a connection between the latter and Fleming-Viot particle systems that allows us to deduce some asymptotic properties from the results of the present paper.

A possible way to tackle this rare event issue is to approximate the quantities at stake through a Fleming-Viot type particle system [2, 14]. Under Assumptions (A) and (B) that will be detailed below, the following process is well defined for any number of particles $N \geqslant 2$:

Definition 1.1 (Fleming-Viot particle system)**.** *The Fleming-Viot particle system $(X_t^1, \cdots, X_t^N)_{t\in[0,T]}$ is the Markov process with state space F^N defined by the following set of rules:*

- *Initialization: consider N i.i.d. particles*

$$X_0^1, \ldots, X_0^N \overset{\text{i.i.d.}}{\sim} \eta_0, \tag{1.2}$$

- *Evolution and killing: each particle evolves independently according to the law of the underlying Markov process X until one of them hits ∂ (or the final time T is reached),*
- *Branching (or rebirth, or splitting): the killed particle is taken from ∂, and is given instantaneously the state of one of the $(N-1)$ other particles (randomly uniformly chosen),*
- *and so on until final time T.*

Finally, we consider the estimators

$$\eta_T^N := \frac{1}{N}\sum_{i=1}^{N} \delta_{X_T^i} \qquad \text{and} \qquad p_T^N := \left(1 - \tfrac{1}{N}\right)^{N\mathcal{N}_T},$$

where $N\mathcal{N}_T$ is the total number of branchings of the particle system until final time T. In other words, $\mathcal{N}_T$ is the empirical mean number of branchings per particle until time T:

$$\mathcal{N}_T := \frac{1}{N}\text{card}\left\{\text{branching times} \leqslant T\right\}.$$

We also define η_t^N, p_t^N and $\mathcal{N}_t$ for $0 < t < T$ analogously.

Under very general assumptions, Villemonais [14] proves among other things that p_T^N (or equivalently $e^{-\mathcal{N}_T}$) converges in probability to p_T when N goes to infinity, and that η_T^N converges in law to η_T. In the present paper, we go one step further and establish central limit results for η_T^N and p_T^N. For this, the key assumptions are that the particle system does not explode in finite time, and that the jump and killing times have atomless distributions. In particular, it includes the case of elliptic diffusive processes killed when hitting the boundary of a given domain. As far as we know, this is the first CLT result in that case of "hard obstacle", also called "hard killing" in the literature.

The rest of the paper is organized as follows. Section 2 details our assumptions, and exposes the main results of the paper (Theorem 2.6 and Corollary 2.7). It also includes some comments on the asymptotic variance, and ends by exposing examples of applications. Section 3 is dedicated to the proof of the central limit theorem, while Section 4 gathers some technical results.

2 Main result

2.1 Notation and assumptions

For any bounded $\varphi : F \to \mathbb{R}$ and $t \in [0, T]$, we consider the unnormalized measure

$$\gamma_t(\varphi) := p_t \eta_t(\varphi) = \mathbb{E}[\varphi(X_t)\mathbf{1}_{t<\tau_\partial}],$$

with $X_0 \sim \eta_0 = \gamma_0$. Note that for any $t \in [0, T]$, one has $p_t = \mathbb{P}(\tau_\partial > t) = \gamma_t(\mathbf{1}_F)$, and recall that $p_0 = 1$ by assumption. The associated empirical approximation is then given by

$$\gamma_t^N := p_t^N \eta_t^N.$$

We remark that $\gamma_0^N = \eta_0^N$. As will become clear from Theorem 2.6 and Corollary 2.7, the study of the final unnormalized measure γ_T^N allows us to deduce very easily corresponding properties for both p_T^N and η_T^N.

For simplicity, we assume that F is a measurable subset of some reference Polish space, and that for each initial condition, X is a càdlàg process in $F \cup \{\partial\}$ satisfying the time-homogeneous Markov property, with ∂ being an absorbing state. Its transition semi-group is denoted Q, meaning that $(Q^t)_{t\geqslant 0}$ is a semi-group operator defined for any bounded measurable function $\varphi : F \to \mathbb{R}$, any $x \in F$ and any $t \geqslant 0$, by

$$Q^t\varphi(x) := \mathbb{E}[\varphi(X_t)|X_0 = x].$$

By convention, in the above, the test function φ defined on F is extended on $F \cup \{\partial\}$ by setting $\varphi(\partial) = 0$. Thus, we have $Q^t\varphi(\partial) = 0$ for all $t \geqslant 0$. This equivalently defines a sub-Markovian semi-group on F also denoted $(Q^t)_{t\geqslant 0}$. Note that just as Q^t acts on functions on the right, it can act on finite measures on the left, as any Markovian semi-group. From the Markov property, we get that $\gamma_0 Q^t = \gamma_t$, and for all $0 \leqslant t \leqslant T$,

$$\gamma_0(Q^T\varphi) = \gamma_t(Q^{T-t}\varphi) = \gamma_T(\varphi), \tag{2.1}$$

which is similar to what we have for Markovian semi-groups.

Furthermore, for any probability distribution μ on F and any bounded measurable function $\varphi : F \to \mathbb{R}$, the standard notation $\mathbb{V}_\mu(\varphi)$ stands for the variance of the random variable $\varphi(Y)$ when Y is distributed according to μ, i.e.

$$\mathbb{V}_\mu(\varphi) := \mathbb{V}(\varphi(Y)) = \mathbb{E}[\varphi(Y)^2] - \mathbb{E}[\varphi(Y)]^2 = \mu(\varphi^2) - \mu(\varphi)^2.$$

Our fundamental assumptions can now be detailed. The first one ensures that two different particles never jump nor branch at the same time.

Assumption (A). *This assumption has two parts:*

(i) For any initial condition $x \in F$, the jump times of the càdlàg Markov process $t \mapsto X_t \in F \cup \{\partial\}$ have an atomless distribution:

$$\mathbb{P}(X_{t^-} \neq X_t | X_0 = x) = 0 \qquad \forall t \geqslant 0.$$

(ii) There exists a linear subspace $\mathcal{D}$ of the space $C_b(F)$ of bounded measurable real-valued functions on F, which contains at least the indicator function $\mathbf{1}_F$, and such that for any $\varphi \in \mathcal{D}$, the mapping $(x,t) \mapsto Q^t(\varphi)(x)$ is continuous on $F \times \mathbb{R}_+$.

The following elementary result will be useful.

Lemma 2.1. *Under Assumption (A), the non-increasing mapping $t \mapsto p_t = \mathbb{P}(\tau_\partial > t)$ is continuous and strictly positive on $[0,T]$.*

Proof. In fact, this result can be deduced using either (i) or (ii) of Assumption (A). We just give a proof based on (i). For any initial condition $x \in F$, (i) implies that the killing time τ_∂ has an atomless distribution in $[0,\infty)$. Indeed, if $t = \tau_\partial$ then obviously $X_{t^-} \neq X_t$ and we conclude that this event happens with probability 0 at any deterministic time t. Note that $\tau_\partial = \infty$ may have positive probability. However, under (i), the continuity result now comes from the relation

$$p_t = \mathbb{P}(\tau_\partial > t) = \int_F \mathbb{P}(\tau_\partial > t | X_0 = x)\eta_0(dx).$$

Besides, recall that p_T is strictly positive by assumption. □

Our second assumption ensures the existence of the particle system at all time.

Assumption (B). *The particle system of Definition 1.1 is well-defined in the sense that* $\mathbb{P}(\mathcal{N}_T < \infty) = 1$.

Since $p_t^N = (1-1/N)^{N\mathcal{N}_t}$, the mapping $t \mapsto p_t^N$ is clearly non-increasing, and $\mathbb{P}(p_t^N = 0) = \mathbb{P}(\mathcal{N}_T = \infty)$, hence we get the following result.

Lemma 2.2. *Under Assumption (B), the non-increasing jump process* $t \mapsto p_t^N$ *is strictly positive on* $[0, T]$.

Remark 2.3 (About the "minimal" assumptions)**.** The main results of this paper, Theorem 2.6 and Corollary 2.7 below, are in fact established under Assumption (B) and a different Assumption (A'), detailed in Section 3.2. In a nutshell, (A') amounts to supposing that when constructing the Fleming-Viot particle system, both the jumps of some specific martingales and the killing times can not occur at the same time. In Section 4.3, we prove that (A) implies (A'). Assumption (A') is thus a weaker assumption than (A), but it is more cumbersome to describe and it might be more difficult to check in practice, hence we chose to present Assumption (A) first.

Remark 2.4 (The results of [3] are weaker)**.** In [3], we prove Theorem 2.6 and Corollary 2.7 under two different assumptions, respectively called (CC) ("carré-du-champ operator") and (SK) ("soft killing") assumptions. In section 4.4 we show that these assumptions imply (A') and (B). This means that the results of the present article encompass those of [3]. However, we refer the interested reader to [3] for further details and examples, as well as for a more elementary proof of Theorem 2.6 in a simpler context.

2.2 Main result

Recall that $(X_t^1, \dots, X_t^N)_{t \geqslant 0}$ denotes the Fleming-Viot particle system, introduced in Section 1.

Definition 2.5. *For any* $n \in \{1, \dots, N\}$ *and any* $k \geqslant 0$*, we denote by* $\tau_{n,k}$ *the* k*-th killing time of particle* n*, with the convention* $\tau_{n,0} = 0$*. Moreover, for any* $j \geqslant 0$*, we denote by* τ_j *the* j*-th killing time for the whole system of particles, with the convention* $\tau_0 = 0$.

Accordingly, the processes $\mathcal{N}_t^n := \sum_{k \geqslant 1} \mathbf{1}_{\tau_{n,k} \leqslant t}$ and

$$\mathcal{N}_t := \frac{1}{N} \sum_{n=1}^{N} \mathcal{N}_t^n = \frac{1}{N} \sum_{j \geqslant 1} \mathbf{1}_{\tau_j \leqslant t}$$

are càdlàg counting processes that correspond respectively to the number of branchings of particle n before time t, and to the total number of branchings per particle of the whole particle system before time t.

As mentioned before, we can then define the empirical measure associated to the particle system as $\eta_t^N := \frac{1}{N}\sum_{n=1}^N \delta_{X_t^n}$, while the estimate of the probability that the process is still not killed at time t is denoted $p_t^N := (1-\frac{1}{N})^{N\mathcal{N}_t}$, and the unnormalized empirical measure is defined as $\gamma_t^N := p_t^N \eta_t^N$.

As will be recalled in Proposition 3.13 and already noticed by Villemonais in [14], their large N limits are respectively

$$\eta_t(\varphi) := \mathbb{E}[\varphi(X_t)|X_t \neq \partial],\ p_t := \mathbb{P}(X_t \neq \partial),\ \text{and}\ \gamma_t(\varphi) := \mathbb{E}[\varphi(X_t)\mathbf{1}_{X_t\neq\partial}].$$

We clearly have $\eta_t(\varphi) = \gamma_t(\varphi)/\gamma_t(\mathbf{1}_F) = \gamma_t(\varphi)/p_t$ and $\gamma_t(\varphi) = \eta_0(Q^t\varphi)$. We can now expose the main result of the present paper. Recall that $\mathbb{V}_\eta(\varphi)$ denotes the variance of φ with respect to the distribution η.

Theorem 2.6. *Let us denote by $\overline{\mathcal{D}}$ the closure in $C_b(F)$ with respect to the norm $\|\cdot\|_\infty$ of the space $\mathcal{D}$ satisfying Condition (ii) of Assumption (A). Then, under Assumptions (A) and (B), for any φ in $\overline{\mathcal{D}}$, one has the convergence in distribution*

$$\sqrt{N}\left(\gamma_T^N(\varphi) - \gamma_T(\varphi)\right) \xrightarrow[N\to\infty]{d} (0, \sigma_T^2(\varphi)),$$

where $\sigma_T^2(\varphi)$ is defined by

$$\sigma_T^2(\varphi) := p_T^2\mathbb{V}_{\eta_T}(\varphi) - p_T^2\ln(p_T)\,\eta_T(\varphi)^2 - 2\int_0^T \mathbb{V}_{\eta_t}(Q^{T-t}(\varphi))p_t dp_t.$$

The variance $\sigma_T^2(\varphi)$ is linked to the asymptotic variance of the final value of the martingale $\gamma_t^N\left(Q^{T-t}(\varphi)\right)$. A sketch of the proof is given in Section 3.1.

For now, just notice that since $\mathbf{1}_F \in \mathcal{D}$ by assumption, and remembering that $\gamma_T(\mathbf{1}_F) = p_T$, the CLT for η_T^N is then a straightforward application of this result by considering the decomposition

$$\sqrt{N}\left(\eta_T^N(\varphi) - \eta_T(\varphi)\right) = \frac{1}{\gamma_T^N(\mathbf{1}_F)}\sqrt{N}\left(\gamma_T^N(\varphi - \eta_T(\varphi)) - \gamma_T(\varphi - \eta_T(\varphi))\right),$$

and the fact that $p_T^N = \gamma_T^N(\mathbf{1}_F)$ converges in probability to $p_T = \gamma_T(\mathbf{1}_F)$.

Corollary 2.7. *Under Assumptions (A) and (B), for any φ in $\overline{\mathcal{D}}$, one has the convergence in distribution*

$$\sqrt{N}\left(\eta_T^N(\varphi) - \eta_T(\varphi)\right) \xrightarrow[N\to\infty]{d} (0, \sigma_T^2(\varphi - \eta_T(\varphi))/p_T^2).$$

In addition,

$$\sqrt{N}\left(p_T^N - p_T\right) \xrightarrow[N\to\infty]{d} (0, \sigma^2),$$

where

$$\sigma^2 := \sigma_T^2(\mathbf{1}_F) = -p_T^2 \ln(p_T) - 2\int_0^T \mathbb{V}_{\eta_t}(Q^{T-t}(\mathbf{1}_F))p_t dp_t.$$

Remark 2.8 (Non independent initial conditions)**.** As will be clear from Step (i) in the proof of Proposition 3.13 and from the proof of part (a) of Proposition 3.24, Theorem 2.6 and Corollary 2.7 both still hold true when the i.i.d. assumption on the initial condition (1.2) is relaxed and replaced by the following set of conditions: (i) the initial particle system $(X_0^1, \ldots, X_0^N)$ is exchangeable, (ii) its empirical distribution $\eta_0^N = \gamma_0^N$ satisfies

$$\mathbb{E}\left[\left(\eta_0^N(Q^T(\varphi)) - \eta_0(Q^T(\varphi))\right)^2\right] \leqslant c\frac{\|\varphi\|_\infty^2}{N},$$

for some constant $c > 0$, and (iii) the following CLT is satisfied: for any $\varphi \in \mathcal{D}$,

$$\sqrt{N}\left(\eta_0^N(Q^T(\varphi)) - \eta_0(Q^T(\varphi))\right) \xrightarrow[N\to\infty]{d} (0, \mathbb{V}_{\eta_0}(Q^T(\varphi))).$$

In the next subsection, we propose to focus our attention on the estimator p_T^N in order to discuss the asymptotic variance given by Corollary 2.7.

2.3 Some comments on the asymptotic variance

Recall that

$$\sqrt{N}\left(p_T^N - p_T\right) \xrightarrow[N\to\infty]{d} (0, \sigma^2),$$

where

$$\sigma^2 = -p_T^2 \ln(p_T) - 2\int_0^T \mathbb{V}_{\eta_t}(Q^{T-t}(\mathbf{1}_F))p_t dp_t. \tag{2.2}$$

In this expression, notice that

$$\mathbb{V}_{\eta_t}(Q^{T-t}(\mathbf{1}_F)) = \mathbb{V}(\mathbb{P}(X_T \neq \partial | X_t)) = \mathbb{E}\left[\left(\mathbb{P}(X_T \neq \partial | X_t) - \frac{p_T}{p_t}\right)^2\right]. \tag{2.3}$$

Here $\mathbb{P}(X_T \neq \partial | X_t)$ is a random variable with values between 0 and 1, and expectation p_T/p_t. Hence the maximal possible value for the variance is obtained for a Bernoulli random variable with parameter p_T/p_t, so that

$$0 \leqslant \mathbb{V}_{\eta_t}(Q^{T-t}(\mathbf{1}_F)) \leqslant \frac{p_T}{p_t}\left(1 - \frac{p_T}{p_t}\right).$$

Taking into account that the integral term in (2.2) is non positive, we finally get the following bounds for the asymptotic variance of the probability estimate:

$$-p_T^2 \ln(p_T) \leqslant \sigma^2 \leqslant 2p_T(1-p_T) + p_T^2 \ln(p_T). \tag{2.4}$$

According to (2.3), the lower bound is reached when, for each $t \in [0,T]$, the probability of being still alive at time T is constant on the support of the law η_t. This situation includes, but is not limited to, the trivial case where the killing intensity is constant and equal to λ on the whole space F, meaning that for any initial condition, τ_∂ has an exponential distribution with parameter λ. Then, obviously, we get $\mathbb{V}_{\eta_t}(Q^{T-t}(\mathbf{1}_F)) = 0$. In fact, in this elementary framework, one can be much more precise about the estimator

$$p_T^N = \left(1 - \tfrac{1}{N}\right)^{N\mathcal{N}_T}.$$

Indeed, a moment thought reveals that $(N\mathcal{N}_T)_{t\geqslant 0}$ is just a Poisson process with intensity $N\lambda$, so that $N\mathcal{N}_T$ has a Poisson distribution with parameter $N\lambda T$, and p_T^N is a discrete random variable with law

$$\mathbb{P}\left(p_T^N = \left(1-\tfrac{1}{N}\right)^k\right) = e^{-N\lambda T}\frac{(N\lambda T)^k}{k!} \qquad \forall k \in \mathbb{N}.$$

In particular, it is readily seen that this estimator is unbiased:

$$\mathbb{E}[p_T^N] = e^{-\lambda T} = \mathbb{P}(X_T \neq \partial) = p_T,$$

with variance

$$\mathbb{V}(p_T^N) = (p_T)^2\left(e^{\lambda T/N} - 1\right) \implies \lim_{N\to\infty} N\mathbb{V}(p_T^N) = -p_T^2\ln(p_T),$$

which is exactly the lower bound in (2.4). Let us also notice also that in this ideal case, the asymptotic variance is much better than the one of the crude Monte Carlo estimator $\hat{p}_T^N$ as given by (1.1). Clearly, when $p_T \approx 0$,

$$-p_T^2\ln(p_T) \ll p_T(1-p_T) \approx p_T.$$

By contrast, the upper bound in (2.4) may be surprising at first sight. Indeed, by (1.1), the crude Monte Carlo estimator $\hat{p}_T^N$ satisfies

$$\sqrt{N}\left(\hat{p}_T^N - p_T\right) \xrightarrow[N\to\infty]{d} (0, p_T(1-p_T)).$$

As $2p_T(1-p_T) + p_T^2\ln(p_T) \geqslant p_T(1-p_T)$ for any $p_T \in [0,1]$, this suggests that there are some situations where the Fleming-Viot estimator is less precise than the crude Monte Carlo estimator. More precisely, if p_T is small, then

$2p_T(1-p_T)+p_T^2\ln(p_T)\approx 2p_T(1-p_T)$, that is almost twice less precise in terms of asymptotic variance.

Although counterintuitive, this phenomenon can in fact be observed on a toy example. Take $F=\{0,1\}$ for the state space, $\eta_0=p\delta_0+(1-p)\delta_1$ with $0<p<1$ for the initial distribution, $\lambda_0=0<\lambda_1$ for the killing rates, and consider the process $X_t=X_0$ until time τ_∂. In other words, nothing happens before killing and the process can be killed if and only if $X_0=1$. Suppose that our goal is to estimate the probability p_1 that the process is still alive at time $T=1$. Clearly, for all $t\geqslant 0$, one has $p_t=p+(1-p)\exp(-\lambda_1 t)$ and the law of X_t given that the process is still alive at time t writes

$$\eta_t=\frac{1}{p_t}\left(p\delta_0+(1-p)\exp(-\lambda_1 t)\delta_1\right).$$

Since, for any $t\in[0,1]$,

$$\mathbb{P}(X_1\neq\partial|X_t)=\mathbf{1}_{X_t=0}+\exp(-\lambda_1(1-t))\mathbf{1}_{X_t=1},$$

we deduce that

$$\mathbb{V}_{\eta_t}(Q^{T-t}(\mathbf{1}_F))=\frac{p+(1-p)e^{-\lambda_1(2-t)}}{p_t}-\left(\frac{p_1}{p_t}\right)^2=\frac{(p_1-p)^2}{p_t(p_t-p)}+\frac{p}{p_t}-\left(\frac{p_1}{p_t}\right)^2.$$

Therefore, taking $T=1$ in (2.2), the asymptotic variance is equal to

$$\sigma^2=2p(1-p_1)+2(p_1-p)^2\ln\frac{1-p}{p_1-p}+p_1^2\ln p_1.$$

Finally, remark that p_1 can be made arbitrarily close to p by taking λ_1 sufficiently large, which in turn leads to a variance that is arbitrarily close to the upper bound in (2.4).

Therefore, the take-home message is that we can exhibit pathological examples where the application of Fleming-Viot particle systems is in fact counterproductive compared to a crude Monte Carlo method. Intuitively, the branching process in Fleming-Viot simulation improves the focus on rare events, but creates dependency between trajectories, which may be strong.

2.4 Connection with sequential Monte Carlo methods

In this section we will compare our results with what happens with the following discrete time algorithm, which corresponds to a sequential Monte Carlo algorithm (see for example [5]). We start with a given finite set of

times $t_0 = 0 < t_1 < \cdots < t_n = T$. Let us assume to simplify that the t_j's are evenly spaced in terms of survival probability, that is $p_{t_j}/p_{t_{j-1}} = p(n)$ for all j, with $p(n) \to 1$ when $n \to \infty$.

We start with N independent copies of the process X and run them until time t_1. The ones having reached ∂ are then killed, and for each one killed, we randomly choose one that is not and duplicate it. Then we run the new and old (not killed) trajectories until time t_2, and iterate until we reach time $t_n = T$. If at some point all the trajectories are killed, i.e. if they all have reached ∂, then we consider that the run of the algorithm has failed and we call this phenomenon an extinction.

This discrete version of the algorithm falls in the framework of [5], so we can apply the results therein. Among various convergence results, we will specifically focus on CLT type theorems and compare them to our setting. Let us also mention that the extinction probability is small when N is large: specifically, there exist positive constants a and b such that the probability of extinction is less than $a\exp(-N/b)$ (Theorem 7.4.1 in [5]).

At each t_k, we denote by $\tilde{\eta}_k^N$ the empirical measure of the particles just before the resampling. We can estimate the probability $\mathbb{P}(\tau_\partial > T)$ by

$$\prod_{k=1}^{n} \tilde{\eta}_k^N(\mathbf{1}_F) = \tilde{\gamma}_n^N(\mathbf{1}_F)\tilde{\eta}_n^N(\mathbf{1}_F) \qquad \text{with} \qquad \tilde{\gamma}_n^N(\mathbf{1}_F) = \prod_{k=1}^{n-1} \tilde{\eta}_k^N(\mathbf{1}_F).$$

We also define the unnormalized measures through their action on test functions φ by $\tilde{\gamma}_n^N(\varphi) = \tilde{\gamma}_n^N(\mathbf{1}_F)\tilde{\eta}_n^N(\varphi)$. As previously, we will assume that $\varphi(\partial) = 0$, which implies that for all $t \geqslant 0$, $Q^t(\varphi)(\partial) = 0$. The following CLT is then a straightforward generalization of Theorem 12.2.2 and the following pages of [5] :

$$\sqrt{N}\left(\mathbf{1}_{\tau^N > n}\tilde{\gamma}_n^N(\varphi) - \gamma_T(\varphi)\right) \xrightarrow[N\to\infty]{d} (0, \tilde{\sigma}_n^2(\varphi)),$$

with τ^N the extinction iteration of the particle system, and $\tilde{\sigma}_n^2(\varphi) = a_n - b_n$, where

$$a_n = \eta_0((Q^T\varphi - \eta_0(Q^T\varphi))^2) + \sum_{j=1}^{n} \gamma_{t_{j-1}}(\mathbf{1}_F)^2\tilde{\eta}_{t_j}((Q^{T-t_j}\varphi - \tilde{\eta}_{t_j}(Q^{T-t_j}\varphi))^2),$$

and

$$b_n = \sum_{j=1}^{n} \gamma_{t_{j-1}}(\mathbf{1}_F)^2\tilde{\eta}_{t_{j-1}}(\mathbf{1}_F(Q^{T-t_{j-1}}\varphi - \tilde{\eta}_{t_j}(Q^{T-t_j}\varphi))^2),$$

with $\tilde{\eta}_{t_j} = p(n)\eta_{t_j} + (1-p(n))\delta_\partial$. We do not have exactly η_{t_j} because it is an updated measure, while the CLT of [5] applies to predicted measures (see [5] Sections 2.7.1 and 2.7.2 for a discussion on the difference). After some very basic algebra, this asymptotic variance can be written as

$$\begin{aligned}\tilde{\sigma}_n^2(\varphi) = \sum_{j=1}^{n} \gamma_{t_{j-1}}(\mathbf{1}_F)^2 \Big(& p(n)(\eta_{t_j}((Q^{T-t_j}\varphi)^2) - \eta_{t_{j-1}}((Q^{T-t_{j-1}}\varphi)^2)) \\ & - 2p(n)^2 \eta_{t_j}(Q^{T-t_j}\varphi)(\eta_{t_j}(Q^{T-t_j}\varphi) - \eta_{t_{j-1}}(Q^{T-t_{j-1}}\varphi)) \\ & + p(n)^2 \eta_{t_j}(Q^{T-t_j}\varphi)^2 (1 - \eta_{t_{j-1}}(Q^{t_j - t_{j-1}}\mathbf{1}_F)) \Big) \\ & + \eta_0((Q^T\varphi - \eta_0(Q^T\varphi))^2).\end{aligned}$$

Now, we should remember that $p_t = \gamma_t(\mathbf{1}_F)$, and that

$$1 - \eta_{t_{j-1}}(Q^{t_j - t_{j-1}}\mathbf{1}_F) = 1 - \frac{\gamma_{t_j}(\mathbf{1}_F)}{\gamma_{t_{j-1}}(\mathbf{1}_F)} = \frac{p_{t_{j-1}} - p_{t_j}}{p_{t_{j-1}}}.$$

If we make $n \to \infty$, which implies that $\sup_j(t_j - t_{j-1}) \to 0$, we have, at least formally, that $\tilde{\sigma}_n^2(\varphi) \to \tilde{\sigma}_\infty^2(\varphi)$ with

$$\begin{aligned}\tilde{\sigma}_\infty^2(\varphi) = & \int_0^T p_t^2 \frac{d}{dt}(\eta_t((Q^{T-t}\varphi)^2))dt - 2\int_0^T p_t^2 \eta_t(Q^{T-t}\varphi)\frac{d}{dt}(\eta_t(Q^{T-t}\varphi))dt \\ & - \int_0^T p_t \eta_t(Q^{T-t}\varphi)^2 p_t' dt + \eta_0((Q^T\varphi - \eta_0(Q^T\varphi))^2).\end{aligned}$$

By integrating by parts the first two integrals, noticing that

$$p_t p_t' \eta_t(Q^{T-t}\varphi)^2 = p_T^2 \eta_T(\varphi)^2 \frac{p_t'}{p_t}$$

in the third one, and replacing $p_t' dt$ with dp_t, we get that

$$\tilde{\sigma}_\infty^2(\varphi) = p_T^2 \mathbb{V}_{\eta_T}(\varphi) - p_T^2 \ln(p_T)\, \eta_T(\varphi)^2 - 2\int_0^T \mathbb{V}_{\eta_t}(Q^{T-t}(\varphi)) p_t dp_t,$$

which is exactly the asymptotic variance $\sigma_T^2(\varphi)$ in Theorem 2.6.

In other words, the asymptotic variance of the continuous time algorithm can be interpreted as the limit of the asymptotic variance of the discrete time algorithm, when the time mesh becomes infinitely fine, i.e. when the number of resamplings goes to infinity.

2.5 Example: Diffusion process with hard obstacle

Before proceeding with the proof of Theorem 2.6, let us give some examples of applications. We show in this section how our CLT applies to Fleming-Viot particle systems based on a Feller process killed when hitting a hard obstacle. As far as we know, this is the first CLT result in that case of "hard obstacle". Yet, there is a cluster of papers studying the hard obstacle case where the underlying process is a diffusion in a bounded domain of $\mathbb{R}^d$ killed when it hits the domain boundary. Among other questions, the convergence of the empirical measures as N goes to infinity is addressed in [1, 7, 11] (see also references therein). This case is also included in the general convergence results of [14].

Let $t \mapsto \widetilde{X}_t$ be a Feller process in a locally compact Polish space E, and let F be a bounded open domain with boundary $\partial F = \overline{F} \setminus F$. Let τ_∂ be the hitting time of $E \setminus \overline{F}$, and set

$$X_t = \begin{cases} \widetilde{X}_t & \text{if } t < \tau_\partial \\ \partial & \text{if } t \geqslant \tau_\partial \end{cases}$$

We consider the set of continuous and bounded functions $\mathcal{D} = C_b(F)$ extended as usual to $F \cup \{\partial\}$ by setting $\varphi(\partial) = 0$ if $\varphi \in \mathcal{D}$. Note that $\mathbf{1}_F \in \mathcal{D}$. The difficulty in checking Assumption (A) is the continuity with respect to t of the mapping $(x,t) \mapsto Q^t(\varphi)(x)$ because of the indicator function in

$$Q^t\varphi(x) = \mathbb{E}[\varphi(X_t)\mathbf{1}_{t<\tau_\partial}|X_0 = x].$$

However, we have the following general result:

Proposition 2.9. *Assume that F is open, that the process $\widetilde{X}$ is Feller, and the following two conditions:*

(i) For all $x \in F$ and all $t \geqslant 0$, $\mathbb{P}(\widetilde{X}_t \in \partial F | \widetilde{X}_0 = x) = 0$.

(ii) For all $x \in \partial F$, $\mathbb{P}(\tau_\partial > 0 | \widetilde{X}_0 = x) = 0$.

Then Assumption (A) is fulfilled with $\mathcal{D} = C_b(F)$.

The proof is given in Section 4.1. Using the latter, we can prove Assumption (A) for regular elliptic diffusions.

Proposition 2.10. *Assume that F is open and bounded in $\mathbb{R}^d$ with smooth boundary ∂F, and that $\widetilde{X}$ is a diffusion with smooth and uniformly elliptic coefficients. Then Assumption (A) holds true.*

Proof. This is a direct application of Proposition 2.9. First, the fact that $\widetilde{X}$ is a Feller process can be found for example in [6], Chapter 8, Theorem 1.6. Next, point (i) is obviously true because the first passage time through ∂F of an elliptic diffusion has a density with respect to Lebesgue's measure. Finally, point (ii) is also satisfied since the entrance time in the interior of a smooth domain from its boundary by an elliptic diffusion is 0. This classical fact can for example be proved by applying Itô's formula to a smooth level function defining the domain, and then the law of the iterated logarithm for the Brownian motion. □

Assumption (B) is more technical and depends on the specific case at hand. For instance, in [8], the authors prove that it is satisfied for regular diffusions and smooth boundary. Specifically, they give a general set of sufficient assumptions for non explosion, some of them being further generalized in [14]. The upcoming result is exactly Theorem 1 of Section 2.1 in [8], in the simple case of smooth domains.

Proposition 2.11. *Assume that F is open and bounded in $\mathbb{R}^d$ with smooth boundary ∂F, and that $\widetilde{X}$ is a diffusion with smooth and uniformly elliptic coefficients. Then Assumption (B) is satisfied.*

Putting all things together, we conclude that if F is open and bounded in $\mathbb{R}^d$ with smooth boundary ∂F, and if the diffusion $\widetilde{X}$ has smooth and uniformly elliptic coefficients, then one can apply the CLT results of the present paper.

2.6 Other examples

The assumptions in [3] are more restrictive than the ones required in the present paper. Therefore, we refer the reader to [3] for examples of Piecewise Deterministic Markov Processes and Diffusions with "Soft Killing" to which Theorem 2.6 and Corollary 2.7 apply as well.

3 Proofs

3.1 Overview

The key object of the proof is the càdlàg martingale

$$t \mapsto \gamma_t^N(Q) := \gamma_t^N\big(Q^{T-t}(\varphi)\big),$$

the fixed parameters T and φ being implicit in order to lighten the notation. Since $\gamma_0^N = \eta_0^N$ and $\gamma_0 = \eta_0$, the difference

$$\gamma_T^N(\varphi) - \gamma_T(\varphi) = \Big(\gamma_T^N(Q) - \gamma_0^N(Q)\Big) + \Big(\eta_0^N(Q^T(\varphi)) - \eta_0(Q^T(\varphi))\Big)$$

is the final value of the centered martingale $\gamma_t^N(Q) - \gamma_0^N(Q)$, with the addition of a second term depending on the initial condition. Note that this second term satisfies a CLT since the initial conditions $X_0^1, \ldots, X_0^N$ are i.i.d. according to Definition 1.1. We will handle the distribution of $\gamma_T^N(Q)$ in the limit $N \to \infty$ by using a Central Limit Theorem for continuous time martingales, namely Theorem 3.22. However, this requires several intermediate steps, mainly for the calculation of the quadratic variation $N[\gamma^N(Q), \gamma^N(Q)]_t$.

Unfortunately, showing the convergence of this quadratic variation is not easy. Specifically, it is much more difficult than in [3] where, thanks to the so-called "carré-du-champ operator" and "soft killing" assumptions, we could write the predictable quadratic variation as an integral against Lebesgue's measure in time, with bounded integrand. We could then easily show the pointwise convergence of the integrand and apply dominated convergence. Here we cannot do that. Instead, the key idea is to replace the quadratic variation by an adapted increasing process i_t^N such that $N[\gamma^N(Q), \gamma^N(Q)]_t - i_t^N$ is a local martingale. Finally, the convergence of i_T^N to $\sigma_T^2(\varphi) - \mathbb{V}_{\eta_0}\big(Q^T(\varphi)\big)$ requires some appropriate timewise integrations by parts formulas, as well as the uniform convergence in time of p_t^N to p_t.

Let us finally mention that, in the sequel, we will make extensive use of stochastic calculus for càdlàg semimartingales, as presented for example in [12] chapter II or [9].

3.2 Well-posedness and non-simultaneity of jumps

In the remainder, we adopt the standard notation $\Delta X_t = X_t - X_{t^-}$ and, to lighten the notation, we will denote for $l = 1, 2$,

$$\gamma_t^N\big(Q^l\big) := \gamma_t^N\Big(\big[\, Q^{T-t}(\varphi)\,\big]^l\Big). \tag{3.1}$$

Recall that X_t^n stands for the position of the particle with index n at time t. Let us fix T and φ, and denote for each $1 \leqslant n \leqslant N$ and any $t \in [0, T]$,

$$\mathbb{L}_t^n := Q^{T-t}(\varphi)(X_t^n) \qquad \text{and} \qquad \mathbb{L}_t := \frac{1}{N}\sum_{n=1}^N \mathbb{L}_t^n,$$

where the parameters T, φ and N are omitted in order to lighten the notation.

We first present a weaker but less practical version of Assumption (A), named Assumption (A'). Lemma 3.1 ensures that (A) implies (A'). All the results of the present paper are obtained under Assumption (A'). This technical assumption is in fact the minimal requirement on the non simultaneity of the

branching and jump times. In particular, Condition (A')(i) states that a single particle branches at each killing time, making the Fleming-Viot branching rule well-defined.

Assumption (A'). *There exists a space $\mathcal{D}$ of bounded measurable real-valued functions on F, which contains at least the indicator function $\mathbf{1}_F$, and such that for any $\varphi \in \mathcal{D}$, $t \mapsto \mathbb{L}_t^n$ is càdlàg for each $1 \leqslant n \leqslant N$, and:*

(i) Only one particle is killed at each branching time: if $m \neq n$, then $\tau_{m,j} \neq \tau_{n,k}$ almost surely for any $j, k \geqslant 1$.

(ii) The processes $\mathbb{L}_t^m$ and $\mathbb{L}_t^n$ never jump at the same time: if $m \neq n$, then

$$\mathbb{P}(\exists t \geqslant 0,\, \Delta\mathbb{L}_t^m \Delta\mathbb{L}_t^n \neq 0) = 0.$$

(iii) The process $\mathbb{L}_t^n$ never jumps at a branching time of another particle: if $m \neq n$, then

$$\mathbb{P}(\exists j \geqslant 0,\, \Delta\mathbb{L}_{\tau_{m,j}}^n \neq 0) = 0.$$

Recall that, by Lemma 2.1, under Condition (i) or (ii) of Assumption (A), the non-increasing mapping $t \mapsto p_t = \mathbb{P}(\tau_\partial > t)$ is continuous and strictly positive on $[0, T]$. This property still holds under Assumption (A')(i) and the proof is similar. Besides, as will be shown in Section 4.3, it turns out that (A) implies (A'). This is the purpose of the following lemma.

Lemma 3.1. *Under Assumption (A), the system of particles satisfies Assumption (A') with the same set $\mathcal{D}$ of test functions.*

Then, under Assumption (A) or (A'), it is easy to upper-bound the jumps of $\gamma_t^N(Q)$ and $\gamma_t^N(Q^2)$.

Corollary 3.2. *Under Assumption (A'), one has $|\Delta\gamma_t^N(Q)| \leqslant \frac{3\|\varphi\|_\infty}{N}$ as well as $|\Delta\gamma_t^N(Q^2)| \leqslant \frac{2\|\varphi\|_\infty^2}{N}$.*

Proof. Since $\gamma_t^N(Q) = p_t^N \mathbb{L}_t$, one has

$$|\Delta\gamma_t^N(Q)| = |\Delta(p_t^N \mathbb{L}_t)| = |p_t^N \mathbb{L}_t - p_{t^-}^N \mathbb{L}_{t^-}|.$$

First case: if $p_{t^-}^N \neq p_t^N$, this implies that $L_{t^-} \neq L_t = L_{t^-} + \Delta L_t$ and that $p_t^N = (1 - 1/N) p_{t^-}^N$, with $0 \leqslant p_{t^-}^N \leqslant 1$, hence we have

$$|\Delta\gamma_t^N(Q)| = p_{t^-}^N |(1-1/N)L_t - L_{t^-}| \leqslant |(1-1/N)L_t - L_{t^-}| \leqslant |\Delta L_t| + |L_{t^-}|/N.$$

Recall that $L_t = (\sum_{n=1}^N \mathbb{L}_t^n)/N$, and since the jumps of L_t^n and L_t^m don't coincide, we deduce that

$$|\Delta L_t| \leqslant \frac{1}{N} \max_{1\leqslant n\leqslant N} |\Delta \mathbb{L}_t^n| = \frac{1}{N} \max_{1\leqslant n\leqslant N} \left|\Delta Q^{T-t}(\varphi)(X_t^n)\right| \leqslant \frac{2\|\varphi\|_\infty}{N}.$$

Since $|L_{t^-}| \leqslant \|\varphi\|_\infty$, we finally get

$$|\Delta \gamma_t^N(Q)| \leqslant \frac{3\|\varphi\|_\infty}{N}.$$

Second case: if $p_{t^-}^N = p_t^N$ but $L_{t^-} \neq L_t$, then by the same reasoning as above,

$$|\Delta \gamma_t^N(Q)| = |\Delta(p_t^N \mathbb{L}_t)| = |p_{t^-}^N \mathbb{L}_t - p_{t^-}^N \mathbb{L}_{t^-}| \leqslant |\Delta L_t| \leqslant \frac{2\|\varphi\|_\infty}{N}.$$

The last and trivial situation is $p_{t^-}^N = p_t^N$ and $L_{t^-} = L_t$, in which case $|\Delta \gamma_t^N(Q)| = 0$. Finally, since $0 \leqslant (Q^{T-t}(\varphi)(X_t^n))^2 \leqslant \|\varphi\|_\infty^2$, one has

$$\left|\Delta\{(Q^{T-t}(\varphi)(X_t^n))^2\}\right| \leqslant \|\varphi\|_\infty^2,$$

and the same reasoning applies to $|\Delta \gamma_t^N(Q^2)|$. □

The rest of the paper is mainly devoted the proof of the following result. We recall that $\overline{\mathcal{D}}$ is the closure of $\mathcal{D}$ in $C_b(F)$ with respect to the norm $\|\cdot\|_\infty$.

Proposition 3.3. *Under Assumptions (A') and (B), for any φ in $\overline{\mathcal{D}}$, one has*

$$\sqrt{N}\left(\gamma_T^N(\varphi) - \gamma_T(\varphi)\right) \xrightarrow[N\to\infty]{d} (0, \sigma_T^2(\varphi)),$$

where

$$\sigma_T^2(\varphi) = p_T^2 \mathbb{V}_{\eta_T}(\varphi) - p_T^2 \ln(p_T)\, \eta_T(\varphi)^2 - 2\int_0^T \mathbb{V}_{\eta_t}(Q^{T-t}(\varphi)) p_t dp_t.$$

Thanks to Lemma 3.1, Proposition 3.3 implies Theorem 2.6.

3.3 Martingale decomposition

This section will build upon the martingale representation of [14]. We decompose the process $t \mapsto \gamma_t^N(Q)$ into the martingale contributions of the Markovian evolution of particle n between branchings k an $k+1$, which will be denoted $t \mapsto \mathbb{M}_t^{n,k}$, and the martingale contributions of the k-th branching of particle n, which will be denoted $t \mapsto \mathcal{M}_t^{n,k}$.

Remark 3.4. Throughout the paper, all the local martingales are local with respect to the sequence of stopping times $(\tau_j)_{j\geqslant 1}$. As required, this sequence of stopping times satisfies $\lim_{j\to\infty}\tau_j > T$ almost surely by Assumption (B).

Recall that we have defined for each $1 \leqslant n \leqslant N$ and any $t \in [0,T]$, $\mathbb{L}_t^n := Q^{T-t}(\varphi)(X_t^n)$, and $\mathbb{L}_t := \frac{1}{N}\sum_{n=1}^N \mathbb{L}_t^n$, so that

$$\gamma_t^N(Q) = \gamma_t^N(Q^{T-t}(\varphi)) = p_t^N \mathbb{L}_t.$$

If $\widetilde{X}_t$ is any particle evolving according to the dynamic of the underlying Markov process for (and only for) $t < \tau_\partial$, then it is still true that $Q^{T-t}(\varphi)(\widetilde{X}_t)\mathbf{1}_{t<\tau_\partial}$ is a martingale. As a consequence, for any $n \in \{1,\dots,N\}$ and any $k \geqslant 1$, Doob's optional sampling theorem ensures that by construction of the particle system, the process

$$\mathbb{M}_t^{n,k} := \Big(\mathbf{1}_{t<\tau_{n,k}}\mathbb{L}_t^n - \mathbb{L}_{\tau_{n,k-1}}^n\Big)\mathbf{1}_{t\geqslant\tau_{n,k-1}} = \begin{cases} 0 & \text{if } t < \tau_{n,k-1} \\ \mathbb{L}_t^n - \mathbb{L}_{\tau_{n,k-1}}^n & \text{if } \tau_{n,k-1} \leqslant t < \tau_{n,k} \\ -\mathbb{L}_{\tau_{n,k-1}}^n & \text{if } \tau_{n,k} \leqslant t \end{cases} \tag{3.2}$$

is a bounded martingale. Accordingly, under Assumption (B), the processes

$$\mathbb{M}_t^n := \sum_{k=1}^\infty \mathbb{M}_t^{n,k} = \mathbb{L}_t^n - \sum_{0\leqslant\tau_{n,k}\leqslant t} \mathbb{L}_{\tau_{n,k}}^n, \tag{3.3}$$

$$\mathbb{M}_t := \frac{1}{\sqrt{N}}\sum_{n=1}^N \mathbb{M}_t^n, \tag{3.4}$$

are local martingales.

For any $n \in \{1,\dots,N\}$ and any $k \geqslant 1$, we also consider the process

$$\mathcal{M}_t^{n,k} := \Big(1-\frac{1}{N}\Big)\Big(\mathbb{L}_{\tau_{n,k}}^n - \frac{1}{N-1}\sum_{m\neq n}\mathbb{L}_{\tau_{n,k}}^m\Big)\mathbf{1}_{t\geqslant\tau_{n,k}} = \Big(\mathbb{L}_{\tau_{n,k}}^n - \mathbb{L}_{\tau_{n,k}}\Big)\mathbf{1}_{t\geqslant\tau_{n,k}}, \tag{3.5}$$

which by Lemma 4.6 is a piecewise constant martingale with a single jump at $t = \tau_{n,k}$, and which is clearly bounded by $2\left\|\varphi\right\|_\infty$. Then, under Assumption (B), the processes

$$\mathcal{M}_t^n := \sum_{k=1}^\infty \mathcal{M}_t^{n,k} = \sum_{0\leqslant\tau_{n,k}\leqslant t} (\mathbb{L}_{\tau_{n,k}}^n - \mathbb{L}_{\tau_{n,k}}),$$

$$\mathcal{M}_t := \frac{1}{\sqrt{N}}\sum_{n=1}^N \mathcal{M}_t^n,$$

are also local martingales. Recalling the notation

$$\mathcal{N}_t^n := \sum_{k\geqslant 1} \mathbf{1}_{\tau_{n,k}\leqslant t}$$

for the number of branchings of particle n before time t and

$$\mathcal{N}_t := \frac{1}{N}\sum_{n=1}^{N} \mathcal{N}_t^n = \frac{1}{N}\sum_{j\geqslant 1} \mathbf{1}_{\tau_j\leqslant t}$$

the total number of branchings per particle before time t, (3.2) and (3.5) respectively imply that for each $1 \leqslant n \leqslant N$

$$d\mathbb{M}_t^n = d\mathbb{L}_t^n - \mathbb{L}_t^n d\mathcal{N}_t^n \tag{3.6}$$
$$d\mathcal{M}_t^n = \big(\mathbb{L}_t^n - \mathbb{L}_t\big) d\mathcal{N}_t^n, \tag{3.7}$$

so that the sum yields

$$d\mathbb{M}_t + d\mathcal{M}_t = \sqrt{N}(d\mathbb{L}_t - \mathbb{L}_t\, d\mathcal{N}_t). \tag{3.8}$$

Let us emphasize that, in the above equations, $\mathbb{L}_t = \mathbb{L}_{t^+}$ since the process $\mathbb{L}_t$ is right-continuous.

Remark 3.5. By definition, the jumps of the martingales $\mathbb{M}_t^n$ are included in the union of the set of the jumps of $\mathbb{L}_t^n$, and the set of the branching times $\tau_{n,k}$ for $k \geqslant 1$. The jumps of $\mathcal{M}_t^n$ are included in the set of the branching times $\tau_{n,k}$ for $k \geqslant 1$. Therefore, Assumption (A') implies that for $m \neq n$, the jumps of $\mathbb{M}_t^m$ and $\mathbb{M}_t^n$ can't happen at the same time, the same being true for $\mathbb{M}_t^m$ and $\mathcal{M}_t^n$.

The following rule will be useful throughout the paper.

Lemma 3.6. *Recalling that $p_t^N = (1-\frac{1}{N})^{N\mathcal{N}_t}$, it holds that*

$$dp_t^N = -p_{t^-}^N d\mathcal{N}_t. \tag{3.9}$$

Proof. When $\mathcal{N}_t \neq \mathcal{N}_{t^-}$, one has $\Delta p_t^N = (1-\frac{1}{N})^{N\mathcal{N}_t} - (1-\frac{1}{N})^{N\mathcal{N}_t-1} = (1-\frac{1}{N})^{N\mathcal{N}_t-1}\big(1-\frac{1}{N}-1\big)$ while $\Delta\mathcal{N}_t = \frac{1}{N}$ for $t=\tau_j$, $j\geqslant 1$. Hence the result follows. □

The upcoming result attests that the process $t \mapsto \gamma_t^N(Q)$ is indeed a martingale and details its decomposition.

Lemma 3.7. *We have the decomposition*

$$\gamma_t^N(Q) = \gamma_0^N(Q) + \frac{1}{\sqrt{N}} \int_0^t p_{u^-}^N (d\mathbb{M}_u + d\mathcal{M}_u). \tag{3.10}$$

Proof. Recalling that p_t^N is a piecewise constant process, one has by plain integration by parts

$$\gamma_t^N(Q) = p_t^N \mathbb{L}_t = \gamma_0^N(Q) + \int_0^t \big(p_{u^-}^N d\mathbb{L}_u + \mathbb{L}_u dp_u^N \big),$$

where we emphasize that in the above equation, the last integrand is indeed $\mathbb{L}_u = \mathbb{L}_{u^+}$. Besides, by (3.9), we are led to

$$\gamma_t^N(Q) - \gamma_0^N(Q) = \int_0^t p_{u^-}^N (d\mathbb{L}_u - \mathbb{L}_u d\mathcal{N}_u).$$

The result is then a direct consequence of (3.8). □

Remark 3.8. Since $\gamma_T(\varphi) = \gamma_0(Q^T\varphi)$, this implies the unbiasedness property $\mathbb{E}\left[\, \gamma_T^N(\varphi) \,\right] = \gamma_T(\varphi)$ for all $N \geqslant 2$. In particular, the case $\varphi = \mathbf{1}_F$ gives $\mathbb{E}\left[\, p_T^N(\varphi) \,\right] = p_T > 0$.

3.4 Quadratic variation estimates for $\mathbb{M}$ and $\mathcal{M}$

The remarkable fact is that the $2N$ martingales $\{\mathbb{M}_t^n, \mathcal{M}_t^m\}_{1 \leqslant n,m \leqslant N}$ are mutually orthogonal. We recall that two local martingales are orthogonal if their quadratic covariation is again a local martingale.

Lemma 3.9. *Under Assumptions (A') and (B), the $2N$ local martingales $\{\mathbb{M}_t^n, \mathcal{M}_t^m\}_{1 \leqslant n,m \leqslant N}$ are mutually orthogonal. In addition,*

$$[\, \mathcal{M}, \mathcal{M} \,]_t = \frac{1}{N} \sum_{n=1}^N [\, \mathcal{M}^n, \mathcal{M}^n \,]_t \,.$$

Orthogonality implies that the process $[\, \mathbb{M}, \mathcal{M} \,]_t$ is a local martingale, and denoting

$$\mathbb{A}_t := \frac{1}{N} \sum_{n=1}^N [\, \mathbb{M}^n, \mathbb{M}^n \,]_t \,,$$

that the process $[\, \mathbb{M}, \mathbb{M} \,]_t - \mathbb{A}_t$ is also a local martingale. In addition, the jumps of $\mathbb{A}$ are controlled by

$$\Delta \mathbb{A}_t \leqslant \frac{\|\varphi\|_\infty^2}{N}. \tag{3.11}$$

Proof. By Assumption (A') (see also Remark 3.5), for $n \neq m$, the piecewise constant martingales $\mathcal{M}_t^n$ and $\mathcal{M}_t^m$ do not vary at the same times, so that $[\,\mathcal{M}^n, \mathcal{M}^m\,]_t = 0$ and the two martingales are *a fortiori* orthogonal.

In the same manner, for $n \neq m$, the martingales $\mathcal{M}_t^n$ and $\mathbb{M}_t^m$ do not vary at the same times, so that $[\,\mathcal{M}^n, \mathbb{M}^m\,]_t = 0$ and the two martingales are *a fortiori* orthogonal.

Moreover, since $\mathcal{M}^n$ is a pure jump martingale, we have by definition of $\mathbb{M}_t^n$

$$d\,[\,\mathbb{M}^n, \mathcal{M}^n\,]_t = \Delta\mathbb{M}_t^n d\mathcal{M}_t^n = -\mathbb{L}_{t^-}^n d\mathcal{M}_t^n,$$

which defines a martingale, so that $\mathbb{M}^n$ and $\mathcal{M}^n$ are orthogonal.

Next, we claim that the product $\mathbb{M}^m\mathbb{M}^n$ is a martingale, implying the orthogonality. Indeed, for a given $s \in [0,T]$, let us define $\sigma_i := (\tau_i \wedge T) \vee s$ the stopping time in $[s,T]$ closest to the i-th branching time. For any $i \geqslant 1$, conditional to $\mathcal{F}_{\sigma_{i-1}}$, $(\mathbb{M}_t^n 1_{t<\sigma_i})_{t\geqslant 0}$ and $(\mathbb{M}_t^m 1_{t<\sigma_i})_{t\geqslant 0}$ are by construction independent, hence we have

$$\mathbb{E}\left[\mathbb{M}_{\sigma_i^-}^m \mathbb{M}_{\sigma_i^-}^n \,\middle|\, \mathcal{F}_{\sigma_{i-1}}\right] = \mathbb{M}_{\sigma_{i-1}}^m \mathbb{M}_{\sigma_{i-1}}^n.$$

In addition, since the martingales $\mathbb{M}^m$ and $\mathbb{M}^n$ do not jump simultaneously, it yields $\mathbb{M}_{\sigma_i}^m \mathbb{M}_{\sigma_i}^n = \mathbb{M}_{\sigma_i}^m \mathbb{M}_{\sigma_i^-}^n + \mathbb{M}_{\sigma_i^-}^m \mathbb{M}_{\sigma_i}^n - \mathbb{M}_{\sigma_i^-}^m \mathbb{M}_{\sigma_i^-}^n$, so that

$$\begin{aligned}\mathbb{E}[\mathbb{M}_{\sigma_i}^m \mathbb{M}_{\sigma_i}^n | \mathcal{F}_{\sigma_i^-}] &= \mathbb{E}[\mathbb{M}_{\sigma_i^-}^n (\mathbb{M}_{\sigma_i}^m - \mathbb{M}_{\sigma_i^-}^m) + \mathbb{M}_{\sigma_i^-}^m (\mathbb{M}_{\sigma_i}^n - \mathbb{M}_{\sigma_i^-}^n) + \mathbb{M}_{\sigma_i^-}^m \mathbb{M}_{\sigma_i^-}^n | \mathcal{F}_{\sigma_i^-}] \\ &= \mathbb{M}_{\sigma_i^-}^m \mathbb{M}_{\sigma_i^-}^n,\end{aligned}$$

and combining these equations gives

$$\mathbb{E}\left[\mathbb{M}_{\sigma_i}^m \mathbb{M}_{\sigma_i}^n \,\middle|\, \mathcal{F}_{\sigma_{i-1}}\right] = \mathbb{M}_{\sigma_{i-1}}^m \mathbb{M}_{\sigma_{i-1}}^n.$$

By iterating on $i \geqslant 1$ and taking into account that $\sigma_0 = s$ and $\lim_{i\to\infty} \sigma_i = T$, we obtain

$$\mathbb{E}\left[\mathbb{M}_T^m \mathbb{M}_T^n \middle| \mathcal{F}_s\right] = \mathbb{M}_s^m \mathbb{M}_s^n,$$

which shows the claimed result.

For the last point, Assumption (A') guarantees that

$$\Delta\mathbb{A}_t = \frac{1}{N} \max_{1\leqslant n\leqslant N} \Delta[\mathbb{M}^n, \mathbb{M}^n]_t = \frac{1}{N} \max_{1\leqslant n\leqslant N} \left(\Delta\mathbb{M}_t^n\right)^2,$$

and the indicated result is now a direct consequence of (3.2) and (3.3). □

In the same way as in (3.1), for each $t \in [0,T]$, we use in the upcoming lemma the notation

$$\mathbb{V}_{\eta_t^N}(Q) := \mathbb{V}_{\eta_t^N}(Q^{T-t}(\varphi)) = \frac{1}{N}\sum_{n=1}^{N}(\mathbb{L}_t^n)^2 - \left(\frac{1}{N}\sum_{n=1}^{N}\mathbb{L}_t^n\right)^2. \tag{3.12}$$

Lemma 3.10. *One has*

$$d\,[\,\mathcal{M},\mathcal{M}\,]_t \leqslant 4\|\varphi\|_\infty^2 d\mathcal{N}_t. \tag{3.13}$$

Moreover, there exist a piecewise constant local martingale $\widetilde{\mathcal{M}}_t$ and a piecewise constant process $\mathcal{R}_t$, both with jumps at branching times, such that

$$d\,[\,\mathcal{M},\mathcal{M}\,]_t = \mathbb{V}_{\eta_{t^-}^N}(Q)d\mathcal{N}_t + \frac{1}{N}d\mathcal{R}_t + \frac{1}{\sqrt{N}}d\widetilde{\mathcal{M}}_t, \tag{3.14}$$

with the following estimate

$$|\Delta\mathcal{R}_t| \leqslant \frac{14\,\|\varphi\|_\infty^2}{N}. \tag{3.15}$$

Proof. Considering the orthogonality property in Lemma 3.9, and taking into account that the martingales $\mathcal{M}^{n,k}$ are piecewise constant with a single jump at time $\tau_{n,k}$, we have

$$[\,\mathcal{M},\mathcal{M}\,]_t = \frac{1}{N}\sum_{n=1}^{N}\sum_{k=1}^{\infty}\Big(\mathcal{M}_{\tau_{n,k}}^{n,k}\Big)^2\mathbf{1}_{t\geqslant\tau_{n,k}}.$$

This implies (3.13) since $|\mathcal{M}_{\tau_{n,k}}^{n,k}| = |\mathbb{L}_{\tau_{n,k}}^n - \mathbb{L}_{\tau_{n,k}}| \leqslant 2\|\varphi\|_\infty$. This equation also implies that (3.14) holds true with

$$\widetilde{\mathcal{M}}_t := \frac{1}{\sqrt{N}}\sum_{n=1}^{N}\sum_{k=1}^{\infty}\Big(\big(\mathcal{M}_{\tau_{n,k}}^{n,k}\big)^2 - \mathbb{E}\left[\big(\mathcal{M}_{\tau_{n,k}}^{n,k}\big)^2\big|\mathcal{F}_{\tau_{n,k}^-}\right]\Big)\mathbf{1}_{t\geqslant\tau_{n,k}}$$

$$\mathcal{R}_t := \sum_{n=1}^{N}\sum_{k=1}^{\infty}\left(\mathbb{E}\left[\big(\mathcal{M}_{\tau_{n,k}}^{n,k}\big)^2\big|\mathcal{F}_{\tau_{n,k}^-}\right] - \mathbb{V}_{\eta_{\tau_{n,k}^-}^N}(Q)\right)\mathbf{1}_{t\geqslant\tau_{n,k}}.$$

On the one hand, Lemma 4.6 ensures that $\widetilde{\mathcal{M}}_t$ is a càdlàg local martingale. On the other hand, by Assumption (A'), we have $\mathbb{L}_{\tau_{n,k}}^l = \mathbb{L}_{\tau_{n,k}^-}^l$ for all $l \neq n$, so that (3.5) becomes

$$\mathcal{M}_{\tau_{n,k}}^{n,k} = \big(1-\tfrac{1}{N}\big)\left(\mathbb{L}_{\tau_{n,k}}^n - \tfrac{1}{N-1}\sum_{l\neq n}\mathbb{L}_{\tau_{n,k}^-}^l\right).$$

Then, by construction of the branching rule, given $\mathcal{F}_{\tau_{n,k}^-}$, $\mathbb{L}^n_{\tau_{n,k}}$ is uniformly drawn among the $(\mathbb{L}^m_{\tau_{n,k}})_{m\neq n}$, which yields

$$\mathbb{E}\left[\left(\mathcal{M}^{n,k}_{\tau_{n,k}}\right)^2\middle|\mathcal{F}_{\tau^-_{n,k}}\right] = \tfrac{1}{N-1}\sum_{m\neq n}\left(1-\tfrac{1}{N}\right)^2\left(\mathbb{L}^m_{\tau^-_{n,k}} - \tfrac{1}{N-1}\sum_{l\neq n}\mathbb{L}^l_{\tau^-_{n,k}}\right)^2.$$

If we temporarily denote the empirical distribution without particle n by

$$\eta^{(n)}_t := \frac{1}{N-1}\sum_{m\neq n}\delta_{X^m_t},$$

we can now reformulate the latter using notation (3.12) as

$$\mathbb{E}\left[\left(\mathcal{M}^{n,k}_{\tau_{n,k}}\right)^2\middle|\mathcal{F}_{\tau^-_{n,k}}\right] = \left(1-\tfrac{1}{N}\right)^2\,\mathbb{V}_{\eta^{(n)}_{\tau^-_{n,k}}}(Q).$$

In other words, we have

$$\mathcal{R}_t = \sum_{n=1}^{N}\sum_{k=1}^{\infty}\left(\left(1-\tfrac{1}{N}\right)^2\,\mathbb{V}_{\eta^{(n)}_{\tau^-_{n,k}}}(Q) - \mathbb{V}_{\eta^{N}_{\tau^-_{n,k}}}(Q)\right)\mathbf{1}_{t\geqslant\tau_{n,k}}.$$

For the last statement, notice that for two probability measures μ and ν with total variation distance $\|\mu-\nu\|_{tv}$ and for any test function f,

$$|\mathbb{V}_\mu(f)-\mathbb{V}_\nu(f)| \leqslant |(\mu-\nu)(f^2)| + |(\mu-\nu)(f)(\mu+\nu)(f)| \leqslant 6\|\mu-\nu\|_{tv}\|f\|^2_\infty$$

so that, for any n and k,

$$\begin{aligned}\left|\Delta\mathcal{R}_{\tau_{n,k}}\right| &\leqslant \left(1-\tfrac{1}{N}\right)^2\left|\mathbb{V}_{\eta^{(n)}_{\tau^-_{n,k}}}(Q) - \mathbb{V}_{\eta^{N}_{\tau^-_{n,k}}}(Q)\right| + \left(1-\left(1-\tfrac{1}{N}\right)^2\right)\mathbb{V}_{\eta^{N}_{\tau^-_{n,k}}}(Q)\\ &\leqslant 6\left(1-\tfrac{1}{N}\right)^2\left((N-1)(\tfrac{1}{N-1}-\tfrac{1}{N})+\tfrac{1}{N}\right)\|\varphi\|^2_\infty + \frac{2}{N}\|\varphi\|^2_\infty\\ &\leqslant \frac{14\,\|\varphi\|^2_\infty}{N}.\end{aligned}$$

□

Remark 3.11. A byproduct of the previous proof is the following equation, which will prove helpful in Definition 3.15 below:

$$\frac{1}{N}\mathcal{R}_t + \int_0^t \mathbb{V}_{\eta^N_{s^-}}(Q)d\mathcal{N}_s = \left(1-\tfrac{1}{N}\right)^2\frac{1}{N}\sum_{n=1}^{N}\int_0^t \mathbb{V}_{\eta^{(n)}_{s^-}}(Q)d\mathcal{N}^n_s. \tag{3.16}$$

This identity will be useful to show that the left hand side is increasing because the right hand side obviously is. This will be used to construct a suitable quadratic variation to the martingale in (3.10).

The next lemma is a very important step of the analysis. It relates the quadratic variation of the local martingale $t \mapsto \mathbb{M}_t$ - given, up to a martingale additive term, by the increasing process $t \mapsto \mathbb{A}_t$ defined in Lemma 3.7 -, with the process $t \mapsto \gamma_t^N(Q^2)$. This will yield estimates on $\mathbb{A}_t$. Note that this idea is inspired by the fact that by definition of the quadratic variation, for any Markov process X, the process $t \mapsto \left[Q^{T-t}(\varphi)(X_t) \right]^2$ equals the quadratic variation of the martingale $t \mapsto Q^{T-t}(\varphi)(X_t)$ up to a martingale additive term.

Lemma 3.12. *There exists a local martingale $(\widetilde{\mathbb{M}}_t)_{t \geqslant 0}$ such that*

$$d\gamma_t^N(Q^2) = p_{t^-}^N d\mathbb{A}_t + \frac{1}{\sqrt{N}} p_{t^-}^N d\widetilde{\mathbb{M}}_t. \tag{3.17}$$

In particular, this implies that

$$\mathbb{E}\left[\int_0^t p_{s^-}^N d\mathbb{A}_s \right] = \mathbb{E}\left[\gamma_t^N(Q^2) - \gamma_0^N(Q^2) \right] \leqslant \|\varphi\|_\infty^2 . \tag{3.18}$$

Moreover, we have

$$\mathbb{E}\left[\int_0^t p_{u^-}^N d[\widetilde{\mathbb{M}}, \widetilde{\mathbb{M}}]_u \right] \leqslant 5\|\varphi\|_\infty^4, \tag{3.19}$$

as well as

$$|\Delta\widetilde{\mathbb{M}}_u| \leqslant \frac{5\|\varphi\|_\infty^2}{\sqrt{N}}. \tag{3.20}$$

Proof. Differentiating $\gamma_t^N(Q^2) := p_t^N \frac{1}{N} \sum_{n=1}^N (\mathbb{L}_t^n)^2$ yields

$$d\gamma_t^N(Q^2) = \frac{1}{N} \sum_{n=1}^N p_{t^-}^N d\big((\mathbb{L}_t^n)^2\big) + (\mathbb{L}_t^n)^2 dp_t^N.$$

Since $dp_t^N = -p_{t^-}^N d\mathcal{N}_t$, one gets

$$d\gamma_t^N(Q^2) = \frac{1}{N} \sum_{n=1}^N p_{t^-}^N \left(d\big((\mathbb{L}_t^n)^2\big) - (\mathbb{L}_t^n)^2 d\mathcal{N}_t \right). \tag{3.21}$$

Next we claim that

$$d(\mathbb{L}_t^n)^2 - \big(\mathbb{L}_t^n\big)^2 d\mathcal{N}_t^n = d[\mathbb{M}^n, \mathbb{M}^n]_t + 2\mathbb{L}_{t^-}^n d\mathbb{M}_t^n. \tag{3.22}$$

We know from (3.6) that $d\mathbb{M}_t^n = d\mathbb{L}_t^n - \mathbb{L}_t^n d\mathcal{N}_t^n$, so that the bilinearity of the quadratic variation gives

$$\begin{aligned} d[\mathbb{M}^n, \mathbb{M}^n]_t &= d[\mathbb{L}^n, \mathbb{L}^n]_t + \big(\mathbb{L}_t^n\big)^2 d\mathcal{N}_t^n - 2d\Big[\int \mathbb{L}^n d\mathcal{N}^n, \mathbb{L}^n\Big]_t \\ &= d[\mathbb{L}^n, \mathbb{L}^n]_t + \big(\mathbb{L}_t^n\big)^2 d\mathcal{N}_t^n - 2(\Delta\mathbb{L}_t^n)\,\mathbb{L}_t^n d\mathcal{N}_t^n \\ &= d[\mathbb{L}^n, \mathbb{L}^n]_t + \mathbb{L}_t^n\big(2\mathbb{L}_{t^-}^n - \mathbb{L}_t^n\big) d\mathcal{N}_t^n. \end{aligned}$$

Then, using again (3.6) through $d\mathbb{L}_t^n = d\mathbb{M}_t^n + \mathbb{L}_t^n d\mathcal{N}_t^n$, it comes

$$\begin{aligned} d(\mathbb{L}_t^n)^2 &= 2\mathbb{L}_{t^-}^n d\mathbb{L}_t^n + d[\mathbb{L}^n, \mathbb{L}^n]_t \\ &= \Big(2\mathbb{L}_{t^-}^n d\mathbb{M}_t^n + 2\mathbb{L}_{t^-}^n \mathbb{L}_t^n d\mathcal{N}_t^n\Big) + \Big(d[\mathbb{M}^n, \mathbb{M}^n]_t - \mathbb{L}_t^n\big(2\mathbb{L}_{t^-}^n - \mathbb{L}_t^n\big) d\mathcal{N}_t^n\Big), \end{aligned}$$

which immediately simplifies into (3.22).

Putting (3.21) and (3.22) together, considering in Lemma 3.9 the definition $\mathbb{A} := \frac{1}{N}\sum_n [\mathbb{M}^n, \mathbb{M}^n]$, and recalling that $\mathcal{N} := \frac{1}{N}\sum_n \mathcal{N}^n$, we obtain

$$\begin{aligned} d\gamma_t^N(Q^2) &= p_{t^-}^N d\mathbb{A}_t + \frac{p_{t^-}^N}{N}\sum_{n=1}^N \big[\, (\mathbb{L}_t^n)^2 (d\mathcal{N}_t^n - d\mathcal{N}_t) + 2\mathbb{L}_{t^-}^n d\mathbb{M}_t^n \,\big] \\ &= p_{t^-}^N d\mathbb{A}_t + \frac{p_{t^-}^N}{N}\sum_{n=1}^N \left[\Big((\mathbb{L}_t^n)^2 - \frac{1}{N}\sum_{m=1}^N (\mathbb{L}_t^m)^2\Big) d\mathcal{N}_t^n + 2\mathbb{L}_{t^-}^n d\mathbb{M}_t^n \right], \end{aligned}$$

and we see that (3.17) is satisfied with

$$d\widetilde{\mathbb{M}}_t = \frac{1}{\sqrt{N}}\sum_{n=1}^N J_t^n d\mathcal{N}_t^n + 2\mathbb{L}_{t^-}^n d\mathbb{M}_t^n, \tag{3.23}$$

where we have defined

$$J_t^n := (\mathbb{L}_t^n)^2 - \frac{1}{N}\sum_{m=1}^N (\mathbb{L}_t^m)^2 = (1-1/N)\left((\mathbb{L}_t^n)^2 - \frac{1}{N-1}\sum_{m\neq n}(\mathbb{L}_t^m)^2\right).$$

Note that, in the same fashion as $(\mathcal{M}_t^n)_{t\geqslant 0}$, $\Big(\int_0^t J_s^n d\mathcal{N}_s^n\Big)_{t\geqslant 0}$ is a local martingale since, for each $k \geqslant 1$, $t \mapsto J_t^n \mathbf{1}_{t\geqslant \tau_{n,k}}$ is a bounded martingale by definition of the branching rule and Lemma 4.6. Moreover, in the same way as in Lemma 3.7, the $2N$ local martingales

$$\left\{ \left(\int_0^t \mathbb{L}_{s^-}^m d\mathbb{M}_s^m\right)_{t\geqslant 0}, \left(\int_0^t J_s^n d\mathcal{N}_s^n\right)_{t\geqslant 0} \right\}_{1\leqslant n,m\leqslant N}$$

are all orthogonal to each other. Indeed, for any pair n, m, (i) $[\mathbb{M}^n, \mathbb{M}^m]$ is a martingale by Lemma 3.7, (ii) $[\mathcal{M}^n, \mathbb{M}^m] = 0$ and $[\mathcal{M}^n, \mathcal{M}^m] = 0$ if $n \neq m$ by Assumption (A'). The only new point to check is that the quadratic covariation

$$d\left[\int \mathbb{L}^n_{s^-} d\mathbb{M}^n_s, \int J^n_s d\mathcal{N}^n_s\right]_t = -(\mathbb{L}^n_{t^-})^2 J^n_t d\mathcal{N}^n_t$$

is indeed a local martingale, which is a consequence of the branching rule implying $\mathbb{E}\left[J^n_{\tau_{n,k}} | \mathcal{F}_{\tau^-_{n,k}}\right] = 0$ and Lemma 4.6.

To establish (3.19) and (3.20), we recall that for any $1 \leqslant n \leqslant N$ $\sup_{t\geqslant 0} |J^n_t| \leqslant \|\varphi\|^2_\infty$, $\sup_{t\geqslant 0} |\mathbb{L}^n_{t^-}| \leqslant \|\varphi\|_\infty$, and $\sup_{t\geqslant 0} |\Delta\mathbb{M}^n_t| \leqslant 2\|\varphi\|_\infty$.

For (3.19), we apply Itô's isometry to (3.23) and use orthogonality to get

$$\begin{aligned}
\mathbb{E}\int_0^t p^N_{u^-} d[\widetilde{\mathbb{M}}, \widetilde{\mathbb{M}}]_u &= \frac{1}{N}\sum_{n=1}^N \mathbb{E}\left[\int_0^u p^N_{u^-}(J^n_u)^2 d\mathcal{N}^n_u + 4\int_0^t p^N_{u^-}(\mathbb{L}^n_{t^-})^2 d[\mathbb{M}^n, \mathbb{M}^n]_u\right] \\
&\leqslant \|\varphi\|^4_\infty \mathbb{E}\left[\int_0^t p^N_{u^-} d\mathcal{N}_u\right] + 4\|\varphi\|^2_\infty \mathbb{E}\left[\int_0^t p^N_{u^-} d\mathbb{A}_u\right] \\
&\leqslant \|\varphi\|^4_\infty \frac{1}{N}\sum_{j=1}^\infty \left(1 - \tfrac{1}{N}\right)^{j-1} + 4\|\varphi\|^2_\infty \mathbb{E}\left[\gamma^N_t(Q^2) - \gamma^N_0(Q^2)\right] \\
&\leqslant 5\|\varphi\|^4_\infty.
\end{aligned}$$

In order to obtain (3.20), consider (3.23) and recall from Assumption (A') that, for $n \neq m$, $\Delta\mathcal{N}^n_t \Delta\mathcal{N}^m_t = 0$ and $\Delta\mathbb{M}^n_t \Delta\mathbb{M}^m_t = 0$. We then deduce that

$$|\Delta\widetilde{\mathbb{M}}_t| \leqslant \frac{1}{\sqrt{N}}\left(\sup_{t\geqslant 0}|J^n_t| + 2\sup_{t\geqslant 0}|\mathbb{L}^n_{t^-}| \sup_{t\geqslant 0}|\Delta\mathbb{M}^n_t|\right) \leqslant \frac{5\|\varphi\|^2_\infty}{\sqrt{N}}.$$

□

3.5 $\mathbb{L}^2$-estimate for $\gamma_T(\varphi)$

The convergence of $\gamma^N_T(\varphi)$ to $\gamma_T(\varphi)$ when N goes to infinity is now a direct consequence of the previous results. This kind of estimate was already noticed by Villemonais in [14].

Proposition 3.13. *For any $\varphi \in \mathcal{D}$, we have*

$$\mathbb{E}\left[\left(\gamma^N_T(\varphi) - \gamma_T(\varphi)\right)^2\right] \leqslant \frac{6\,\|\varphi\|^2_\infty}{N}.$$

Proof. Thanks to Lemma 3.7 and the fact that $\gamma_T(\varphi) = \gamma_0(Q^T\varphi)$, we have the orthogonal decomposition

$$\gamma_T^N(\varphi)-\gamma_T(\varphi) = \frac{1}{\sqrt{N}}\int_0^T p_{t^-}^N\, d\mathbb{M}_t + \frac{1}{\sqrt{N}}\int_0^T p_{t^-}^N\, d\mathcal{M}_t + \gamma_0^N(Q^T\varphi)-\gamma_0(Q^T\varphi),$$

and it is easy to upper-bound the individual contribution of each term to the total variance.
(i) Initial condition. Since $\gamma_0 = \eta_0$ and $\gamma_0^N = \eta_0^N$, we have

$$\mathbb{E}\left[\,\left(\gamma_0^N(Q^T\varphi) - \gamma_0(Q^T\varphi)\right)^2\,\right] = \tfrac{1}{N}\mathbb{V}_{\eta_0}(Q^T(\varphi)(X)) \leqslant \tfrac{1}{N}\|Q^T(\varphi)\|_\infty^2 \leqslant \tfrac{1}{N}\|\varphi\|_\infty^2.$$

(ii) $\mathcal{M}$-terms. Using Itô's isometry and (3.13), we obtain

$$\begin{aligned}\mathbb{E}\left[\left(\int_0^T p_{t^-}^N d\mathcal{M}_t\right)^2\right] &= \mathbb{E}\left[\int_0^T \left(p_{t^-}^N\right)^2 d[\mathcal{M},\mathcal{M}]_t\right] \\ &\leqslant 4\|\varphi\|_\infty^2 \frac{1}{N}\sum_{j=1}^{\infty}\left(1-\tfrac{1}{N}\right)^{2(j-1)} \leqslant 4\|\varphi\|_\infty^2.\end{aligned}$$

(iii) $\mathbb{M}$-terms. In the same way, applying Itô's isometry and (3.17), we get

$$\begin{aligned}\mathbb{E}\left[\left(\int_0^T p_{t^-}^N d\mathbb{M}_t\right)^2\right] &= \mathbb{E}\left[\int_0^T \left(p_{t^-}^N\right)^2 d[\mathbb{M},\mathbb{M}]_t\right] \\ &\leqslant \mathbb{E}\left[\int_0^T p_{t^-}^N d\mathbb{A}_t\right] = \mathbb{E}\left[\gamma_T^N(Q^2)\right] \leqslant \|\varphi\|_\infty^2.\end{aligned}$$

□

In particular, Proposition 3.13 implies that for any φ in $\mathcal{D}$, $\gamma_t^N(\varphi)$ converges in probability to $\gamma_t(\varphi)$ when N goes to infinity. Since we have assumed that $\mathbf{1}_F$ belongs to $\mathcal{D}$, the probability estimate p_t^N goes to its deterministic target p_t in probability. The next subsection provides a stronger result.

3.6 Time uniform estimate for p_t

In this section, we prove the convergence of $\sup_{t\in[0,T]}\left|p_t^N - p_t\right|$ to 0 in probability by using the time marginal convergence of Proposition 3.13. Recall that, by Assumption (A) or (A'), the mapping $t \mapsto p_t$ is continuous (see Lemma 2.1). Hence, the proof only uses this argument and the monotonicity of $t \mapsto p_t^N$. One can merely see it as a Dini-like result.

Lemma 3.14. *One has*

$$\sup_{t\in[0,T]} \left|p_t^N - p_t\right| \xrightarrow[N\to\infty]{\mathbb{P}} 0.$$

Proof. Since the mapping $t \mapsto p_t$ is continuous on $[0,T]$ by Lemma 2.1, it is uniformly continuous. Hence, for any $\varepsilon > 0$, there exists a subdivision $\{t_0 = 0 < t_1 < \cdots < t_J = T\}$ such that, for any $1 \leqslant j \leqslant J$ and any t in $[t_{j-1}, t_j]$, one has

$$\max(|p_t - p_{t_{j-1}}|, |p_t - p_{t_j}|) \leqslant \varepsilon.$$

Hence, since $t \mapsto p_t^N$ is decreasing, it is readily seen that

$$|p_t^N - p_t| \leqslant \max(|p_{t_{j-1}}^N - p_t|, |p_t - p_{t_j}^N|) \leqslant \varepsilon + \max(|p_{t_{j-1}}^N - p_{t_{j-1}}|, |p_{t_j}^N - p_{t_j}|).$$

Consequently, with probability 1, uniformly in $t \in [0,T]$, we get

$$|p_t^N - p_t| \leqslant \varepsilon + \max_{0\leqslant j\leqslant J} |p_{t_j}^N - p_{t_j}|.$$

Taking $\varphi = \mathbf{1}_F$ and $T = t_j$ in Proposition 3.13 ensures that

$$\max_{0\leqslant j\leqslant J} |p_{t_j}^N - p_{t_j}| \xrightarrow[N\to\infty]{\mathbb{P}} 0.$$

Therefore, we have

$$\mathbb{P}(\sup_{t\in[0,T]} |p_t^N - p_t| > 2\varepsilon) \leqslant \mathbb{P}(\max_{0\leqslant j\leqslant J} |p_{t_j}^N - p_{t_j}| > \varepsilon) \xrightarrow[N\to\infty]{} 0.$$

Since ε is arbitrary, we get the desired result. □

3.7 Approximation of the quadratic variation of $\gamma^N(Q)$

As will become clear later, the following process represents a useful approximation of $N\left[\gamma^N(Q), \gamma^N(Q)\right]_t$.

Definition 3.15. *For each $\varphi \in \mathcal{D}$ and $T > 0$, we define for $t \in [0,T]$ the càdlàg increasing process*

$$i_t^N := \int_0^t \left(p_{u^-}^N\right)^2 d\mathbb{A}_u - \int_0^t \mathbb{V}_{\eta_{u^-}^N}(Q) p_{u^-}^N dp_u^N + \frac{1}{N}\int_0^t \left(p_{u^-}^N\right)^2 d\mathcal{R}_u. \tag{3.24}$$

The fact that this process is increasing comes from (3.16) and $dp_t^N = -p_{t^-}^N d\mathcal{N}_t$, which yields the alternative formulation

$$-\mathbb{V}_{\eta_{t^-}^N}(Q) p_{t^-}^N dp_t^N + \frac{1}{N}\left(p_{t^-}^N\right)^2 d\mathcal{R}_t = \left(p_{t^-}^N\right)^2 \frac{(1-1/N)^2}{N} \sum_{n=1}^N \mathbb{V}_{\eta_{t^-}^{(n)}}(Q) d\mathcal{N}_t^n$$

where the empirical distribution without particle n is denoted by $\eta_t^{(n)} := \frac{1}{N-1}\sum_{m\neq n}\delta_{X_t^m}$.

The estimation of i_t^N is in fact easier than the estimation of $N\left[\,\gamma^N(Q),\gamma^N(Q)\,\right]_t$ and these two increasing processes are equal up to a martingale term.

Lemma 3.16. *The process $t\mapsto N\left[\,\gamma^N(Q),\gamma^N(Q)\,\right]_t - i_t^N$ is a local martingale.*

Proof. From (3.10) and Lemma 3.9, we know that

$$N\left[\,\gamma^N(Q),\gamma^N(Q)\,\right]_t - \int_0^t \big(p_{u^-}^N\big)^2 d\mathbb{A}_u - \int_0^t \big(p_{u^-}^N\big)^2 d\left[\,\mathcal{M},\mathcal{M}\,\right]_u \tag{3.25}$$

is a local martingale. The result is then a direct consequence of (3.14). □

The next step is just a reformulation of i_t^N through an integration by parts.

Lemma 3.17. *The increasing process i_t^N can be decomposed as*

$$\begin{aligned} i_t^N = {} & p_t^N\gamma_t^N\big(Q^2\big) - \gamma_0^N\big(Q^2\big) + \left[\,\gamma_t^N(Q)\,\right]^2 \ln p_t^N - 2\int_0^t \gamma_{u^-}^N(Q^2)dp_u^N \\ & + m_t^N + \ell_t^N + O\left(\frac{1}{N}\right), \end{aligned}$$

where

$$m_t^N := -\frac{1}{\sqrt{N}}\int_0^t \big(p_{u^-}^N\big)^2 d\widetilde{\mathbb{M}}_u$$

is a local martingale, and

$$\ell_t^N := -\int_0^t \ln p_{u^-}^N d\big(\gamma_u^N(Q)\big)^2.$$

Proof. Starting from (3.24), we apply Lemma 3.12 to get

$$i_t^N = \int_0^t p_{u^-}^N d\gamma_u^N\big(Q^2\big) - \int_0^t \mathbb{V}_{\eta_{u^-}^N}(Q)p_{u^-}^N dp_u^N + m_t^N + \frac{1}{N}\int_0^t \big(p_{u^-}^N\big)^2 d\mathcal{R}_u.$$

Using (3.15), we are led to

$$\Big|\int_0^t \big(p_{u^-}^N\big)^2 d\mathcal{R}_u\Big| \leqslant \frac{14\|\varphi\|_\infty^2}{N}\sum_{i=0}^{\infty}(1-\tfrac{1}{N})^{2i} \leqslant 7\|\varphi\|_\infty^2,$$

We claim now that a first timewise integration by parts yields

$$\int_0^t p_{u^-}^N d\gamma_u^N\big(Q^2\big) = -\int_0^t \gamma_{u^-}^N\big(Q^2\big)dp_u^N + \gamma_t^N\big(Q^2\big)p_t^N - \gamma_0^N\big(Q^2\big) + O\Big(\frac{1}{N}\Big).$$

Indeed, Corollary 3.2 ensures that $|\Delta\gamma_{\tau_j}^N(Q^2)| \leqslant 2\,\|\varphi\|_\infty^2/N$, so that Condition (i) of Lemma 4.7 is satisfied with $z_t^N = \gamma_t^N(Q^2)$ and integration by parts rule (4.3) can therefore be applied.

Next, remarking that

$$\int_0^t \mathbb{V}_{\eta_{u^-}^N}(Q)p_{u^-}^N dp_u^N = \int_0^t \gamma_{u^-}^N\big(Q^2\big)dp_u^N - \int_0^t \big(\gamma_{u^-}^N(Q)\big)^2\big(p_{u^-}^N\big)^{-1}dp_u^N,$$

a second timewise integration by parts yields

$$\begin{aligned}\int_0^t \big(\gamma_{u^-}^N(Q)\big)^2\big(p_{u^-}^N\big)^{-1}dp_u^N &= \int_0^t \big(\gamma_{u^-}^N(Q)\big)^2 d\log p_u^N + O(\frac{1}{N})\\ &= \big[\,\gamma_t^N(Q)\,\big]^2\ln p_t^N - \int_0^t \ln p_{u^-}^N d\big(\gamma_u^N(Q)\big)^2 + O(\frac{1}{N}).\end{aligned}$$

Indeed, Assumption (A') also implies, for $N \geqslant 2$,

$$|\gamma_{\tau_j^-}^N(Q)| \leqslant 2\,\|\varphi\|_\infty\,(1-1/N)^j \quad \text{and} \quad |\Delta\gamma_{\tau_j}^N(Q)| \leqslant \frac{6\,\|\varphi\|_\infty}{N}(1-1/N)^{j-1}.$$

As a consequence, Conditions (ii) and (iii) of Lemma 4.7 are satisfied for $z_t^N = \big(\gamma_t^N(Q)\big)^2$, so that we can apply successively rules (4.4) and (4.5) of Lemma 4.7. Finally, putting all estimates together gives the desired result. □

Lemma 3.18. *One has* $\mathbb{E}\,\big[\,\big(m_t^N\big)^2\,\big] = O(1/N)$ *as well as* $\mathbb{E}\big|\ell_t^N\big| = O\Big(1/\sqrt{N}\Big)$.

Proof. The first assertion is an immediate consequence of Itô's isometry for martingales, together with (3.19). For the second one, Itô's formula yields

$$\ell_t^N = -2\int_0^t \ln p_{u^-}^N \gamma_{u^-}^N(Q)d\gamma_u^N(Q) - \int_0^t \ln p_{u^-}^N d\,\big[\,\gamma^N(Q),\gamma^N(Q)\,\big]_u.$$

Therefore,

$$\mathbb{E}\,\big|\ell_t^N\big| \leqslant 2\mathbb{E}\left[\left|\int_0^t \ln p_{u^-}^N \gamma_{u^-}^N(Q)d\gamma_u^N(Q)\right|\right] + \mathbb{E}\left[\int_0^t |\ln p_{u^-}^N| d\,\big[\,\gamma^N(Q),\gamma^N(Q)\,\big]_u\right].$$

Then, Cauchy-Schwarz inequality and Itô's isometry provide

$$\mathbb{E}\left|\ell_t^N\right| \leqslant 2\left(\mathbb{E}\left[\int_0^t \left(\ln p_{u^-}^N \gamma_{u^-}^N(Q)\right)^2 d\left[\,\gamma^N(Q),\gamma^N(Q)\,\right]_u\right]\right)^{1/2} + \mathbb{E}\left[\int_0^t |\ln p_{u^-}^N| d\left[\,\gamma^N(Q),\gamma^N(Q)\,\right]_u\right].$$

Since $p^2|\ln p| \leqslant 1$ for any $p \in (0,1]$, we have

$$\left(\ln p_{u^-}^N \gamma_{u^-}^N(Q)\right)^2 = \left(\ln p_{u^-}^N \times p_{u^-}^N \eta_{u^-}^N(Q)\right)^2 \leqslant |\ln p_{u^-}^N| \times \|\varphi\|_\infty^2.$$

Hence, if we denote

$$c(N) := \mathbb{E}\left[\int_0^T \left|\ln p_{u^-}^N\right| d\left[\,\gamma^N(Q),\gamma^N(Q)\,\right]_u\right],$$

it comes

$$\mathbb{E}\left|\ell_t^N\right| \leqslant 2\|\varphi\|_\infty \sqrt{c(N)} + c(N).$$

Next, the basic decomposition of Lemma 3.7 yields

$$d\left[\,\gamma^N(Q),\gamma^N(Q)\,\right]_t = \frac{1}{N}\left(p_{t^-}^N\right)^2 d\left[\,\mathbb{M},\mathbb{M}\,\right]_t + \frac{1}{N}\left(p_{t^-}^N\right)^2 d\left[\,\mathcal{M},\mathcal{M}\,\right]_t,$$

so that the orthogonality property of Lemma 3.9 allows us to reformulate $c(N)$ as

$$c(N) = \frac{1}{N}\mathbb{E}\left[\int_0^T \left|\ln p_{u^-}^N\right| \left(p_{u^-}^N\right)^2 (d\mathbb{A}_u + d\left[\,\mathcal{M},\mathcal{M}\,\right]_u)\right].$$

Using the fact that $p|\ln p| \leqslant 1$ together with (3.13), this yields

$$c(N) \leqslant \frac{1}{N}\mathbb{E}\left[\int_0^T p_{t^-}^N d\mathbb{A}_t\right] + \frac{4}{N^2}\|\varphi\|_\infty^2 \sum_{j\geqslant 1}(1-\tfrac{1}{N})^j.$$

Finally (3.18) gives $c(N) \leqslant \frac{5}{N}\|\varphi\|_\infty^2$ and the proof is complete. □

3.8 Convergence of i_t^N

For the forthcoming calculations we recall that

$$\begin{aligned} p_t \mathbb{V}_{\eta_t}\left(Q^{T-t}(\varphi)\right) &= \gamma_t\left[\left(Q^{T-t}(\varphi)\right)^2\right] - p_t^{-1}\left[\gamma_t\left(Q^{T-t}(\varphi)\right)\right]^2 \\ &= \gamma_t\left[\left(Q^{T-t}(\varphi)\right)^2\right] - p_t^{-1}(\gamma_T(\varphi))^2 \\ &= \gamma_t\left(Q^2\right) - p_t^{-1}(\gamma_T(\varphi))^2. \end{aligned}$$

with the notation $\gamma_t(Q^2) := \gamma_t\left[\left(Q^{T-t}(\varphi)\right)^2\right]$. In the same way, we write $\gamma_t(Q) := \gamma_t\left[\left(Q^{T-t}(\varphi)\right)\right]$, and the asymptotic variance formula will be denoted as follows:

Definition 3.19. *For any $t \in [0,T]$ and any $\varphi \in \mathcal{D}$, let us define*

$$i_t(\varphi) := p_t\gamma_t(Q^2) - \gamma_0(Q^2) + [\,\gamma_t(Q)\,]^2 \ln p_t - 2\int_0^t \gamma_u\big(Q^2\big)dp_u. \tag{3.26}$$

Our next purpose is to show that $i_t(\varphi)$ corresponds to the asymptotic variance of interest, as suggested by Lemma 3.17.

Proposition 3.20. *For any $t \in [0,T]$, one has*

$$i_t^N \xrightarrow[N\to\infty]{\mathbb{P}} i_t(\varphi).$$

Proof. We recall that, when X is a time-homogeneous Markov process, the process $t \mapsto Q^{T-t}(\varphi)(X_t)$ is a càdlàg martingale on $[0,T]$ with respect to the natural filtration of X, hence $\gamma_t\big(Q^{T-t}(\varphi)\big) = \gamma_t(Q) = \gamma_T(\varphi)$. Therefore, by Lemma 3.17, we can write

$$\begin{aligned} i_t^N - i_t(\varphi) =& \Big(p_t^N\gamma_t^N\big(Q^2\big) - p_t\gamma_t\big(Q^2\big)\Big) - \Big(\gamma_0^N\big(Q^2\big) - \gamma_0\big(Q^2\big)\Big) \\ &+ \Big(\gamma_t^N(Q)^2\ln p_t^N - \gamma_t(Q)^2\ln p_t\Big) + m_t^N + \ell_t^N + O(1/N) \\ &- 2\left(\int_0^t \gamma_{u^-}^N(Q^2)dp_u^N - \int_0^t \gamma_u(Q^2)dp_u\right). \end{aligned}$$

Clearly, by Proposition 3.13 and Lemma 3.18, the boundary terms and the rest terms all tend to 0 in probability. So we just have to show that

$$\int_0^t \gamma_{u^-}^N(Q^2)dp_u^N - \int_0^t \gamma_u(Q^2)dp_u = a_t^N + b_t^N$$

goes to 0 as well, where we have defined

$$a_t^N := \int_0^t \gamma_{u^-}^N(Q^2)d(p_u^N - p_u) \qquad \text{and} \qquad b_t^N := \int_0^t \big(\gamma_{u^-}^N(Q^2) - \gamma_u(Q^2)\big)dp_u.$$

To prove that $b_t^N \xrightarrow{\mathbb{P}} 0$, notice that

$$\big|b_t^N\big| \leqslant \int_0^t \big|\gamma_{u^-}^N(Q^2) - \gamma_u^N(Q^2)\big|\,dp_u + \int_0^t \big|\gamma_u^N(Q^2) - \gamma_u(Q^2)\big|\,dp_u.$$

Corollary 3.2 ensures that

$$\int_0^t \big|\gamma_{u^-}^N(Q^2) - \gamma_u^N(Q^2)\big|\,dp_u \leqslant \frac{2\|\varphi\|_\infty^2}{N} \xrightarrow[N\to\infty]{} 0.$$

Additionally, Proposition 3.13 and Lebesgue's dominated convergence theorem yield

$$\mathbb{E}\left[\int_0^t \left|\gamma_u^N(Q^2) - \gamma_u(Q^2)\right| dp_u\right] = \int_0^t \mathbb{E}\left[\left|\gamma_u^N(Q^2) - \gamma_u(Q^2)\right|\right] dp_u \xrightarrow[N\to\infty]{} 0,$$

which implies that

$$\int_0^t \left|\left(\gamma_u^N(Q^2) - \gamma_u(Q^2)\right)\right| dp_u \xrightarrow[N\to\infty]{\mathbb{P}} 0.$$

The proof of $a_t^N \xrightarrow{\mathbb{P}} 0$ requires more attention. Since $|\Delta\gamma_{\tau_j}^N(Q^2)| \leqslant 2\|\varphi\|_\infty^2/N$ by Corollary 3.2, the timewise integration by parts rule (4.3) of Lemma 4.7 enables us to rewrite the first term as

$$a_t^N = -\int_0^t (p_{u^-}^N - p_u)\, d\,\gamma_u^N(Q^2) + \gamma_t^N(Q^2)(p_t^N - p_t) + O(1/N),$$

where we have used that $p_0^N = p_0 = 1$. Since $\gamma_t^N(Q^2)$ is bounded, the boundary term goes to 0 by Proposition 3.13. For the integral term, equation (3.17) leads to the decomposition

$$\int_0^t (p_{u^-}^N - p_u)\, d\,\gamma_u^N(Q^2) = \int_0^t (p_{u^-}^N - p_u) p_{u^-}^N d\mathbb{A}_u + \frac{1}{\sqrt{N}} \int_0^t (p_{u^-}^N - p_u) p_{u^-}^N d\widetilde{\mathbb{M}}_u. \tag{3.27}$$

Since $\mathbb{A}$ is an increasing process, it comes

$$\left|\int_0^t (p_{u^-}^N - p_u) p_{u^-}^N d\mathbb{A}_u\right| \leqslant \sup_u |p_{u^-}^N - p_u| \times \left(\int_0^t p_{u^-}^N d\mathbb{A}_u\right). \tag{3.28}$$

The supremum term goes to 0 in probability by Lemma 3.14 and, by (3.17),

$$\mathbb{E}\left[\int_0^t p_{u^-}^N d\mathbb{A}_u\right] = \mathbb{E}\left[\gamma_t^N(Q^2)\right] \leqslant \|\varphi\|_\infty^2.$$

So the right hand side of (3.28) is the product of an $\mathrm{o}_{\mathbb{P}}(1)$ with an $\mathrm{O}_{\mathbb{P}}(1)$, which is classically an $\mathrm{o}_{\mathbb{P}}(1)$ (see for example [13], Theorem 7.15, for a general version of this result), and the first term of (3.27) goes to zero in probability.

For the second term in (3.27), just notice that $\left|p_{u^-}^N - p_u\right| p_{u^-}^N \leqslant 1$, so that Itô's isometry and (3.19) yield

$$\mathbb{E}\left[\left(\int_0^t (p_{u^-}^N - p_u) p_{u^-}^N d\widetilde{\mathbb{M}}_u\right)^2\right] = \mathbb{E}\left[\int_0^t (p_{u^-}^N - p_u)^2 \left(p_{u^-}^N\right)^2 d[\widetilde{\mathbb{M}}, \widetilde{\mathbb{M}}]_u\right] \leqslant 5\|\varphi\|_\infty^4$$

and a_t^N tends to zero in probability as well. □

3.9 Another formulation of $\sigma_T^2(\varphi)$

In order to retrieve the expression of Theorem 2.6, one can then simplify the variance at final time T as follows.

Lemma 3.21. *Define*

$$\sigma_T^2(\varphi) := \mathbb{V}_{\eta_0}\big(Q^T(\varphi)\big) + i_T(\varphi),$$

with $i_T(\varphi)$ like in (3.26)*, then*

$$\sigma_T^2(\varphi) = p_T^2 \mathbb{V}_{\eta_T}(\varphi) - p_T^2 \ln(p_T)\, \eta_T(\varphi)^2 - 2\int_0^T \mathbb{V}_{\eta_t}(Q^{T-t}(\varphi)) p_t dp_t. \quad (3.29)$$

Proof. Since $\gamma_T(Q^2) = \gamma_T(\varphi^2)$,

$$i_T = p_T\gamma_T(\varphi^2) - \gamma_0(Q^2) + \gamma_T(\varphi)^2 \ln p_T - 2\int_0^T \gamma_t\big(Q^2\big) dp_t. \quad (3.30)$$

Furthermore, by definition,

$$p_t^{-1}\gamma_t\big(Q^2\big) = \eta_t(Q^{T-t}(\varphi)^2) = \mathbb{V}_{\eta_t}(Q^{T-t}(\varphi)) + p_t^{-2}\gamma_t(Q^{T-t}(\varphi))^2. \quad (3.31)$$

Recall that from equation (2.1), $\gamma_t(Q^{T-t}(\varphi)) = \gamma_T(\varphi)$, so that reporting the latter identity into (3.31), and then (3.31) into (3.30) gives

$$i_T = p_T\gamma_T(\varphi^2) - \gamma_0(Q^2) - \gamma_T(\varphi)^2 \ln p_T - 2\int_0^T \mathbb{V}_{\eta_t}(Q^{T-t}(\varphi)) p_t dp_t.$$

In the same way, $p_T\gamma_T(\varphi^2) = p_T^2\mathbb{V}_{\eta_T}(\varphi) + \gamma_T(\varphi)^2$ and $\mathbb{V}_{\eta_0}\big(Q^T(\varphi)\big) = \gamma_0(Q^2) - (\gamma_T(\varphi))^2$, hence the result holds true. □

3.10 Martingale Central Limit Theorem

The following result is an adaptation of Theorem 1.4 page 339 in [6] to our specific context. The main difference is about the initial condition.

Theorem 3.22. *On a filtered probability space, let $t \mapsto z_t^N$ denote a sequence of càdlàg local martingales indexed by $N \geqslant 1$. Assume moreover that*

(i) $z_0^N \xrightarrow[N\to+\infty]{\mathcal{D}} \mu_0$*, where μ_0 is a given probability on $\mathbb{R}$.*

(ii) One has $\lim_{N\to\infty} \mathbb{E}[\sup_{t\in[0,T]} \big|\Delta z_t^N\big|^2] = 0$.

(iii) *For each N, there exists an increasing càdlàg process $t \mapsto i_t^N$ such that $t \mapsto \left(z_t^N - z_0^N\right)^2 - i_t^N$ is a local martingale.*

(iv) *The process $t \mapsto i_t^N$ satisfies $\lim_{N\to\infty} \mathbb{E}\left[\sup_{t\in[0,T]} \Delta i_t^N\right] = 0$.*

(v) *There exists a continuous and increasing deterministic function $t \mapsto i_t$ such that, for all $t \in [0,T]$,*

$$i_t^N \xrightarrow[N\to+\infty]{\mathbb{P}} i_t.$$

Then $(z_t^N)_{t\in[0,T]}$ converges in law (under the Skorokhod topology) to $(Z_t)_{t\in[0,T]}$, where $Z_0 \sim \mu_0$ and $(Z_t - Z_0)_{t\in[0,T]}$ is a Gaussian process, independent of Z_0, with independent increments and variance function i_t.

Proof. First, we notice that Theorem 1.4 with condition (b) in [6] is exactly the present result in the special case where $z_0^N = 0$ and $\mu_0 = \delta_0$. See also Section 5, Chapter 7 of [9] in which, again, the case of a general initial condition is left to the reader.

Second, fix $\psi \in C_b(\mathbb{R})$, and consider the $\mathbb{P}$-absolutely continuous probability defined by

$$\mathbb{P}_\psi = \frac{1}{\mathbb{E}\left[\, e^{\psi(z_0^N)} \,\right]} e^{\psi(z_0^N)}\ \mathbb{P}.$$

We next give the argument that covers the case of a general initial condition. For any ψ, we claim that under $\mathbb{P}_\psi$ with the same filtration, all the assumptions of the present theorem hold for $t \mapsto z_t^N - z_0^N$ instead of $t \mapsto z_t^N$.

By construction, the equivalent probabilities $\mathbb{P}$ and $\mathbb{P}_\psi$ have a likelihood ratio $d\mathbb{P}/d\mathbb{P}_\psi$ that is (a) bounded above and away from 0, and (b) measurable with respect to the initial σ-field (i.e., time 0). Condition (b) immediately implies that any martingale property under $\mathbb{P}$ still holds true under $\mathbb{P}_\psi$, the filtration being kept the same. The processes $t \mapsto z_t^N - z_0^N$ are thus local martingales under $\mathbb{P}_\psi$, with the same localizing stopping times. Condition (a) implies that the upper bound on jumps (ii) is still satisfied under $\mathbb{P}_\psi$. In addition, (a) implies that the process $t \mapsto \left(z_t^N - z_0^N\right)^2 - i_t^N$ is still a local martingale and thus (iii) holds true. Condition (a) implies also that the upper bound on jumps (iv) is satisfied. Finally, (a) implies that convergences in probability under $\mathbb{P}$ or $\mathbb{P}_\psi$ are equivalent, so that (v) is verified.

As a consequence, under each $\mathbb{P}_\psi$ with bounded ψ, the process $t \mapsto z_t^N - z_0^N$ converges in law under the Skorokhod topology to $(M_t)_{t\in[0,T]}$, a Gaussian martingale with initial value $M_0 = 0$ and variance function i_t.

Finally, let F be a continuous functional on the Skorokhod space of càdlàg paths, and ψ a continuous bounded test function. Using the previous reasoning and assumption (i), we have that

$$\mathbb{E}\left[\, e^{\psi(z_0^N)} F(z_t^N - z_0^N,\, t \geqslant 0) \,\right] = \mathbb{E}_\psi\left[\, F(z_t^N - z_0^N,\, t \geqslant 0) \,\right] \mathbb{E}\left[\, e^{\psi(z_0^N)} \,\right]$$
$$\xrightarrow[N\to\infty]{} \mathbb{E}\left[\, F(M_t,\, t \geqslant 0) \,\right] \mu_0(e^\psi).$$

Since F and ψ are arbitrary, the latter limit corresponds to the weak convergence of $(z_t^N)_{t\geqslant 0}$ towards $Z_t := Z_0 + M_t$, where $Z_0 \sim \mu_0$ and $(M_t)_{t\geqslant 0}$ are independent. This is exactly the desired result. □

Remark 3.23. In other words, the limit Gaussian process $(Z_t)_{t\in[0,T]}$ is the solution of the stochastic differential equation

$$\begin{cases} Z_0 & \sim \mu_0 \\ dZ_t & = \sqrt{i_t}\; dWt \end{cases}$$

where $(W_t)_{t\in[0,T]}$ is a standard Brownian motion.

Proposition 3.24. *Under Assumption (B), for any bounded φ such that Assumption (A') is satisfied, the sequence of martingales $(z_t^N)_{0\leqslant t\leqslant T}$ defined by*

$$z_t^N = \sqrt{N}\left(\gamma_t^N(Q^{T-t}(\varphi)) - \gamma_0(Q^{T-t}(\varphi))\right)$$

converges in law towards a Gaussian process $(Z_t)_{t\in[0,T]}$ with independent increments, initial distribution $\mathcal{N}(0, \mathbb{V}_{\eta_0}(Q^T(\varphi)))$ and variance function $\sigma_t^2(\varphi) = \mathbb{V}_{\eta_0}(Q^T(\varphi)) + i_t(\varphi)$, with $i_t(\varphi)$ defined by (3.26).

Proof. We just have to check that the assumptions of Theorem 3.22 are satisfied in our framework. Before proceeding, let us remind that since φ belongs to $\mathcal{D}$, it is necessarily bounded.

(i) Recall that $(X_0^1, \ldots, X_0^N)$ are i.i.d. with law $\eta_0 = \gamma_0$, so that clearly

$$z_0^N = \sqrt{N}\left(\gamma_0^N(Q^T(\varphi)) - \gamma_0(Q^T(\varphi))\right) \xrightarrow[N\to\infty]{d} (0, \mathbb{V}_{\eta_0}(Q^T(\varphi))).$$

(ii) This is a simple consequence of Corollary 3.2.

(iii) This is the purpose of Lemma 3.16.

(iv) By Definition 3.15, we have

$$i_t^N = \int_0^t \left(p_{u^-}^N\right)^2 d\mathbb{A}_u - \int_0^t \mathbb{V}_{\eta_{u^-}^N}(Q) p_{u^-}^N dp_u^N + \frac{1}{N}\int_0^t \left(p_{u^-}^N\right)^2 d\mathcal{R}_u,$$

so that

$$\Delta i_t^N \leqslant \Delta \mathbb{A}_t + \|\varphi\|_\infty^2 |\Delta p_t^N| + \frac{1}{N} |\Delta \mathcal{R}_t|.$$

It remains to see that $|\Delta p_t^N| \leqslant 1/N$ and to apply the bounds given in (3.11) and (3.15) to deduce that

$$\mathbb{E}\left[\sup_{0 \leqslant t \leqslant T} \Delta i_t^N\right] \leqslant \frac{2\|\varphi\|_\infty^2}{N} + \frac{14\|\varphi\|_\infty^2}{N^2} \xrightarrow[N\to\infty]{} 0.$$

(v) By Proposition 3.20, for all $t \in [0,T]$, i_t^N goes to $i_t(\varphi)$ in probability. Thus, the fact that the mapping $t \mapsto i_t(\varphi)$ is increasing is just a consequence of the same property for i_t^N, which is clear from the alternative formulation of i_t^N given just after Definition 3.15. It remains to prove the continuity of $t \mapsto i_t(\varphi)$, where

$$i_t(\varphi) = p_t \gamma_t(Q^2) - \gamma_0(Q^2) + [\,\gamma_t(Q)\,]^2 \ln p_t - 2\int_0^t \gamma_u\big(Q^2\big) dp_u.$$

Since $t \mapsto p_t$ is continuous by Lemma 2.1, it suffices to prove that the same holds true for the mapping $t \mapsto \gamma_t(Q^2)$. First, as explained in Section 4.3, Assumption (A) ensures that, for any initial condition $x \in F$, the jumps of the bounded martingale $t \mapsto M_t := Q^{T-t}(\varphi)(X_t)$ have an atomless distribution, hence M_t is almost surely continuous in t, and so is M_t^2. Second, fix $0 \leqslant t \leqslant T$, then by dominated convergence,

$$\lim_{h\to 0} \gamma_{t+h}(Q^2) = \lim_{h\to 0} \mathbb{E}\left[M_{t+h}^2\right] = \mathbb{E}\left[\lim_{h\to 0} M_{t+h}^2\right] = \mathbb{E}\left[M_t^2\right] = \gamma_t(Q^2),$$

which proves the continuity of $t \mapsto \gamma_t(Q^2)$.

□

Let us assume in the following discussion that Assumption (B) is satisfied. If we marginalize on the final time in Proposition 3.24, we obtain that, for any bounded φ such that Assumption (A') is satisfied,

$$\sqrt{N}\left(\gamma_T^N(\varphi) - \gamma_T(\varphi)\right) \xrightarrow[N\to\infty]{d} (0, \sigma_T^2(\varphi)).$$

We can now complete the proof of Proposition 3.3. Indeed, we can extend this result to any function φ in the $\|\cdot\|_\infty$-closure $\overline{\mathcal{D}}$ of $\mathcal{D}$, and thus establish Proposition 3.3, and in turn Theorem 2.6.

Lemma 3.25. *Under Assumptions (A') and (B), for any $\varphi \in \overline{\mathcal{D}}$, we have*

$$\mathbb{E}\left[\,\left(\gamma_T^N(\varphi) - \gamma_T(\varphi)\right)^2\,\right] \leqslant \frac{18\,\|\varphi\|_\infty^2}{N}.$$

Proof. For any φ in $\overline{\mathcal{D}}$, consider a sequence (φ_n) in $\mathcal{D}$ converging to φ with respect to the supremum norm. In particular, $(\|\varphi_n\|_\infty)$ goes to $\|\varphi\|_\infty$. Since $|\gamma_T(f)| \leqslant \|f\|_\infty$ and $|\gamma_T^N(f)| \leqslant \|f\|_\infty$, we have

$$\begin{aligned}\big(\gamma_T^N(\varphi) - \gamma_T(\varphi)\big)^2 &\leqslant 3\left\{\gamma_T^N(\varphi-\varphi_n)^2 + \big(\gamma_T^N(\varphi_n) - \gamma_T(\varphi_n)\big)^2 + \gamma_T(\varphi_n-\varphi)^2\right\}\\ &\leqslant 3\big(\gamma_T^N(\varphi_n) - \gamma_T(\varphi_n)\big)^2 + 6\|\varphi-\varphi_n\|_\infty^2.\end{aligned}$$

Now, Proposition 3.13 implies

$$\mathbb{E}\left[\,\big(\gamma_T^N(\varphi) - \gamma_T(\varphi)\big)^2\,\right] \leqslant \frac{18\,\|\varphi_n\|_\infty^2}{N} + 6\|\varphi-\varphi_n\|_\infty^2 \xrightarrow[n\to\infty]{} \frac{18\,\|\varphi\|_\infty^2}{N}.$$

□

We close this section with the proof of Proposition 3.3.

Proof. We will use the simplified version (3.29) of the asymptotic variance. Let us denote by Φ any bounded Lipschitz function, G a centered Gaussian variable with variance $\sigma_T^2(\varphi)$ for an arbitrary function $\varphi \in \overline{\mathcal{D}}$.

For any $\varepsilon > 0$, we can find φ_ε in $\mathcal{D}$ such that $\|\varphi - \varphi_\varepsilon\|_\infty \leqslant \varepsilon$. We can also assume that $\gamma_T(\varphi_\varepsilon) = \gamma_T(\varphi)$. Note that we can also choose φ_ε such that $|\sigma_T^2(\varphi_\varepsilon) - \sigma_T^2(\varphi)| \leqslant \varepsilon$. Indeed, it is easy to check by dominated convergence that $\varphi \mapsto \sigma_T^2(\varphi)$ is continuous for the norm $\|\cdot\|_\infty$. Hence, let us denote by G_ε a centered Gaussian variable with variance $\sigma_T^2(\varphi_\varepsilon)$.

Then we may write

$$\begin{aligned}&|\mathbb{E}[\Phi(\sqrt{N}(\gamma_T^N(\varphi) - \gamma_T(\varphi))] - \mathbb{E}[\Phi(G)]|\\ &\qquad\leqslant \mathbb{E}[|\Phi(\sqrt{N}(\gamma_T^N(\varphi) - \gamma_T(\varphi))) - \Phi(\sqrt{N}(\gamma_T^N(\varphi_\varepsilon) - \gamma_T(\varphi)))|]\\ &\qquad\quad + |\mathbb{E}[\Phi(\sqrt{N}(\gamma_T^N(\varphi_\varepsilon) - \gamma_T(\varphi)))] - \mathbb{E}[\Phi(G_\varepsilon)]|\\ &\qquad\quad + |\mathbb{E}[\Phi(G_\varepsilon)] - \mathbb{E}[\Phi(G)]|.\end{aligned}$$

For the first term, by Lemma 3.25, Jensen's inequality and remembering that $\gamma_T(\varphi - \varphi_\varepsilon) = 0$, we have

$$\mathbb{E}|\Phi(\sqrt{N}(\gamma_T^N(\varphi) - \gamma_T(\varphi))) - \Phi(\sqrt{N}(\gamma_T^N(\varphi_\varepsilon) - \gamma_T(\varphi)))| \leqslant 3\sqrt{2}\|\Phi\|_{\mathrm{Lip}}\|\varphi - \varphi_\varepsilon\|_\infty.$$

Hence, for any given $\delta > 0$, we can choose ε such that this first term is less than δ. Clearly, the same property holds for the third term as well. Besides, since φ_ε is in $\mathcal{D}$, for N large enough, the second term can also be made less than δ by Corollary 3.24. As this result holds for any bounded Lipschitz function Φ, we conclude using the Portmanteau theorem.

□

Remark 3.26. This corollary might be useful in practice: to obtain the CLT associated with any observable φ, it is sufficient to check Assumption (A) or (A') for appropriately regularized functions.

4 Appendix

4.1 Preliminary on Feller processes

In this section, we recall the definition and some properties of Feller processes (see also for example Section 17 of [10]).

Definition 4.1. *Let E be a locally compact Polish space. Let $C_0(E)$ denote the space of continuous functions that vanish at infinity. A càdlàg time-homogeneous process in E is Feller if and only if each of its probability transitions maps $C_0(E)$ into itself. Formally: for all $\varphi \in C_0(E)$ and $t \geqslant 0$, $z \mapsto \mathbb{E}[\varphi(Z_t)|Z_0 = z] \in C_0(E)$, where $(Z_t)_{t\geqslant 0}$ denotes the Markov process constructed with any given initial condition $Z_0 = z \in E$.*

Feller processes enjoy many useful standard properties including: (i) The associated natural filtration $\mathcal{F}_t^Z := \sigma(Z_{t'},\, 0 \leqslant t' \leqslant t)$ is right-continuous; (ii) Z is strong Markov with respect to $\mathcal{F}^Z$; (iii) Z is quasi-left continuous with respect to $\mathcal{F}^Z$. A characterization of quasi-left continuity is the following ([9], Proposition 2.26): if $(\tau_n)_{n\geqslant 1}$ is any increasing sequence of stopping times, then on the event $\{\lim_n \tau_n < +\infty\}$, one has $\lim_n Z_{\tau_n} = Z_{\lim_n \tau_n}$. Note that taking deterministic sequences implies that quasi-left continuous processes never jump at deterministic times.

We will need a slightly less standard property of Feller processes related to the so-called Skorokhod J_1 topology as defined in the following proposition.

Proposition 4.2 (J_1 topology)**.** *Let d be a metric of the Polish topology of E. Let $\mathbb{D}_E$ denote the space of càdlàg maps from $\mathbb{R}_+$ to E. There is a Polish topology on $\mathbb{D}_E$, called the Skorokhod J_1 topology, characterized by the following property: $\lim_n (z_t^n)_{t\geqslant 0} = (z_t)_{t\geqslant 0}$ in $\mathbb{D}_E$ if and only if there is a sequence $(\lambda^n)_{n\geqslant 0}$ of increasing one-to-one maps of $\mathbb{R}_+$ onto itself such that for each $t_0 \geqslant 0$*

$$\lim_n \sup_{0\leqslant t\leqslant t_0} d\big(z^n_{\lambda^n(t)}, z_t\big) = \lim_n \sup_{0\leqslant t\leqslant t_0} |\lambda^n(t) - t| = 0.$$

If Z is Feller, the distribution of $(Z_t)_{t\geqslant 0} \in (\mathbb{D}_E, J_1)$ is continuous with respect to its initial condition $Z_0 = z$. This is detailed in the following lemma.

Lemma 4.3. *Let $\mathbb{D}_E$ denote the space of càdlàg trajectories endowed with the Skorokhod J_1 topology, and let $(Z_s^z)_{s\geqslant 0}$ denote a Feller process with initial condition $Z_0 = z$. The map $z \mapsto \mathcal{L}(Z_t^z)_{t\geqslant 0}$ defined from E to probabilities on $\mathbb{D}_E$, endowed with convergence in distribution, is continuous.*

Proof. Let $(z^n)_{n\geqslant 0}$ be a sequence of initial conditions with $\lim_n z^n = z$ and denote $Z^n := Z^{z^n}$ as well as $Z := Z^z$. Then, by Condition (iv) of Theorem 17.25 in [10] (Condition (ii), which implies Condition (iv), is trivially true in the present context), the sequence of processes $((Z^n_t)_{t\geqslant 0})_{n\geqslant 0}$ converges in distribution towards $(Z_t)_{t\geqslant 0}$ in the Skorokhod space $(\mathbb{D}_E, J_1)$. □

We then recall the lower and upper continuity of hitting times with respect to the Skorokhod J_1 topology.

Lemma 4.4. *Let $B \subset E$, $(z_t)_{t\geqslant 0} \in \mathbb{D}_E$, and define $t_B(z) := \inf\{t \geqslant 0,\ z_t \in \mathring{B}\}$, as well as $\bar{t}_B(z) := \inf\{t \geqslant 0,\ z_{t^-} \in \overline{B} \text{ or } z_t \in \overline{B}\}$. Consider a converging sequence $\lim_n (z^n_t)_{t\geqslant 0} = (z_t)_{t\geqslant 0}$ in $(\mathbb{D}_E, J_1)$. Then t_B is upper continuous in $(\mathbb{D}_E, J_1)$:*

$$\limsup_n t_B(z^n) \leqslant t_B(z),$$

and $\bar{t}_B$ is lower continuous in $(\mathbb{D}_E, J_1)$:

$$\bar{t}_B(z) \leqslant \liminf_n \bar{t}_B(z^n).$$

Proof. For the upper continuity, without loss of generality, we can assume that $t_B(z) < \infty$. By right continuity of $(z_t)_{t\geqslant 0}$ and by definition of t_B, $z_{t_B(z)+\varepsilon} \in \mathring{B}$ for some arbitrary small enough $\varepsilon > 0$. By definition of the Skorokhod topology, there is a converging sequence $\lim_n t_n = t_B(z) + \varepsilon$ in $\mathbb{R}^+$ such that $\lim_n z^n_{t_n} = z_{t_B(z)+\varepsilon} \in \mathring{B}$. Thus, since $\mathring{B}$ is open, for any n large enough, $z^n_{t_n} \in \mathring{B}$ so that $t_B(z^n) \leqslant t_n$. The result follows by taking the limit $n \to +\infty$ and then $\varepsilon \to 0$, ε being arbitrary.

Concerning the lower continuity, set $t_0 := \liminf_n \bar{t}_B(z^n)$, which we assume finite without loss of generality. By definition of the hitting time functional $\bar{t}_B$, we can construct a sequence $(t_n)_{n\geqslant 1}$ such that, up to extraction, (i) $t_n \leqslant t_0+1$, (ii) $\lim_n t_n = t_0$, and (iii) $\lim_n d(z^n_{t^n}, B) = 0$ where d denotes a distance for the Polish space E. On the other hand, by time uniformity in the definition of the J_1 convergence $(z^n)_{t\geqslant 0} \to (z_t)_{t\geqslant 0}$, the set $\{z^n_t,\ t \leqslant t_0 + 1,\ n \geqslant 0\}$ is bounded. Hence, by compacity, there exists a sub-sequence of $(t_n)_{n\geqslant 1}$ satisfying $\lim_n t_n = t_0$ as well as $z^n_{t_n} \to b$, where $b \in \overline{B}$ by condition (iii) above. The convergence in J_1 topology implies that the extracted limit b necessarily belongs to $\{z_{t_0^-}, z_{t_0}\}$, implying that either $z_{t_0^-} \in \overline{B}$ or $z_{t_0} \in \overline{B}$. By definition of $\bar{t}_B$, this means that $\bar{t}_B(z) \leqslant t_0$. □

We can conclude with the key property that is useful in the proof of Proposition 2.9.

Lemma 4.5. *Let B be a subset of E, Z a Feller process, and $z \in E$ a given initial condition. Denote $\overline{\tau}_B := \inf\{t \geqslant 0,\ Z_{t^-} \in \overline{B} \text{ or } Z_t \in \overline{B}\} \in [0,\infty]$ as well as $\tau_B := \inf\{t \geqslant 0,\ Z_t \in \mathring{B}\} \in [0,\infty]$. Besides, assume that*

$$\mathbb{P}(\overline{\tau}_B = \tau_B | Z_0 = z) = 1. \tag{4.1}$$

Let $\lim_n z^n = z$ be a given converging sequence of initial conditions. Then the distribution of $\tau_B \in [0,\infty]$ under $\mathbb{P}(\ .\ | Z_0 = z^n)$ converges when $n \to \infty$ towards its distribution under $\mathbb{P}(\ .\ | Z_0 = z)$. Moreover if $\mathbb{P}(\tau_B < +\infty | Z_0 = z) > 0$, then the distribution of (Z_{τ_B}, τ_B) under $\mathbb{P}(\ .\ | Z_0 = z^n,\ \tau_B < \infty)$ converges when $n \to \infty$ towards its distribution under $\mathbb{P}(\ .\ | Z_0 = z,\ \tau_B < \infty)$.

Proof. Using Lemma 4.3 and a Skorokhod embedding argument, we can construct a sequence $(Z^n_t)_{t\geqslant 0}$ of processes with initial conditions $(z^n)_{n\geqslant 0}$ such that $\lim_n Z^n = Z$ in $(\mathbb{D}_E, J_1)$ almost surely. We claim that (i) $\lim_n \tau^n_B = \tau_B$, and (ii) $\lim_n Z^n_{\tau^n_B} = Z_{\tau_B}$ on the event $\{\tau_B < \infty\}$, which enable to conclude by dominated convergence.

On the one hand, Lemma 4.4 together with (4.1) directly implies (i).

On the other hand, let us work on the event $\{\tau_B < \infty\}$. The definition of the Skorokhod topology implies that the sequence $(Z^n_{\tau^n_B})_{n\geqslant 1}$ has its accumulation points included in $\{Z_{\tau^\infty_B -}, Z_{\tau^\infty_B}\}$. Since by construction $Z^n_{\tau^n_B} \in \overline{B}$, these accumulation points are also included in $\overline{B}$. We now claim that by quasi-left continuity of Z and condition (4.1), $Z_{\tau^-_B} \in \overline{B} \Rightarrow Z_{\tau^-_B} = Z_{\tau_B}$, which in turn implies from the discussion above that $\lim_n Z^n_{\tau^n_B} = Z_{\tau_B}$, and hence proves (ii) above. Indeed, defining $B_k = \{x,\ d(x,B) < 1/k\}$, one has by construction $\lim_k \tau_{B_k} = \overline{\tau}_B$ which also equals $\overline{\tau}_B = \tau_B$ by Assumption (4.1). Then quasi-left continuity implies that $\lim_k Z_{\tau_{B_k}} = Z_{\lim_k \tau_{B_k}} = Z_{\tau_B}$, while $Z_{\tau^-_B} \in \overline{B}$ implies $\tau_{B_k} < \tau_B$ so that $\lim_k Z_{\tau_{B_k}} = Z_{\tau^-_B}$, hence we get the claimed result. □

Proof of Proposition 2.9. The Feller property classically implies the quasi-left continuity of $t \mapsto \widetilde{X}$ and thus Condition (i) of Assumption (A) for all jump times except perhaps τ_∂.

Let (x^n, t^n) be a sequence in $F \times [0,T]$ converging to $(x,t) \in F \times [0,T]$. We claim that

$$\lim_n \mathbb{E}_{x^n}\left[\varphi(\widetilde{X}^n_{t_n})\mathbf{1}_{\tau^n_\partial > t^n}\right] = \mathbb{E}_x\left[\varphi(\widetilde{X}_t)\mathbf{1}_{\tau_\partial > t}\right], \tag{4.2}$$

which will ensure Condition (ii) of Assumption (A).

First, we claim that $\mathbb{P}_x(\tau_\partial = t) = 0$. Indeed, since $\widetilde{X}$ is Feller hence quasi left continuous, it cannot jump at a given $t \geqslant 0$ so that $\{\tau_\partial = t\} = \{\tau_\partial =$

t and $\widetilde{X}_t = \widetilde{X}_{t^-}\}$. Thus $\{\tau_\partial = t\}$ implies $\widetilde{X}_t \in \partial F$, which has probability zero by Condition (i) in Proposition 2.9.

Second, we claim that

$$\mathbb{P}_x(\overline{\tau}_\partial = \tau_\partial) = 1.$$

where $\overline{\tau}_\partial := \inf\{t,\ \widetilde{X}_{t^-} \in E \setminus F \text{ or } \widetilde{X}_t \in E \setminus F\} < \tau_\partial$. Indeed, by the strong Markov property of Feller processes, it is enough to prove that $\mathbb{P}_{\widetilde{X}_{\overline{\tau}_\partial}}(\tau_\partial > 0) = 0$, which is just a consequence of Condition (ii) in Proposition 2.9.

Finally, according to Lemma 4.3, a Skorokhod embedding argument shows that we can assume the almost sure convergence $\lim_n \widetilde{X}^n = \widetilde{X}$ in $(\mathbb{D}_E, J_1)$. Since $\widetilde{X}$ is Feller hence quasi left continuous, $\lim_n \widetilde{X}^n_{t_n} = \widetilde{X}_t$. To obtain (4.2), it remains to show that $\lim_n \tau^n_\partial = \tau_\partial$. This follows from Lemma 4.5 by simply taking $B = E \setminus \overline{F}$.

□

4.2 Stopping times and martingales

Lemma 4.6. *Let τ be a stopping time on a filtered probability space, and U an integrable and $\mathcal{F}_\tau$ measurable random variable such that $\mathbb{E}[\,U|\mathcal{F}_{\tau^-}\,] = 0$. Then the process $t \mapsto U\mathbf{1}_{t\geqslant\tau}$ is a càdlàg martingale.*

Proof. Let $t > s$ be given. First remark that $\mathbf{1}_{t\geqslant\tau} = \mathbf{1}_{s\geqslant\tau} + \mathbf{1}_{s<\tau}\mathbf{1}_{t\geqslant\tau}$. Then by definition of $\mathcal{F}_\tau$, $U\mathbf{1}_{s\geqslant\tau}$ is $\mathcal{F}_s$-measurable, so that

$$\mathbb{E}[\,U\mathbf{1}_{t\geqslant\tau}|\mathcal{F}_s\,] = U\mathbf{1}_{s\geqslant\tau} + \mathbb{E}[\,U\mathbf{1}_{t\geqslant\tau}|\mathcal{F}_s\,]\,\mathbf{1}_{s<\tau}.$$

Next, by definition of $\mathcal{F}_{\tau^-}$, $\mathbb{E}[\,U\mathbf{1}_{t\geqslant\tau}|\mathcal{F}_s\,]\,\mathbf{1}_{s<\tau}$ and $\mathbf{1}_{t\geqslant\tau}$ are $\mathcal{F}_{\tau^-}$-measurable, therefore

$$\mathbb{E}[\,U\mathbf{1}_{t\geqslant\tau}|\mathcal{F}_s\,]\,\mathbf{1}_{s<\tau} = \mathbb{E}[\,\mathbb{E}[\,U|\mathcal{F}_{\tau^-}\,]\,\mathbf{1}_{t\geqslant\tau}|\mathcal{F}_s\,]\,\mathbf{1}_{s<\tau} = 0.$$

The result follows. □

4.3 Proof of Lemma 3.1: (A) ⇒ (A')

The following obvious weakening of Assumption (A) is the raw condition that is required in the proof of Lemma 3.1.

(1) For any initial condition $x \in F$, the killing time has an atomless distribution, that is

$$\mathbb{P}(\tau_\partial = t|X_0 = x) = 0 \qquad \forall t \geqslant 0.$$

(2) There exists a space $\mathcal{D}$ of bounded measurable real-valued functions on F, which contains at least the indicator function $\mathbf{1}_F$, and such that for any $\varphi \in \mathcal{D}$, for any initial condition $x \in F$, the jumps of the càdlàg version of the martingale $t \mapsto Q^{t_0-t}(\varphi)(X_t)$ have an atomless distribution:

$$\mathbb{P}\big(\Delta Q^{t_0-t}(\varphi)(X_t) \neq 0 | X_0 = x\big) = 0 \qquad \forall 0 \leqslant t \leqslant t_0.$$

Our goal now is to prove that conditions (1) and (2) above imply Assumption (A'). Throughout the proof, let $1 \leqslant m \neq n \leqslant N$ and $j, k \geqslant 0$ be given integers. We recall that, by convention, $\tau_{n,0} = \tau_{m,0} = 0$.

(i) It is sufficient to prove that $\mathbb{P}(\tau_{n,k+1} = \tau_{m,j+1} \ \& \ \tau_{m,j} \leqslant \tau_{n,k}) = 0$, since taking the countable union of such events over $j, k \geqslant 0$ and $1 \leqslant m \neq n \leqslant N$ will yield the result. Conditionally on $\mathcal{F}_{\tau_{n,k}}$ and $\{\tau_{m,j} \leqslant \tau_{n,k}\}$, the two branching times $\tau_{n,k+1}$ and $\tau_{m,j+1}$ are independent. Moreover, Assumption (1) implies that conditionally on $\mathcal{F}_{\tau_{n,k}}$, $\tau_{n,k+1}$ has an atomless distribution. We deduce that

$$\mathbb{P}(\tau_{n,k+1} = \tau_{m,j+1} \ \& \ \tau_{m,j} \leqslant \tau_{n,k} | \mathcal{F}_{\tau_{n,k}}) = 0.$$

(ii) According to Proposition 1.3 in [9], we can define a countable sequence of stopping times $\sigma_{m,a}$ with $a \geqslant 1$ that exhaust the jumps of $\mathbb{L}^m_t$ for $\tau_{m,j} \vee \tau_{n,k} \leqslant t \leqslant \tau_{m,j+1}$. Conditionally on $\mathcal{F}_{\tau_{n,k}}$ and $\{\tau_{m,j} \leqslant \tau_{n,k}\}$, the two processes $(\mathbb{L}^n_t)_{t<\tau_{n,k+1}}$ and $(\mathbb{L}^m_t)_{t<\tau_{m,j+1}}$ are independent. Moreover, Assumption (2) implies that conditionally on $\mathcal{F}_{\tau_{n,k}}$, $\big(\mathbb{L}^n_t = Q^{T-t}(\varphi)(X^n_t)\big)_{\tau_{n,k} \leqslant t < \tau_{n,k+1}}$ has jumps with atomless distribution. As a consequence, for each $a \geqslant 1$,

$$\mathbb{P}\Big(\Delta \mathbb{L}^n_{\sigma_{m,a}} \neq 0 \ \& \ \tau_{m,j} \leqslant \tau_{n,k} \Big| \mathcal{F}_{\tau_{n,k}} \Big) = 0.$$

Taking the countable union of such events over $a \geqslant 1$, $j, k \geqslant 0$ and $1 \leqslant m \neq n \leqslant N$ gives the result.

(iii) One can apply the same reasoning as for (ii) with $\tau_{m,j+1}$ instead of $\sigma_{m,a}$.

4.4 About Soft Killing and Carré-du-Champ operator

As mentioned in Remark 2.4, we prove in [3] Theorem 2.6 and Corollary 2.7 under two different assumptions, respectively called (CC) (carré-du-champ operator) and (SK) (soft killing) assumptions. We refer the interested reader to [3] for the explicit formulations of (CC) and (SK). Without going into details, let us just justify why this is a more restrictive framework than the one presented here, meaning that (CC) and (SK) imply (A') and (B).

The fact that (SK) implies (B) is established in Lemma 3.2 of [3], where B_t stands for $N\mathcal{N}_t$. Moreover, Lemma 3.1 and Lemma 3.4 in [3] respectively ensure that (SK) implies (A')(i), and that (CC) and (SK) imply (A')(ii) and (A')(iii).

4.5 Integration rules

Remember that $p_t^N = (1-1/N)^{N\mathcal{N}_t}$, so that $p_0^N = 1$. Recall that $\sum_{j=1}^{\infty}(1-1/N)^{j-1} = N$.

Lemma 4.7. *Assume $N \geqslant 2$. Let $t \mapsto z_t^N$ be a càdlàg semi-martingale, $c > 0$ a deterministic constant, and consider the following conditions, satisfied for any branching time τ_j, $j \geqslant 1$:*

$$(i)\ |\Delta z_{\tau_j}^N| \leqslant c/N,\ (ii)\ |z_{\tau_j^-}^N| \leqslant c(1-1/N)^j,\ (iii)\ |\Delta z_{\tau_j}^N| \leqslant c(1-1/N)^j/N.$$

If (i) holds true, one has

$$\int_0^t p_{s^-}^N dz_s^N = p_t^N z_t^N - z_0^N - \int_0^t z_{s^-}^N dp_s^N + O(1/N). \tag{4.3}$$

If (ii) holds true, one has

$$\int_0^t z_{s^-}^N \left(p_{s^-}^N\right)^{-1} dp_s^N = \int_0^t z_{s^-}^N d\ln p_s^N + O(1/N). \tag{4.4}$$

Finally, if (iii) holds true, one has

$$\int_0^t z_{s^-}^N d\ln p_s^N = z_t^N \ln p_t^N - \int_0^t \ln p_{s^-}^N dz_s^N + O(1/N). \tag{4.5}$$

In all equations above, the O notation only depends on the deterministic constant c.

Proof. Equation (4.3) comes from the integration by parts formula defining the quadratic variation

$$p_t^N z_t^N - p_0^N z_0^N = \int_0^t z_{s^-}^N dp_s^N + \int_0^t p_{s^-}^N dz_s^N + [p^N, z^N]_t,$$

and the fact that $\Delta p_{\tau_j}^N = -(1-1/N)^{j-1}/N$ for all $j \geqslant 1$ so that

$$[p^N, z^N]_t = \sum_{j \geqslant 1} \Delta p_{\tau_j}^N \Delta z_{\tau_j}^N = O(1/N).$$

For (4.4), notice that for any jump time τ_j, $j \geqslant 1$, one has $\left(p^N_{\tau_j^-}\right)^{-1}\Delta p^N_{\tau_j} = -\frac{1}{N}$ as well as $\Delta \ln p^N_{\tau_j} = \log(1-\frac{1}{N})$, implying that

$$\left| \int_0^t z^N_{s^-} \left[\left(p^N_{s^-}\right)^{-1} dp^N_s - d\ln p^N_s \right] \right| \leqslant \sum_j |z^N_{\tau_j^-}| \left| \log(1-\frac{1}{N}) + \frac{1}{N} \right| = O(1/N).$$

Similarly to (4.3), Equation (4.5) is merely an integration by parts formula, with this time

$$[\ln p^N, z^N]_t = \sum_{j\geqslant 1} \Delta \ln p^N_{\tau_j} \Delta z^N_{\tau_j} = \log(1-\tfrac{1}{N}) \sum_j \Delta z^N_{\tau_j} = O(1/N). \qquad \square$$

References

[1] M. Bieniek, K. Burdzy, and S. Finch. Non-extinction of a Fleming-Viot particle model. *Probab. Theory Related Fields*, 153(1-2):293–332, 2012.

[2] K. Burdzy, R. Holyst, D. Ingerman, and P. March. Configurational transition in a Fleming-Viot-type model and probabilistic interpretation of Laplacian eigenfunctions. *Journal of Physics A: Mathematical and General*, 29(11):2633, 1996.

[3] F. Cérou, B. Delyon, A. Guyader, and M. Rousset. A Central Limit Theorem for Fleming-Viot Particle Systems with Soft Killing. *ArXiv:1611.00515v2*, 2016.

[4] F. Cérou, B. Delyon, A. Guyader, and M. Rousset. On the Asymptotic Normality of Adaptive Multilevel Splitting. *SIAM/ASA Journal on Uncertainty Quantification*, 7(1):1–30, 2019.

[5] P. Del Moral. *Feynman-Kac formulae, Genealogical and interacting particle systems with applications*. Springer-Verlag, New York, 2004.

[6] S.N. Ethier and T.G. Kurtz. *Markov processes*. John Wiley & Sons, Inc., New York, 1986.

[7] I. Grigorescu and M. Kang. Hydrodynamic limit for a Fleming-Viot type system. *Stochastic Process. Appl.*, 110(1):111–143, 2004.

[8] I. Grigorescu and M. Kang. Immortal particle for a catalytic branching process. *Probab. Theory Related Fields*, 153(1-2):333–361, 2012.

[9] J. Jacod and A.N. Shiryaev. *Limit theorems for stochastic processes*, volume 288. Springer-Verlag, Berlin, second edition, 2003.

[10] O. Kallenberg. *Foundations of Modern Probability*. Probability and Its Applications. Springer New York, 2002.

[11] J.-U. Löbus. A stationary Fleming-Viot type Brownian particle system. *Math. Z.*, 263(3):541–581, 2009.

[12] P.E. Protter. *Stochastic integration and differential equations*, volume 21 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, second edition, 2005.

[13] M.J. Schervish. *Theory of statistics*. Springer Series in Statistics. Springer-Verlag, New York, 1995.

[14] D. Villemonais. General approximation method for the distribution of Markov processes conditioned not to be killed. *ESAIM Probab. Stat.*, 18:441–467, 2014.