On Synchronized Fleming-Viot Particle Systems

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

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

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

Abstract

This article presents a variant of Fleming-Viot particle systems, which are a standard way to approximate the law of a Markov process with killing as well as related quantities. Classical Fleming-Viot particle systems proceed by simulating N trajectories, or particles, according to the dynamics of the underlying process, until one of them is killed. At this killing time, the particle is instantaneously branched on one of the $(N-1)$ other ones, and so on until a fixed and finite final time T. In our variant, we propose to wait until K particles are killed and then rebranch them independently on the $(N-K)$ alive ones. Specifically, we focus our attention on the large population limit and the regime where K/N has a given limit when N goes to infinity. In this context, we establish consistency and asymptotic normality results. The variant we propose is motivated by applications in rare event estimation problems through its connection with Adaptive Multilevel Splitting and Subset Simulation.

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

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. X is called a killed Markov process with cemetery point ∂. 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 is alive at time T, i.e., $\mathcal{L}(X_T | \tau_\partial > T)$, and in the probability of this event, that is $p_T := \mathbb{P}(\tau_\partial > T)$, with the natural assumption that $p_T > 0$ (see Figure 1). Without loss of generality, we will assume for simplicity that $\mathbb{P}(X_0 = \partial) = 0$, that is $p_0 = 1$. Let us stress that in all this paper, T is held fixed and finite.

We will also assume – as a consequence of Assumption (A) below – that the non-increasing function $t \mapsto p_t$ is continuous on $[0, T]$, and we will consider the approximation ρ_t of p_t defined by

$$t \mapsto \rho_t := \theta^{\lfloor \log p_t / \log \theta \rfloor},$$

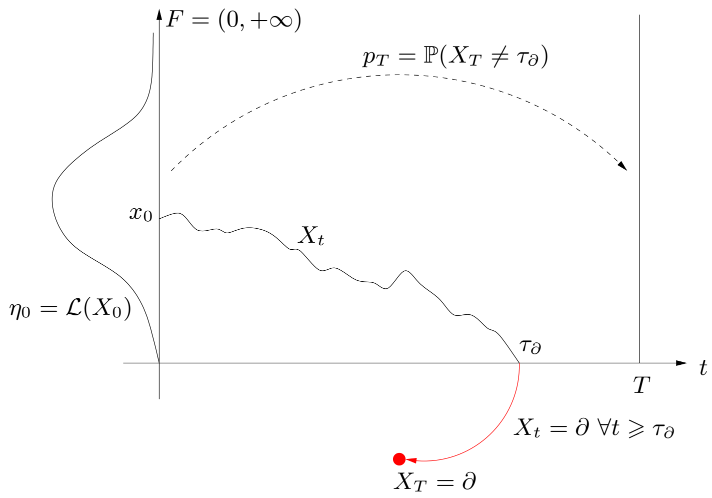

Figure 1: A Markov process with killing.

where $\theta \in (0, 1)$ is a given probability typically much larger than p_T. To fix ideas, one can think of p_T being lower than 10^{-6} while $\theta = 1/2$. In other words, $t \mapsto \rho_t$ is the right continuous and piecewise constant function that coincides with p_t for each power of θ, that is (see Figure 2)

$$\rho_t = \theta^j \iff \theta^{j+1} < p_t \leqslant \theta^j, \; j \in \mathbb{N}.$$

Let us denote

$$j_{\max} = \lfloor \log p_T / \log \theta \rfloor.$$

For simplifying technical reasons, we assume that $\log p_T / \log \theta$ is not an integer, so that $p_T = r\theta^{j_{\max}}$ with $\theta < r < 1$. Moreover, we suppose that for each $0 \leqslant j \leqslant j_{\max}$, there exists a unique t_j such that

$$p_{t_j} = \mathbb{P}(\tau_\partial > t_j) = \theta^j.$$

This is obviously true if the non-increasing function $t \mapsto p_t$ is in fact strictly decreasing on $[0, T]$ (see Figure 2).

Under Assumptions (A) and (B) that will be detailed below, the following process is well defined for any number of particles $N \geqslant 2$ and any $1 \leqslant K < N$. We propose to call it a Synchronized Fleming-Viot Particle System (see Figure 3). The "classical" Fleming-Viot Particle System corresponds to the case where $K = 1$, see for example [5, 18, 23, 19, 3, 27, 7, 8] and references therein.

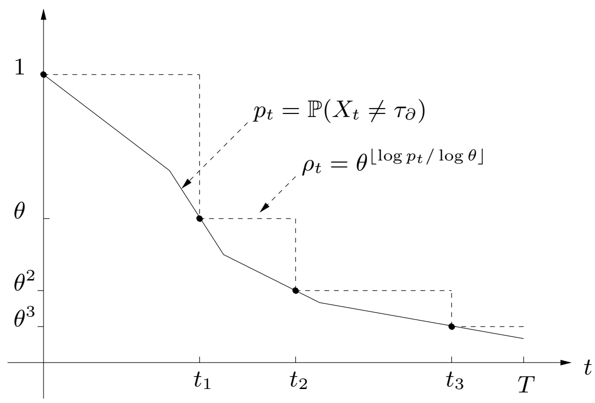

Figure 2: The mappings $t \mapsto p_t$ and $t \mapsto \rho_t$ when $\theta = 1/2$.

Definition 1.1 (Synchronized Fleming-Viot Particle System)**.** *Let X denote a killed Markov process in F with cemetery point ∂. The associated synchronized Fleming-Viot particle system $(X_t^1, \cdots, X_t^N)_{t\in[0,T]}$ with K synchronized branchings is the Markov process with state space $(F \times \{\partial\})^N$ defined by the following set of rules:*

- *Initialization: consider N i.i.d. particles*
$$X_0^1, \ldots, X_0^N \overset{\text{i.i.d.}}{\sim} \eta_0 = \mathcal{L}(X_0),$$
- *Evolution and killing: each particle evolves independently according to the law of the underlying Markov process X until K of them hit ∂ (or the final time T is reached),*
- *Branching (or rebirth, or splitting): the K killed particles are taken from ∂, and are independently and instantaneously given the state of one of the $(N - K)$ other particles (randomly uniformly chosen),*
- *and so on until final time T.*

As will be proved in Proposition 2.2, it turns out that the sequence of quantiles $(t_j)_{0<j\leqslant j_{\max}}$ are approximated by the sequence of successive branchings times $(\tau_j)_{0<j\leqslant j_{\max}}$ of the synchronized Fleming-Viot particle system when $K := K_N$ satisfies
$$1 - \frac{K_N}{N} \xrightarrow[N\to\infty]{} \theta. \tag{1.1}$$
Additionally, we define the right continuous counting process
$$t \mapsto \mathcal{B}_t = \mathcal{B}_t^N := \text{card}\,\{\text{branching times} \leqslant t\}\,.$$

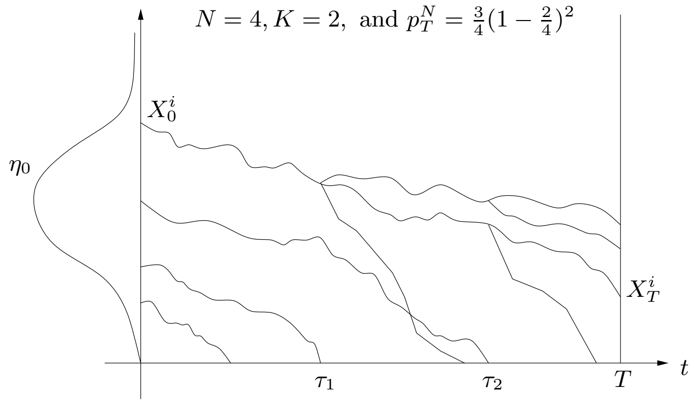

Figure 3: A Synchronized Fleming-Viot Particle System.

and then consider the estimators

$$\eta_t^N := \frac{1}{N}\sum_{n=1}^{N} \delta_{X_t^n} \qquad \text{and} \qquad \rho_t^N := \left(1 - \tfrac{K}{N}\right)^{\mathcal{B}_t},$$

of

$$\eta_t := \mathcal{L}(X_t | \tau_\partial > t_j) \quad \forall\ t_j \leqslant t < t_{j+1},$$

and ρ_t respectively. Since we have assumed that $p_0 = 1$, we have $p_0 = \rho_0$ and $\eta_0 = \mathcal{L}(X_0)$. Note that the quantiles t_j are implicitly estimated by the jump times τ_j of the process $t \mapsto \rho_t^N$ (see Proposition 2.2). Moreover, we emphasize that η_t is not $\mathcal{L}(X_t|\tau_\partial > t)$, which is the law of the process X_t given that it is still alive at time t. In particular, η_t does not define a probability measure on F, but on $F \cup \{\partial\}$ with a Dirac at ∂ associated with the probability $\mathbb{P}(\tau_\partial \leqslant t | \tau_\partial > t_j)$ for all t such that $t_j \leqslant t < t_{j+1}$.

The distribution of the process restricted to F, that is

$$\gamma_t(\varphi) := \mathbb{E}[\varphi(X_t)] \quad \forall \varphi \ \text{ with } \ \varphi(\partial) = 0,$$

is then estimated by

$$\gamma_t^N(\varphi) := \rho_t^N \times \eta_t^N(\varphi). \tag{1.2}$$

The probability p_t that the process is still alive at time t is estimated by (see also Figure 3)

$$p_t^N := \gamma_t^N(\mathbf{1}_F) = \rho_t^N \times \eta_t^N(\mathbf{1}_F) = \left(1 - \frac{K}{N}\right)^{\mathcal{B}_t} \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}_F(X_t^n),$$

and the distribution of X_t conditioned to be alive, i.e., $\mathcal{L}(X_t|\tau_\partial > t)$, by

$$\frac{\gamma_t^N(\varphi)}{\gamma_t^N(\mathbf{1}_F)} = \frac{\eta_t^N(\varphi)}{\eta_t^N(\mathbf{1}_F)} \quad \forall\varphi \text{ with } \varphi(\partial) = 0.$$

Note that this notation differs from the work in [8] and from classical particle models with fixed population size where the empirical distribution of particles η_t^N estimates the law of X_t conditioned to be alive at time t. In particular, in the present work, the mapping $t \mapsto \eta_t(\mathbf{1}_F)$ is càdlàg with jumps at the quantiles t_j with $\Delta\eta_{t_j}(\mathbf{1}_F) = 1 - \theta$.

The purpose of this paper is to extend the results obtained in [8] for classical Fleming-Viot particle systems to the synchronized Fleming-Viot particle systems that we have just defined under the scaling assumption (1.1), that is when $K = K_N$ is proportional to N. The main result corresponds to Theorem 2.3 and Corollary 2.4, which provide CLT type results for the estimators γ_T^N, p_T^N and $\eta_T^N/\eta_T^N(\mathbf{1}_F)$. We also prove convergence of the branching times τ_j towards the corresponding quantiles t_j in Proposition 2.2.

We refer to [8] for examples where Assumptions (A) and (B) below are verified. In particular, this includes the case where X_t is a regular enough uniformly elliptic diffusive process killed when hitting the smooth boundary of a given compact domain.

Finally, note that this work is motivated by practical applications in rare event estimation problems. We refer to [1, 2, 11, 13, 20, 6] for the presentation of the set of methods we have in mind, called Adaptive Multilevel Splitting or Subset Simulation. More precisely, we prove in [9] that the so-called "last particle" version of splitting algorithms can be interpreted as Fleming-Viot particle systems, using a score function (also called an importance function or a reaction coordinate) as a new time-index. Using the same representation with the same change of time (that we will not detail again here), the present paper enables us to obtain CLT type results for versions of this algorithm where the K particles with lowest scores are killed at each step. In fact, the original version of Adaptive Multilevel Splitting that was proposed in [11], with theoretical results only in dimension 1, corresponds exactly to the case where K particles are killed at each step, when $K = K_N$ is proportional to N. Therefore, the results of the present article allow us to deduce asymptotic results for Adaptive Multilevel Splitting under very general assumptions and in any finite dimension. The interested reader can find applications of Adaptive Multilevel Splitting in various fields in [10, 26, 22, 4, 14, 24] and references therein.

The rest of the paper is organized as follows. Section 2 details our assumptions and exposes the main results of the paper, Section 3 is dedicated to the proofs while Section 4 gathers some supplementary material.

2 Main result

2.1 Assumptions

We assume that F is a measurable subset of some reference Polish space, and that for each initial condition, under this reference topology, X is càdlàg in $F \cup \{\partial\}$ and satisfies the time-homogeneous Markov property. Its probability transition is denoted Q, meaning that there is a semi-group operator $(Q^t)_{t \geqslant 0}$ 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 latter and in all what follows, 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 that is also denoted $(Q^t)_{t \geqslant 0}$.

For any bounded $\varphi : F \to \mathbb{R}$ extended with the convention $\varphi(\partial) = 0$ and any $t \in [0, T]$, we can then consider the unnormalized measure

$$\gamma_t(\varphi) = \mathbb{E}[\varphi(X_t)] = \mathbb{E}[\varphi(X_t)\mathbf{1}_{\tau_\partial > t}] = \eta_0 Q^t \varphi,$$

with $X_0 \sim \eta_0 = \gamma_0$. For any $t \in [0, T]$, one has $p_t = \mathbb{P}(\tau_\partial > t) = \gamma_t(\mathbf{1}_F)$. As mentioned before, the associated empirical approximation is given by

$$\gamma_t^N := \rho_t^N \eta_t^N,$$

so that we can define the limiting measure

$$\eta_t := \gamma_t / \rho_t.$$

We stress again that, contrary to [8], $\eta_t(\mathbf{1}_F) \neq 1$ in general but since we have assumed that the process X_t is càdlàg we still have $\eta_{t_j}(\mathbf{1}_F) = 1$ and

$$\eta_{t_j} = \mathcal{L}\big(X_{t_j} | \tau_\partial > t_j\big).$$

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. Our fundamental assumptions can now be detailed.

Assumption (A). *Let X denote a time-homogeneous Markov process.*

(i) For any initial condition $X_0 = x \in F$, the distribution of the killing time τ_∂ is atomless.

(ii) 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 $\varphi \in \mathcal{D}$, any initial condition $X_0 = x$ and any final time T, the jumps of the càdlàg martingale $t \mapsto Q^{T-t}(\varphi)(X_t)$ have an atomless distribution.

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 $\mathbb{P}(\mathcal{B}_T < +\infty) = 1$.*

Under Assumptions (A) and (B), the non-increasing jump processes $t \mapsto p_t^N$ and $t \mapsto \rho_t^N$ are strictly positive.

Remark 2.1. Condition (i) of Assumption (A) and Lebesgue's continuity theorem imply that, for any initial distribution η_0 on F, the non-increasing mapping $t \mapsto p_t = \mathbb{P}(\tau_\partial > t)$ is continuous.

Our third and final assumption ensures the strict monoticity of $t \mapsto p_t$ at each quantile.

Assumption (C). *The continuous non-increasing mapping $t \mapsto p_t$ is strictly decreasing at each t_j, $j = 1, \ldots, j_{\max}$, that is $p_s > p_{t_j} > p_t$ for all $s < t_j < t$.*

2.2 Main result

We keep the notation of Section 1. In particular, $(X_t^1, \ldots, X_t^N)_{t \geqslant 0}$ denotes the synchronized Fleming-Viot particle system, and

$$\tau_j := j\text{-th branching time of the particle system.}$$

Accordingly, the number $\mathcal{B}_t$ of branchings until time t is $\mathcal{B}_t = \sum_{j=1}^{\infty} \mathbf{1}_{\tau_j \leqslant t}$.

For the upcoming results, we work under Assumptions (A), (B), and (C), and we assume that $K = K_N$ satisfies

$$\frac{K_N}{N} \xrightarrow[N \to \infty]{} 1 - \theta \in (0, 1).$$

We start with the convergence of the branching times towards the quantiles.

Proposition 2.2. *We have*

$$(\tau_1, \ldots, \tau_{j_{\max}}) \xrightarrow[N\to+\infty]{\mathbb{P}} (t_1, \ldots, t_{j_{\max}}) \qquad \text{and} \qquad \mathbb{P}(\tau_{j_{\max}+1} \leqslant T) \xrightarrow[N\to+\infty]{} 0.$$

We can now expose the main result of the present paper. As usual, $\mathcal{N}(m, \sigma^2)$ denotes the normal distribution with mean m and variance σ^2. Furthermore, for any probability distribution μ on F and any test 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.,

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

Theorem 2.3. *Let us denote by $\overline{\mathcal{D}}$ the closure with respect to the norm $\|\cdot\|_\infty$ of the space $\mathcal{D}$ satisfying Condition (ii) of Assumption (A). Then for any φ in $\overline{\mathcal{D}}$ extended with $\varphi(\partial) = 0$, one has the convergence in distribution*

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

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

$$\begin{aligned}\sigma_T^2(\varphi) := \quad & \theta^{2j_{\max}}\mathbb{V}_{\eta_T}(\varphi) + j_{\max}(1/\theta - 1)\theta^{2j_{\max}}\eta_T(\varphi)^2 \\ & + \sum_{j=1}^{j_{\max}} \mathbb{V}_{\eta_{t_j}}\left(Q^{T-t_j}(\varphi)\right)\left(\theta^{2j-1} - \theta^{2j+1}\right).\end{aligned}$$

For classical Fleming-Viot particle systems, we have shown in [8] that, under the same assumptions and denoting $\eta_t = \mathcal{L}(X_t | t > \tau_\partial)$, the asymptotic variance takes the form

$$\sigma_T^2(\varphi) = p_T^2\mathbb{V}_{\eta_T}(\varphi) - p_T^2\log(p_T)\eta_T(\varphi)^2 - 2\int_0^T \mathbb{V}_{\eta_t}\left(Q^{T-t}(\varphi)\right)p_t dp_t. \quad (2.1)$$

Returning to our variant and the result of Theorem 2.3, since $\mathbf{1}_F \in \mathcal{D}$ by assumption, and $\gamma_T(\mathbf{1}_F) = p_T$, the CLT for $\eta_T^N/\eta_T^N(\mathbf{1}_F)$ is then a straightforward application of this result by considering the decomposition

$$\begin{aligned}&\sqrt{N}\left(\frac{\eta_T^N(\varphi)}{\eta_T^N(\mathbf{1}_F)} - \frac{\eta_T(\varphi)}{\eta_T(\mathbf{1}_F)}\right) = \\ &\quad \frac{1}{\gamma_T^N(\mathbf{1}_F)}\sqrt{N}\left(\gamma_T^N(\varphi - \eta_T(\varphi)\mathbf{1}_F/\eta_T(\mathbf{1}_F)) - \gamma_T(\varphi - \eta_T(\varphi)\mathbf{1}_F/\eta_T(\mathbf{1}_F))\right),\end{aligned}$$

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

Corollary 2.4. *One has*

$$\sqrt{N}\left(p_T^N - p_T\right) \xrightarrow[N\to\infty]{d} \mathcal{N}(0, \sigma_T^2(\mathbf{1}_F)).$$

Additionally, for any φ in $\overline{\mathcal{D}}$,

$$\sqrt{N}\left(\frac{\eta_T^N(\varphi)}{\eta_T^N(\mathbf{1}_F)} - \frac{\eta_T(\varphi)}{\eta_T(\mathbf{1}_F)}\right) \xrightarrow[N\to\infty]{d} \mathcal{N}\big(0, \sigma_T^2(\varphi - \eta_T(\varphi)\,\mathbf{1}_F/\eta_T(\mathbf{1}_F))/p_T^2\big).$$

Let us comment on the relative asymptotic variance of p_T^N, i.e., $\sigma_T^2(\mathbf{1}_F)/p_T^2$. Since by definition $p_T = r\theta^{j_{\max}}$, with $j_{\max} = \lfloor \log p_T / \log\theta \rfloor$ and $\theta < r < 1$, we have $\eta_T(\mathbf{1}_F) = r$ and $\mathbb{V}_{\eta_T}(\mathbf{1}_F) = r(1-r)$. Therefore, we obtain

$$\frac{\sigma_T^2(\mathbf{1}_F)}{p_T^2} = j_{\max}\frac{1-\theta}{\theta} + \frac{1-r}{r} + \frac{1}{p_T^2}\sum_{j=1}^{j_{\max}} \mathbb{V}_{\eta_{t_j}}\left(Q^{T-t_j}(\mathbf{1}_F)\right)\left(\theta^{2j-1} - \theta^{2j+1}\right).$$

In the latter, the variance terms may be reformulated as

$$\mathbb{V}_{\eta_{t_j}}\left(Q^{T-t_j}(\mathbf{1}_F)\right) = \mathbb{V}(\mathbb{P}(X_T \neq \partial | X_{t_j})) = \mathbb{E}\left[\left(\mathbb{P}(X_T \neq \partial | X_{t_j}) - \frac{p_T}{p_{t_j}}\right)^2\right].$$

For each $1 \leqslant j \leqslant j_{\max}$, if $X_{t_j} \sim \eta_{t_j}$, $\mathbb{P}(X_T \neq \partial | X_{t_j})$ is a random variable with values between 0 and 1 and expectation p_T/p_{t_j} so that the maximal variance is reached by a Bernoulli random variable with parameter p_T/p_{t_j}. As a consequence,

$$0 \leqslant \mathbb{V}_{\eta_{t_j}}\left(Q^{T-t_j}(\mathbf{1}_F)\right) \leqslant \frac{p_T}{p_{t_j}}\left(1 - \frac{p_T}{p_{t_j}}\right) = r\theta^{j_{\max}-j}\left(1 - r\theta^{j_{\max}-j}\right).$$

Then a straightforward computation yields

$$j_{\max}\frac{1-\theta}{\theta} + \frac{1-r}{r} \leqslant \frac{\sigma_T^2(\mathbf{1}_F)}{p_T^2} \leqslant \frac{1+\theta}{p_T} - \frac{\theta+r}{r} - j_{\max}(1-\theta).$$

On the one side, concerning the upper-bound, we see that

$$\lim_{\theta\to 1^-} \frac{1+\theta}{p_T} - \frac{\theta+r}{r} - j_{\max}(1-\theta) = 2\frac{1-p_T}{p_T} + \log p_T.$$

Interestingly, considering (2.1), this limit corresponds to the upper-bound obtained for classical Fleming-Viot particle systems, that is when $K = 1$, as shown in [8]. In particular, if p_T is very low, this is approximately equal to $2(1-p_T)/p_T$. Clearly, this is twice the relative variance of a naive (or standard) Monte Carlo simulation. Consequently, one has to pay attention to the fact that, in some very specific situations, Fleming-Viot particle systems may lead to very poor estimators.

On the other side, the lower-bound already appeared in the context of Adaptive Multilevel Splitting (see for example [11, 6, 12]). This bound is reached when, for each j, starting with law η_{t_j} at time t_j, the probability of being still alive at time T is constant on the support of η_{t_j}. Everything happens as if one would estimate independently $j_{\max}$ times the probability θ and one time the probability r, all this being done by naive Monte Carlo with independent samples of common size N. Using standard tools, one can see that the resulting product estimator $\hat{p}_T^N = \hat{p}_{t_1}^N \dots \hat{p}_{t_{j_{\max}}}^N \hat{r}^N$ satisfies the CLT

$$\sqrt{N}\,\frac{\hat{p}_T^N - p_T}{p_T} \xrightarrow[N\to\infty]{d} \mathcal{N}\left(0, j_{\max}\frac{1-\theta}{\theta} + \frac{1-r}{r}\right),$$

which is exactly the above-mentioned lower-bound.

Additionally, since $j_{\max} = \lfloor \log p_T / \log \theta \rfloor$ and $\theta < r = p_T \theta^{-j_{\max}} < 1$, this lower-bound may be rewritten, for all $0 < \theta < 1$, as

$$h(\theta) := \left\lfloor \frac{\log p_T}{\log \theta} \right\rfloor \frac{1-\theta}{\theta} + \frac{\theta^{\left\lfloor \frac{\log p_T}{\log \theta} \right\rfloor}}{p_T} - 1.$$

One can check that h is non increasing on $(0,1)$, so that the minimal possible relative asymptotic variance is simply $\lim_{\theta\to 1^-} h(\theta) = -\log p_T$. As explained in [8] and can be deduced from (2.1), this precisely corresponds to the minimal relative asymptotic variance for classical Fleming-Viot particle systems, that is when $K = 1$.

In view of this, one could argue that the best thing to do for estimating p_T is to use classical Fleming-Viot particle systems. However, things are not so simple. First, because as far as we can judge, this is only the case when the variance terms $\mathbb{V}_{\eta_t}\big(Q^{T-t}(\mathbf{1}_F)\big)$ are zero at all time, which is a very particular situation. Second, because as for all Monte Carlo methods, one should not only compare the variances of different methods, but also their respective algorithmic complexities, or costs.

More explicitly, suppose that, by convention, the algorithmic cost for the simulation of a single trajectory/particle until its killing is equal to 1. For classical Fleming-Viot particle systems, we have proved in [9], Corollary 2.9, that the number of resampling is equal to $-N \log p_T + O_p(\sqrt{N})$, so that the total cost is $N(1 - \log p_T) + O_p(\sqrt{N})$, i.e., N initial trajectories plus

$-N \log p_T + O_p(\sqrt{N})$ rebranched ones. For synchronized Fleming-Viot particle systems, Remark 3.12 ensures that the number of resamplings goes to $j_{\max} = \lfloor \log p_T / \log \theta \rfloor$ in probability. Since $K \sim (1-\theta)N$ particles are rebranched at each step, the total cost is asymptotically equivalent to

$$N\left(1 + \left\lfloor \frac{\log p_T}{\log \theta} \right\rfloor (1-\theta)\right),$$

which is less than $N(1 - \log p_T)$ for any $0 < \theta < 1$. Beyond this lower algorithmic cost, it is also worth noting that synchronized Fleming-Viot particle systems can easily be parallelized, contrary to classical ones.

3 Proofs

3.1 Overview

The probability space is filtered by the natural filtration of the particle system, denoted $(\mathcal{F}_t)_{t \geqslant 0}$. Note that $\mathcal{F}_t$ contains all the events related not only to the trajectories of the particles, but also all the auxiliary variables used for the resamplings, up to time t.

The key object of the proof is the càdlàg martingale

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

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

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

is the final value of the latter martingale, with the addition of a second term depending on the initial condition. This second term satisfies a CLT by assumption. 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 Proposition 3.15. However, this requires several intermediate steps, mainly for the calculation of the quadratic variation $N[\gamma^N(Q), \gamma^N(Q)]_t$. In the sequel, we will make extensive use of stochastic calculus for càdlàg semimartingales, as presented for example in [25] chapter II, or [21].

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

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

We will also denote, for each $1 \leqslant n \leqslant N$ and any $t \in [0,T]$,

$$\mathbb{L}_t^n := Q^{T-t}(\varphi)(X_t^n), \qquad\qquad \mathbb{L}_t := \frac{1}{N}\sum_{n=1}^{N} \mathbb{L}_t^n = \eta_t^N(Q), \tag{3.2}$$

where, again, the fixed parameters T and φ are implicit.

3.2 Martingale decomposition

Let us recall that τ_j denotes the j-th branching time of the particle system. We will need some additional notation related to the behavior of the particle system at each branching time.

Definition 3.1.

- *(Individual indexation of branching times) For any $n \in \{1,\dots,N\}$ and any $k \geqslant 0$, we denote by*

$$\tau_{n,k} := \text{ k-th rebirth (or branching) time of particle } n,$$

with the convention $\tau_{n,0} = 0$.

- *(Surviving particles) $X_{\tau_j}^{n,-}$, for $n \in \{1,\dots,N\}$, denotes the state of particle n* after *the last jump on ∂ of the last killed particle at τ_j, but* before *the resampling. We also denote*

$$\begin{aligned}\mathrm{Alive}_j &:= \{(N-K) \text{ particles that are not resampled at time } \tau_j\} \\ &:= \left\{ n \in \{1,\dots,N\} \text{ s.t. } X_{\tau_j}^{n,-} \neq \partial \right\}.\end{aligned}$$

We may also use the individual indexation, for instance $\mathrm{Alive}_{n,k}$ *is the set of particles that are not resampled at time $\tau_{n,k}$.*

- *(σ-field before resampling) For each branching time τ_j we also define $\mathcal{F}_{\tau_j}^- := \mathcal{F}_{\tau_j^-} \vee \sigma(X_{\tau_j}^{n,-}, n \in \{1,\dots,N\})$. We obviously have $\mathcal{F}_{\tau_{j-}} \subset \mathcal{F}_{\tau_j}^- \subset \mathcal{F}_{\tau_j}$.*

Remark 3.2. 1. In general, if $K \geqslant 2$, the branching time $\tau_{n,k}$ is different from the k-th *killing* time of particle n. However, note that the latter belongs to the time interval $(\tau_{n,k-1}, \tau_{n,k}]$, the upper-bound being reached if n is the last particle of $\{1,\dots,N\} \setminus \mathrm{Alive}_{n,k}$ to be killed.

2. In the previous definition, the word "Alive" in "Alive$_j$" is a slight abuse of terminology. Indeed, if for example $j = 1$, all particles are alive at time τ_1. At time τ_1^-, $(K-1)$ are dead (i.e., equal to ∂) and τ_1 is the K-th killing date, at which all of the K "dead" particles are instantaneously resampled, i.e., branched on the $(N-K)$ "alive" ones. This is illustrated on Figure 3. As we will see, there is exactly one particle for which $X^{n,-}_{\tau_j}$ is not equal to $X^n_{\tau_j^-}$, namely the particle which is killed at time τ_j.

This section builds upon the same martingale representation as in [27]. Namely, 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 3.3. 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).

If $\widetilde{X}_t$ is any particle evolving according to the dynamics 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

$$\mathbb{M}_t^{n,k} := \left(\mathbf{1}_{t<\tau_{n,k}}\mathbb{L}_t^n - \mathbb{L}^n_{\tau_{n,k-1}}\right)\mathbf{1}_{t\geqslant\tau_{n,k-1}} = \begin{cases} 0 & \text{if } t < \tau_{n,k-1} \\ \mathbb{L}_t^n - \mathbb{L}^n_{\tau_{n,k-1}} & \text{if } \tau_{n,k-1} \leqslant t < \tau_{n,k} \\ -\mathbb{L}^n_{\tau_{n,k-1}} & \text{if } \tau_{n,k} \leqslant t \end{cases} \tag{3.3}$$

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

$$\mathbb{M}_t^n := \sum_{k=1}^{\infty} \mathbb{M}_t^{n,k} = \mathbb{L}_t^n - \sum_{0\leqslant\tau_{n,k}\leqslant t} \mathbb{L}^n_{\tau_{n,k}}, \tag{3.4}$$

$$\mathbb{M}_t := \frac{1}{\sqrt{N}}\sum_{n=1}^{N} \mathbb{M}_t^n = \sqrt{N}\left(\mathbb{L}_t - \sum_{0\leqslant\tau_{n,k}\leqslant t} \mathbb{L}_{\tau_{n,k}}\right), \tag{3.5}$$

are local martingales. The scaling by a $1/\sqrt{N}$ factor in the definition of $\mathbb{M}$ is there to ensure that the variance of the latter is of order 1 in the large population limit.

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

$$\mathcal{M}_t^{n,k} := \left(1 - \frac{K}{N}\right)\left(\mathbb{L}^n_{\tau_{n,k}} - \frac{1}{N-K} \sum_{m \in \text{Alive}_{n,k}} \mathbb{L}^m_{\tau_{n,k}}\right)\mathbf{1}_{t \geqslant \tau_{n,k}}. \tag{3.6}$$

By Lemma 4.1, this is a piecewise constant martingale with a single jump at $t = \tau_{n,k}$, and it is clearly bounded by $2\,\|\varphi\|_\infty$. Then, under Assumption (B), the processes

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

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

are also local martingales. Again, the scaling by $1/\sqrt{N}$ is chosen to ensure that the variance of the latter is of order 1 in the large population limit.

Lemma 3.4 (About the jumps of the martingales)**.** *Under Assumption (A):*

(i) For each n, the jumps of $\mathcal{M}_t^n$ only happen at branching times, more precisely at times $\tau_{n,k}$ for $k \geqslant 1$.

(ii) For all $j \geqslant 1$, one has $\Delta\mathbb{M}^n_{\tau_j} = 0$ unless n is the only particle in $\{1, \ldots, N\} \setminus \text{Alive}_j$ that is killed exactly at time τ_j. In any case, one has

$$\Delta\mathbb{M}^n_{\tau_j} = -\mathbf{1}_{n \notin \text{Alive}_j} \mathbb{L}^n_{\tau_j^-}.$$

(iii) If $m \neq n$, the jumps of $\mathbb{M}_t^m$ and $\mathbb{M}_t^n$ never happen at the same time.

Proof. (i) and (ii) are direct consequences of the definitions of $\mathcal{M}$ and $\mathbb{M}$. For (iii), by construction, 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 $\tau_{n,k}$. Thus, Assumption (A) ensures that for $n \neq m$, the jumps of $\mathbb{M}_t^n$ and $\mathbb{M}_t^m$ could happen at the same time only if the latter is a branching time. However, by (ii), $\mathbb{M}_t^n$ may jump only in the case where n is the unique particle killed exactly at $\tau_{n,k}$, in which case $\mathbb{M}_t^m$ does not jump, hence (iii). □

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

Lemma 3.5. *We have the decomposition*

$$\gamma_t^N(Q) = \gamma_0^N(Q) + \frac{1}{\sqrt{N}} \int_0^t \rho_{u^-}^N (d\mathbb{M}_u + d\mathcal{M}_u). \tag{3.7}$$

Proof. Considering (1.2) and (3.2), an integration by parts yields

$$\gamma_t^N(Q) = \rho_t^N \eta_t^N(Q^{T-t}(\varphi)) = \rho_t^N \mathbb{L}_t = \gamma_0^N(Q) + \int_0^t (\rho_{u^-}^N d\mathbb{L}_u + \mathbb{L}_u d\rho_u^N),$$

where

$$\rho_t^N = \left(1 - \tfrac{K}{N}\right)^{\mathcal{B}_t} = \left(1 - \tfrac{K}{N}\right)^{\sum_{j=1}^{\infty} \mathbf{1}_{\tau_j \leqslant t}}.$$

Hence, our goal is to show that

$$\rho_{u^-}^N d\mathbb{L}_u + \mathbb{L}_u d\rho_u^N = \frac{1}{\sqrt{N}} \rho_{u^-}^N (d\mathbb{M}_u + d\mathcal{M}_u).$$

Between the branching times τ_j, the result is obviously true since ρ^N and $\mathcal{M}$ are constant processes and, by (3.5), $d\mathbb{M}_u = \sqrt{N} d\mathbb{L}_u$.

At branching time τ_j, on the one hand, we get by definition

$$\Delta\gamma_{\tau_j}^N(Q) = \rho_{\tau_j^-}^N \frac{1}{N} \sum_{n=1}^N \left((1-K/N)\mathbb{L}_{\tau_j}^n - \mathbb{L}_{\tau_j^-}^n\right) = \rho_{\tau_j^-}^N \left[(1-K/N)\mathbb{L}_{\tau_j} - \mathbb{L}_{\tau_j^-}\right].$$

On the other hand, Lemma 3.4 gives

$$\frac{1}{\sqrt{N}} \Delta\mathbb{M}_{\tau_j} = -\frac{1}{N} \sum_{n \notin \text{Alive}_j} \mathbb{L}_{\tau_j^-}^n.$$

Note that, in the latter, only the particle that is killed at time τ_j contributes to the sum. Moreover, in a similar fashion,

$$\frac{1}{\sqrt{N}} \Delta\mathcal{M}_{\tau_j} = \frac{1}{N}(1-K/N)\left(\sum_{n \notin \text{Alive}_j} \mathbb{L}_{\tau_j}^n - \frac{K}{N-K} \sum_{n \in \text{Alive}_j} \mathbb{L}_{\tau_j}^n\right). \quad (3.8)$$

By Assumption (A), $\mathbb{L}^n$ does not jump at τ_j if $n \in \text{Alive}_j$, so that

$$\begin{aligned}
\frac{1}{\sqrt{N}} \Delta\mathcal{M}_{\tau_j} &= \frac{1}{N}(1-K/N)\left(\sum_n \mathbb{L}_{\tau_j}^n - \sum_{n \in \text{Alive}_j} \mathbb{L}_{\tau_j^-}^n - \frac{K}{N-K} \sum_{n \in \text{Alive}_j} \mathbb{L}_{\tau_j^-}^n\right) \\
&= \frac{1}{N}(1-K/N) \sum_n \mathbb{L}_{\tau_j}^n - \frac{1}{N} \sum_{n \in \text{Alive}_j} \mathbb{L}_{\tau_j^-}^n \\
&= (1-K/N)\mathbb{L}_{\tau_j} - \frac{1}{N} \sum_{n \in \text{Alive}_j} \mathbb{L}_{\tau_j^-}^n.
\end{aligned}$$

Computing the sum $\Delta\mathbb{M}_{\tau_j} + \Delta\mathcal{M}_{\tau_j}$ yields the result. □

3.3 Quadratic variation analysis

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 and only if their quadratic covariation is again a local martingale.

Lemma 3.6. *Under Assumptions (A) and (B), the N^2 local martingales $\{\mathbb{M}_t^n, \mathcal{M}_t^m\}_{1\leqslant n,m\leqslant N}$ are mutually orthogonal. As a consequence,*

$$[\,\mathcal{M},\mathcal{M}\,]_t = \frac{1}{N}\sum_{n=1}^{N}[\,\mathcal{M}^n,\mathcal{M}^n\,]_t + \textit{local martingale},$$

$[\,\mathbb{M},\mathcal{M}\,]_t$ *is a local martingale, and*

$$[\,\mathbb{M},\mathbb{M}\,]_t = \frac{1}{N}\sum_{n=1}^{N}[\,\mathbb{M}^n,\mathbb{M}^n\,]_t + \textit{local martingale}.$$

In what follows, we adopt the notation

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

The jumps of $\mathbb{A}$ are controlled by

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

Proof. i) Let us show that $[\,\mathcal{M}^n,\mathcal{M}^m\,]_t$ is a local martingale for $n \neq m$. Indeed, $\mathcal{M}^n$ is piecewise constant outside the branching times so that

$$[\,\mathcal{M}^n,\mathcal{M}^m\,]_t = \sum_j \mathbf{1}_{t\geqslant\tau_j}\Delta\mathcal{M}^n_{\tau_j}\Delta\mathcal{M}^m_{\tau_j},$$

where, by definition, $\Delta\mathcal{M}^n_{\tau_j} = 0$ if $n \in \mathrm{Alive}_j$ while, otherwise,

$$\Delta\mathcal{M}^n_{\tau_j} = \Big(1-\frac{K}{N}\Big)\Big(\mathbb{L}^n_{\tau_j} - \frac{1}{N-K}\sum_{m\in\mathrm{Alive}_j}\mathbb{L}^m_{\tau_j^-}\Big)$$

which by construction has zero average conditionally on $\mathcal{F}^-_{\tau_j}$ (uniform resampling among the living particles). In the same way, $\Delta\mathcal{M}^n_{\tau_j}$ and $\Delta\mathcal{M}^m_{\tau_j}$ are independent for $n \neq m$ conditionally on $\mathcal{F}^-_{\tau_j}$, by conditional independence of the resampling of killed particles. As a consequence,

$$\mathbb{E}\left[\,\Delta\mathcal{M}^n_{\tau_j}\Delta\mathcal{M}^m_{\tau_j}\Big|\,\mathcal{F}_{\tau_j^-}\,\right] = \mathbb{E}\left[\mathbb{E}\left[\,\Delta\mathcal{M}^n_{\tau_j}\Delta\mathcal{M}^m_{\tau_j}\Big|\,\mathcal{F}^-_{\tau_j}\right]\Big|\,\mathcal{F}_{\tau_j^-}\right] = 0,$$

and Lemma 4.1 allows us to conclude the proof of Step i).

ii) We claim that $[\,\mathbb{M}^n, \mathcal{M}^m\,]_t$ is a local martingale. Since $\mathcal{M}^m$ is a pure jump martingale that only jumps at branching times, we have

$$[\,\mathbb{M}^n, \mathcal{M}^m\,]_t = \sum_j \mathbf{1}_{t \geqslant \tau_j} \Delta\mathbb{M}^n_{\tau_j} \Delta\mathcal{M}^m_{\tau_j}.$$

As explained in Lemma 3.4, $\Delta\mathbb{M}^n_{\tau_j}$ can be non zero only when particle n is the single particle killed exactly at time τ_j. Specifically, we have

$$\Delta\mathbb{M}^n_{\tau_j} = -\mathbf{1}_{n\notin \mathrm{Alive}_j} \mathbb{L}^n_{\tau_j^-},$$

which is measurable with respect to $\mathcal{F}^-_{\tau_j}$. Consequently,

$$\mathbb{E}\left[\, \Delta\mathbb{M}^n_{\tau_j} \Delta\mathcal{M}^m_{\tau_j} \middle|\, \mathcal{F}_{\tau_j^-}\,\right] = \mathbb{E}\left[-\mathbf{1}_{n\notin \mathrm{Alive}_j} \mathbb{L}^n_{\tau_j^-} \mathbb{E}\left[\, \Delta\mathcal{M}^m_{\tau_j} \middle|\, \mathcal{F}^-_{\tau_j}\right] \middle|\, \mathcal{F}_{\tau_j^-}\right] = 0,$$

and Lemma 4.1 concludes the proof of Step ii).

iii) We claim that the product $\mathbb{M}^n\mathbb{M}^m$ is a local martingale for $n \neq m$. The proof is similar to the one of Lemma 3.9 in [8]. Let us just briefly mention that it relies on the following facts: $\mathbb{M}^n$ and $\mathbb{M}^m$ are by construction independent between branching times (conditionally on the past), and never jump simultaneously, even at branching times by Lemma 3.4.

For the last point, Assumption (A) guarantees that

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

and the indicated result is now a direct consequence of (3.3) and (3.4). □

Our next objective is to calculate the quadratic variation $\frac{1}{N}\sum_n [\,\mathcal{M}^n, \mathcal{M}^n\,]$. Following (3.1), and remarking that $\mathbb{L}^n_{\tau_j} = \mathbb{L}^n_{\tau_j^-}$ for all $n \in \mathrm{Alive}_j$ by Lemma 3.4, we also adopt the upcoming notation.

Notation 3.7. *The empirical distribution of the particles that are "alive" at branching time τ_j is denoted*

$$\eta^N_{\mathrm{Alive}_j} := \frac{1}{N-K} \sum_{n\in \mathrm{Alive}_j} \delta_{X^n_{\tau_j}}.$$

Accordingly, we have

$$\mathbb{V}_{\eta^N_{\mathrm{Alive}_j}}(Q) = \mathbb{V}_{\eta^N_{\mathrm{Alive}_j}}(Q^{T-\tau_j}(\varphi)) = \frac{1}{N-K} \sum_{n\in \mathrm{Alive}_j} \left[\mathbb{L}^n_{\tau_j^-} - \tfrac{1}{N-K} \sum_{m\in \mathrm{Alive}_j} \mathbb{L}^m_{\tau_j^-} \right]^2.$$

Mutatis mutandis, $\mathbb{V}_{\eta^N_{\mathrm{Alive}_j}}(Q^2)$ is defined in the same manner.

Note that $\mathbb{V}_{\eta^N_{\text{Alive}_j}}(Q)$ and $\mathbb{V}_{\eta^N_{\text{Alive}_j}}(Q^2)$ are measurable with respect to $\mathcal{F}^-_{\tau_j}$.

Lemma 3.8. *There exists a piecewise constant local martingale $\widetilde{\mathcal{M}}_t$ with jumps at branching times, such that*

$$[\mathcal{M},\mathcal{M}]_t = \left(1-\frac{K}{N}\right)^2 \frac{K}{N} \sum_{j\geqslant 1} \mathbf{1}_{\tau_j \leqslant t} \mathbb{V}_{\eta^N_{\text{Alive}_j}}(Q) + \frac{1}{\sqrt{N}} \widetilde{\mathcal{M}}_t.$$

Since $\mathbb{V}_{\eta^N_{\text{Alive}_j}}(Q) \leqslant 2\,\|\varphi\|^2_\infty$, we deduce that

$$d\,[\mathcal{M},\mathcal{M}]_t \leqslant 2\left(1-\frac{K}{N}\right)^2 \frac{K}{N} \|\varphi\|^2_\infty d\mathcal{B}_t + \textit{local martingale}. \tag{3.9}$$

Proof. Considering the orthogonality property in Lemma 3.6, 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

$$[\mathcal{M},\mathcal{M}]_t = \frac{1}{N} \sum_{n=1}^{N} \sum_{j=1}^{+\infty} \left(\Delta \mathcal{M}^n_{\tau_j}\right)^2 \mathbf{1}_{t \geqslant \tau_j}.$$

We can then define

$$\widetilde{\mathcal{M}}_t := \frac{1}{\sqrt{N}} \sum_{n=1}^{N} \sum_{j=1}^{+\infty} \left(\left(\Delta \mathcal{M}^n_{\tau_j}\right)^2 - \mathbb{E}\left[\left(\Delta \mathcal{M}^n_{\tau_j}\right)^2 \middle| \mathcal{F}^-_{\tau_j} \right] \right) \mathbf{1}_{t \geqslant \tau_j}$$

which is indeed a local martingale by Lemma 3.4 and Lemma 4.1. Recall that $\Delta \mathcal{M}^n_{\tau_j} = 0$ if $n \in \text{Alive}_j$. Otherwise, by construction of the branching rule, we obtain

$$\begin{aligned}
\mathbb{E}\left[\left(\Delta \mathcal{M}^n_{\tau_j}\right)^2 \middle| \mathcal{F}^-_{\tau_j} \right] &= \left(1-\tfrac{K}{N}\right)^2 \tfrac{1}{N-K} \sum_{m \in \text{Alive}_j} \left(\mathbb{L}^m_{\tau_j^-} - \tfrac{1}{N-K} \sum_{l \in \text{Alive}_j} \mathbb{L}^l_{\tau_j^-} \right)^2 \\
&= \left(1-\tfrac{K}{N}\right)^2 \mathbb{V}_{\eta^N_{\text{Alive}_j}}(Q),
\end{aligned}$$

which is independent of the choice of the resampled particle n. Since there are exactly K resampled particles at time τ_j, this yields

$$\frac{1}{N} \sum_{n=1}^{N} \mathbb{E}\left[\left(\Delta \mathcal{M}^n_{\tau_j}\right)^2 \middle| \mathcal{F}^-_{\tau_j} \right] = \tfrac{K}{N}\left(1-\tfrac{K}{N}\right)^2 \mathbb{V}_{\eta^N_{\text{Alive}_j}}(Q),$$

hence the result. □

The next lemma is a crucial 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 3.5 -, with the process $t \mapsto \gamma_t^N(Q^2)$. This leads to estimates on $\mathbb{A}_t$. This idea is inspired by the fact that, by definition of the quadratic variation, and 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 3.9. *One has the decomposition*

$$d\gamma_t^N(Q^2) = \rho_{t^-}^N d\mathbb{A}_t + \frac{1}{\sqrt{N}}\rho_{t^-}^N d\widetilde{\mathbb{M}}_t, \tag{3.10}$$

where $(\widetilde{\mathbb{M}}_t)_{t\geqslant 0}$ *is a local martingale satisfying*

$$\mathbb{E}\left[\int_0^t \rho_{u^-}^N d[\widetilde{\mathbb{M}}, \widetilde{\mathbb{M}}]_u\right] \leqslant 6\|\varphi\|_\infty^4. \tag{3.11}$$

Remark 3.10. (3.10) implies that

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

Proof. Since $\mathcal{B}_t$ denotes the number of branching times until time t, it comes

$$d\rho_t^N = -\frac{K}{N}\rho_{t^-}^N d\mathcal{B}_t.$$

If $\mathcal{B}_t^n$ denotes the number of branching times of particle n until time t, we have

$$\mathcal{B}_t = \frac{1}{K}\sum_{n=1}^N \mathcal{B}_t^n$$

since, according to Lemma 3.4, exactly K particles are resampled at each branching time. We will now prove (3.10) and calculate the martingale part $\widetilde{\mathbb{M}}$. Differentiating

$$\gamma_t^N(Q^2) := \rho_t^N \frac{1}{N}\sum_{n=1}^N (\mathbb{L}_t^n)^2$$

yields

$$\begin{aligned} d\gamma_t^N(Q^2) &= \frac{1}{N}\sum_{n=1}^N \rho_{t^-}^N d\big((\mathbb{L}_t^n)^2\big) + (\mathbb{L}_t^n)^2 d\rho_t^N \\ &= \frac{1}{N}\sum_{n=1}^N \rho_{t^-}^N \left(d\big((\mathbb{L}_t^n)^2\big) - \frac{K}{N}(\mathbb{L}_t^n)^2 d\mathcal{B}_t\right). \end{aligned} \tag{3.12}$$

Next, we claim that

$$d(\mathbb{L}_t^n)^2 - \big(\mathbb{L}_t^n\big)^2 d\mathcal{B}_t^n = d[\mathbb{M}^n, \mathbb{M}^n]_t + 2\mathbb{L}_{t^-}^n d\mathbb{M}_t^n. \tag{3.13}$$

First, by definition of $\mathbb{M}^n$ (see (3.4)), $d\mathbb{M}_t^n = d\mathbb{L}_t^n - \mathbb{L}_t^n d\mathcal{B}_t^n$, so that the bilinearity of the quadratic variation gives

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

Then, using again $d\mathbb{L}_t^n = d\mathbb{M}_t^n + \mathbb{L}_t^n d\mathcal{B}_t^n$, this yields

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

which immediately simplifies into (3.13). Putting (3.12), (3.13), and the very definition of $\mathbb{A} = \frac{1}{N}\sum_n [\mathbb{M}^n, \mathbb{M}^n]$ together, we obtain

$$d\gamma_t^N(Q^2) = \rho_{t^-}^N d\mathbb{A}_t + \frac{\rho_{t^-}^N}{N} \sum_{n=1}^N \left[(\mathbb{L}_t^n)^2 \left(d\mathcal{B}_t^n - \frac{K}{N} d\mathcal{B}_t \right) + 2\mathbb{L}_{t^-}^n d\mathbb{M}_t^n \right].$$

Now, by definition of the counting processes $\mathcal{B}^n$ and $\mathcal{B}$,

$$\sum_{n=1}^N (\mathbb{L}_t^n)^2 d\mathcal{B}_t^n = \left[\sum_{n \notin \text{Alive}_t} (\mathbb{L}_t^n)^2 \right] d\mathcal{B}_t,$$

where we have used the notation

$$\text{Alive}_t := \{\text{particles that are not resampled at time } t\}.$$

As a consequence,

$$\begin{aligned} d\gamma_t^N(Q^2) =& \rho_{t^-}^N d\mathbb{A}_t + \frac{\rho_{t^-}^N}{N} \left[(1 - K/N) \sum_{n \notin \text{Alive}_t} (\mathbb{L}_t^n)^2 - \frac{K}{N} \sum_{n \in \text{Alive}_t} (\mathbb{L}_t^n)^2 \right] d\mathcal{B}_t \\ &+ \frac{\rho_{t^-}^N}{N} \sum_{n=1}^N 2\mathbb{L}_{t^-}^n d\mathbb{M}_t^n. \end{aligned}$$

Hence we see that (3.10) is satisfied with

$$d\widetilde{\mathbb{M}}_t = \frac{1}{\sqrt{N}} J_t d\mathcal{B}_t + \frac{1}{\sqrt{N}} \sum_{n=1}^{N} 2\mathbb{L}_{t^-}^n d\mathbb{M}_t^n, \tag{3.14}$$

where we have defined

$$J_t := \frac{1-K/N}{\sqrt{N}} \left[\sum_{n\notin \mathrm{Alive}_t} (\mathbb{L}_t^n)^2 - \frac{K}{N-K} \sum_{n\in \mathrm{Alive}_t} (\mathbb{L}_t^n)^2 \right].$$

It is readily seen that

$$\mathbb{E}\left[J_{\tau_j} \middle| \mathcal{F}_{\tau_j}^- \right] = 0,$$

so that, according to Lemma 4.1, $\widetilde{\mathbb{M}}$ is indeed a local martingale. Using Notation 3.7 and the fact that $\sup_{t\geqslant 0} |\mathbb{L}_{t^-}^n| \leqslant \|\varphi\|_\infty$, we also have

$$\begin{aligned} \mathbb{E}\left[J_{\tau_j}^2 \middle| \mathcal{F}_{\tau_j}^- \right] &= (1-K/N)^2 \frac{K}{N} \mathbb{V}_{\eta_{\mathrm{Alive}_j}^N}(Q^2) \\ &\leqslant 2(1-K/N)^2 \frac{K}{N} \|\varphi\|_\infty^4 . \end{aligned} \tag{3.15}$$

We can now calculate the quadratic variation of $\widetilde{\mathbb{M}}$. In the same way as in Lemma 3.5, the $(N+1)$ local martingales

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

are all orthogonal to each other. Indeed, by Lemma 3.5, $[\mathbb{M}^n, \mathbb{M}^m]$ is a martingale for any pair $n\neq m$. The only new point to check (using again Lemma 3.4) is that the quadratic covariation

$$d\left[\int \mathbb{L}_{s^-}^n d\mathbb{M}_s^n, \int J_s d\mathcal{B}_s \right]_t = -(\mathbb{L}_{t^-}^n)^2 J_t d\mathcal{B}_t$$

is indeed a local martingale, which is again a consequence of $\mathbb{E}[J_{\tau_j}|\mathcal{F}_{\tau_j}^-] = 0$ and Lemma 4.1. To establish (3.11), we apply Itô's isometry to (3.14) and use orthogonality to obtain

$$\mathbb{E}\int_0^t \rho_{u^-}^N d[\widetilde{\mathbb{M}}, \widetilde{\mathbb{M}}]_u = \mathbb{E}\left[\int_0^t \rho_{u^-}^N (J_u)^2 d\mathcal{B}_u + \frac{4}{N}\sum_{n=1}^{N} \int_0^t \rho_{u^-}^N (\mathbb{L}_{t^-}^n)^2 d[\mathbb{M}^n, \mathbb{M}^n]_u \right].$$

On the one hand, using (3.15), we get

$$\mathbb{E}\left[\int_0^t \rho_{u^-}^N (J_u)^2 d\mathcal{B}_u\right] = \mathbb{E}\left[\sum_j \mathbf{1}_{\tau_j \leqslant t} \rho_{\tau_j^-}^N \mathbb{E}\left[J_{\tau_j}^2 \,\middle|\, \mathcal{F}_{\tau_j}^- \right]\right]$$
$$= \sum_{j \geqslant 1} \mathbf{1}_{\tau_j \leqslant t} (1-K/N)^{j+1} \frac{K}{N} \mathbb{E}\left[\mathbb{V}_{\eta_{\mathrm{Alive}_j}^N}(Q^2) \right]$$
$$\leqslant 2(1-K/N)^2 \|\varphi\|_\infty^4,$$

while, on the other hand, (3.11) implies

$$\frac{4}{N}\sum_{n=1}^N \int_0^t \rho_{u^-}^N (\mathbb{L}_{t^-}^n)^2 d[\mathbb{M}^n,\mathbb{M}^n]_u \leqslant 4\|\varphi\|_\infty^2 \mathbb{E}\left[\int_0^t \rho_{u^-}^N d\mathbb{A}_u\right]$$
$$\leqslant 4\|\varphi\|_\infty^2 \mathbb{E}\left[\gamma_t^N(Q^2) - \gamma_0^N(Q^2)\right]$$
$$\leqslant 4\|\varphi\|_\infty^4.$$

Combining both inequalities yields the result.

□

3.4 $\mathbb{L}^2$-estimate

The convergence of $\gamma_T^N(\varphi)$ to $\gamma_T(\varphi)$ when N goes to infinity is now a straightforward consequence of the previous results. This kind of estimate was already noticed by Villemonais in [27] for classical Fleming-Viot particle systems (i.e., in the case where $K = 1$).

Proposition 3.11. *For any $\varphi \in \mathcal{D}$ and any $K \in [1, N]$, we have*

$$\mathbb{E}\left[\left(\gamma_T^N(\varphi) - \gamma_T(\varphi)\right)^2\right] \leqslant \frac{4\,\|\varphi\|_\infty^2}{N}.$$

Proof. Thanks to Lemma 3.5 and the fact that $\gamma_T(\varphi) = \gamma_0(Q^T\varphi)$, we have the orthogonal decomposition

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

and we can upper-bound the 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 (3.9), we obtain

$$\mathbb{E}\left[\left(\int_0^T \rho_{t^-}^N d\mathcal{M}_t\right)^2\right] = \mathbb{E}\left[\int_0^T (\rho_{t^-}^N)^2 d[\mathcal{M},\mathcal{M}]_t\right]$$
$$\leqslant 2\|\varphi\|_\infty^2 \frac{K}{N}\sum_{j=1}^{\infty}\left(1-\frac{K}{N}\right)^{2j} \leqslant 2\|\varphi\|_\infty^2.$$

(iii) $\mathbb{M}$-terms. In the same way, applying Itô's isometry and (3.10), we get

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

□

In particular, Proposition 3.11 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 $p_t^N = \gamma_t^N(\mathbf{1}_F)$ goes to its deterministic target $p_t = \gamma_t(\mathbf{1}_F)$ in probability. An interesting consequence is our first main result Proposition 2.2 that we can now justify.

Proof of Proposition 2.2. Fix $j \in [1, j_{\max}]$ and $\varepsilon > 0$. The strict monotonicity assumption ensures that

$$\delta_1 = p_{t_j-\varepsilon} - p_{t_j} > 0 \qquad \text{and} \qquad \delta_2 = p_{t_j} - p_{t_j+\varepsilon} > 0.$$

We have to prove that $\mathbb{P}(\tau_j \notin [t_j-\varepsilon, t_j+\varepsilon])$ goes to zero when N goes to infinity. Consider first the probability $\mathbb{P}(\tau_j < t_j - \varepsilon)$. We have

$$\{\tau_j < t_j - \varepsilon\} \subset \{p_{t_j-\varepsilon}^N \leqslant (1-K_N/N)^j\} \subset \{p_{t_j-\varepsilon}^N < \theta^j + \delta_1/2\},$$

for N large enough, using $1 - K_N/N \to \theta$. From Proposition 3.11, we know that $p_{t_j-\varepsilon}^N$ converges in probability to $p_{t_j-\varepsilon} = p_{t_j} + \delta_1 = \theta^j + \delta_1$, which implies that $\mathbb{P}(p_{t_j-\varepsilon}^N < \theta^j + \delta_1/2) \to 0$ as N goes to infinity. The term $\mathbb{P}(\tau_j > t_j + \varepsilon)$ is treated similarly.

For the last assertion, let $\delta = p_T - \theta^{j_{\max}+1} > 0$. Then Proposition 3.11 implies that

$$\mathbb{P}\left(\tau_{j_{\max}+1} \leqslant T\right) = \mathbb{P}(p_T^N \leqslant \theta^{j_{\max}+1}) = \mathbb{P}\left(p_T^N \leqslant p_T - \delta\right) \xrightarrow[N\to+\infty]{} 0.$$

□

Remark 3.12. An immediate consequence of Proposition 2.2 is that, if we denote by $j_{\max}^N$ the actual number of resamplings until final time T, we have, for all $\varepsilon > 0$,

$$\mathbb{P}\left(|j_{\max}^N - j_{\max}| > \varepsilon\right) = \mathbb{P}\left(j_{\max}^N \neq j_{\max}\right) \xrightarrow[N\to+\infty]{} 0.$$

3.5 Convergence of empirical measures at branching times

As for (3.2), we denote

$$\gamma_t(Q) = \gamma_t(Q^{T-t}(\varphi)) \qquad \text{and} \qquad \gamma_t(Q^2) = \gamma_t((Q^{T-t}(\varphi))^2)$$

where, again, the parameters T and φ are omitted.

Lemma 3.13. *The function $t \mapsto \gamma_t(Q^2)$ is continuous on $0 \leqslant t \leqslant T$.*

Proof. First, by Assumption (A), the distribution of the times at which the bounded martingale $M_t = Q^{T-t}(\varphi)(X_t)$ jumps is atomless, hence M_t is almost surely continuous in t, and so is M_t^2. Second, fix $0 \leqslant t \leqslant T$. By definition,

$$\gamma_t(Q^2) = \mathbb{E}\left[\mathbf{1}_{\tau_\partial > t}(Q^{T-t}(\varphi)(X_t))^2\right] = \mathbb{E}\left[(Q^{T-t}(\varphi)(X_t))^2\right] = \mathbb{E}\left[M_t^2\right],$$

and 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. □

The proof of the CLT relies on the analysis of the convergence of the quadratic variation of the martingale $\gamma^N(Q)$ when N goes to infinity. This requires to study the convergence of specific quantities related to the empirical measures at branching times, namely $\gamma_{\tau_j}^N(Q)$, $\gamma_{\tau_j^-}^N(Q)$, $\gamma_{\tau_j}^N(Q^2)$, and $\gamma_{\tau_j^-}^N(Q^2)$.

In fact, we will also need the following minor variant of $\gamma_{\tau_j^-}^N$, denoted $\gamma_{\tau_j}^{-,N}$ and defined by

$$\gamma_{\tau_j}^{-,N} := \rho_{\tau_j^-}^N \frac{1}{N} \sum_{n\in\mathrm{Alive}_j} \delta_{X_{\tau_j^-}^n} = (1 - K_N/N)^{j-1} \frac{1}{N} \sum_{n\in\mathrm{Alive}_j} \delta_{X_{\tau_j^-}^n}$$

Lemma 3.14. *For $l = 1, 2$, we have the following convergences:*

$$\gamma_{\tau_j}^N(Q^l) \xrightarrow[N\to+\infty]{\mathbb{P}} \gamma_{t_j}(Q^l),$$

$$\gamma_{\tau_j^-}^N(Q^l) \xrightarrow[N\to+\infty]{\mathbb{P}} \gamma_{t_j}(Q^l),$$

$$\gamma_{\tau_j}^{-,N}(Q^l) \xrightarrow[N\to+\infty]{\mathbb{P}} \gamma_{t_j}(Q^l).$$

Proof. We start by noting that $\gamma^N_{\tau_j^-}$ and $\gamma^{-,N}_{\tau_j}$ only differ by one Dirac measure of mass $1/N$, corresponding to the particle killed exactly at time τ_j. Therefore,

$$\gamma^N_{\tau_j^-}(Q^l) - \gamma^{-,N}_{\tau_j}(Q^l) = O(1/N),$$

and the second convergence will imply the third.

Now we consider the first convergence, with $l = 2$. Let $\varepsilon > 0$. By Lemma 3.13, we can find $\delta > 0$ such that $|\gamma_{t_j+\delta}(Q^2) - \gamma_{t_j-\delta}(Q^2)| \leqslant \varepsilon$. We consider that the event $\mathcal{A}^\delta_j = \{t_j - \delta \leqslant \tau_j \leqslant t_j + \delta\}$ is realised. By Proposition 2.2, this happens with arbitrarily large probability for N large enough. By Lemma 3.9, we have

$$\gamma^N_{\tau_j}(Q^2) - \gamma^N_{t_j-\delta}(Q^2) = \int_{t_j-\delta}^{\tau_j} \rho^N_{t^-} d\mathbb{A}_t + \frac{1}{\sqrt{N}} \int_{t_j-\delta}^{\tau_j} \rho^N_{t^-} d\widetilde{\mathbb{M}}_t.$$

Inequality (3.11) implies that the second term tends to 0 in probability. For the first one, using that $\mathbb{A}$ is increasing and again Lemma 3.9, we get

$$\int_{t_j-\delta}^{\tau_j} \rho^N_{t^-} d\mathbb{A}_t \leqslant \int_{t_j-\delta}^{t_j+\delta} \rho^N_{t^-} d\mathbb{A}_t = \gamma^N_{t_j+\delta}(Q^2) - \gamma^N_{t_j-\delta}(Q^2) + O_p(1/\sqrt{N})$$

with, by Proposition 3.11,

$$\left|\gamma^N_{t_j+\delta}(Q^2) - \gamma^N_{t_j-\delta}(Q^2) + O_p(1/\sqrt{N})\right| \xrightarrow[N\to\infty]{\mathbb{P}} \left|\gamma_{t_j+\delta}(Q^2) - \gamma_{t_j-\delta}(Q^2)\right| \leqslant 2\varepsilon.$$

We have then shown that $\gamma^N_{\tau_j}(Q^2) \to \gamma_{t_j}(Q^2)$ in probability. The second convergence for $l = 2$ is proved the same way, with τ_j^- instead of τ_j.

We consider now $l = 1$. Recall that $\gamma_{t_j+\delta}(Q) = \gamma_{t_j-\delta}(Q)$. Hence, using the same kind of arguments, we obtain, for N large enough so that $\mathbb{P}(\mathcal{A}^\delta_j) > 1-\varepsilon$,

$$\begin{aligned}
&\mathbb{E}\left[|\gamma^N_{\tau_j}(Q) - \gamma^N_{t_j-\delta}(Q)|^2\right] \\
&= 4\|\varphi\|^2_\infty \varepsilon + \mathbb{E}\left[\mathbf{1}_{\mathcal{A}^\delta_j} \frac{1}{N} \int_{t_j-\delta}^{\tau_j} (\rho^N_{s^-})^2 (d[\mathbb{M},\mathbb{M}]_s + d[\mathcal{M},\mathcal{M}]_s)\right] \\
&\leqslant 4\|\varphi\|^2_\infty \varepsilon + \mathbb{E}\left[\int_{t_j-\delta}^{t_j+\delta} (\rho^N_{s^-})^2 (d[\mathbb{M},\mathbb{M}]_s + d[\mathcal{M},\mathcal{M}]_s)\right] \\
&= 4\|\varphi\|^2_\infty \varepsilon + \mathbb{E}\left[\frac{1}{N}(\gamma^N_{t_j+\delta}(Q) - \gamma^N_{t_j-\delta}(Q))^2\right] \\
&\leqslant 4\|\varphi\|^2_\infty \varepsilon + 2\mathbb{E}\left[(\gamma^N_{t_j+\delta}(Q) - \gamma_{t_j+\delta}(Q))^2 + (\gamma^N_{t_j-\delta}(Q) - \gamma_{t_j-\delta}(Q))^2\right] \\
&\leqslant 4\|\varphi\|^2_\infty \varepsilon + \frac{16\|\varphi\|^2_\infty}{N},
\end{aligned}$$

the last inequality coming from Proposition 3.11. The latter implies the convergence in probability. The second convergence is again treated similarly, with τ_j^- instead of τ_j.

$\square$

3.6 Stretching the time

The martingale $\gamma_t^N(Q^{T-t}(\varphi))$ has a quadratic variation with both continuous time and discrete time features. In order to show a CLT for its final value $\gamma_T^N(\varphi)$, we have to apply a general CLT for martingales with jumps. The problem is that the jumps at the resampling times do not get smaller when $N \to \infty$. To circumvent this difficulty, we first show a CLT specifically tailored for our purpose.

Proposition 3.15. *Let $T > 0$ denote a fixed time horizon. For each $N \geqslant 1$, we consider on a filtered probability space the following random objects: first, a sequence of increasing stopping times τ_j, $1 \leqslant 1 \leqslant j_{\max}$ with the convention $\tau_0 = t_0 = 0$; and, second, a càdlàg local martingale $t \mapsto M_t$ that can be decomposed as the sum of two càdlàg local martingales, namely $M_t = M_0 + M_t^0 + M_t^1$, where $M_0^0 = M_0^1 = 0$, and M^1 is a pure jump martingale which jumps only at the stopping times $\tau_1, \ldots, \tau_{j_{\max}}$ with jumps of the form*

$$\Delta M_{\tau_j}^1 = \sum_{m=1}^{K_N} \Delta_j^m \qquad j = 1, \ldots j_{\max},$$

where K_N is deterministic and $(\Delta_j^m)_{1 \leqslant m \leqslant K_N}$ are integrable martingale increments with respect to a discrete filtration $\left(\mathcal{F}_j^m\right)_{0 \leqslant m \leqslant K_N}$ verifying

$$\mathcal{F}_{\tau_j^-} \subset \mathcal{F}_j^0 \subset \mathcal{F}_j^1 \subset \cdots \subset \mathcal{F}_j^{K_N} = \mathcal{F}_{\tau_j};$$

that is, for $1 \leqslant m \leqslant K_N$, Δ_j^m is $\mathcal{F}_j^m$-mesurable, and $\mathbb{E}\left[\Delta_j^m | \mathcal{F}_j^{m-1}\right] = 0$. We also assume that $M_{\tau_j}^0$ is $\mathcal{F}_j^0$ mesurable, making M^0 and M^1 orthogonal local martingales.

We then assume that all these objects satisfy the following properties:

1. *M_0 converges in distribution towards μ_0, a probability measure on $\mathbb{R}$.*
2. *For $1 \leqslant j \leqslant j_{\max}$, $\tau_j \xrightarrow[N\to\infty]{\mathbb{P}} t_j$, for some deterministic sequence $0 < t_1 < \cdots < t_j < \cdots < t_{j_{\max}} < T$.*

3. There exists a càdlàg increasing process $(v_t^N)_{0\leqslant t\leqslant T}$ *such that* $((M_t^0)^2 - v_t^N)_{0\leqslant t\leqslant T}$ *is a local martingale. There is a deterministic continuous increasing function* $v(t)$*,* $0 \leqslant t \leqslant T$*, such that* $v(0) = 0$*, and for* $0 \leqslant t \leqslant T$*,*

$$v_t^N \xrightarrow[N\to\infty]{\mathbb{P}} v(t).$$

4. We have

$$\lim_{N\to\infty} \mathbb{E}\left[\sup_{0\leqslant t\leqslant T} \left|M_t^0 - M_{t-}^0\right|^2\right] = 0,$$

and

$$\lim_{N\to\infty} \mathbb{E}\left[\sup_{0\leqslant t\leqslant T} \left|v_t^N - v_{t-}^N\right|\right] = 0.$$

5. The sequence (K_N) *goes to infinity and, for each* $1 \leqslant j \leqslant j_{\max}$*, we have*

$$\lim_{N\to\infty} \mathbb{E}\left[\max_{1\leqslant m\leqslant K_N} |\Delta_j^m|^2\right] = 0.$$

6. For each $1 \leqslant j \leqslant j_{\max}$*, there is a deterministic continuous increasing function* $\alpha \mapsto v_j(\alpha)$ *on* $[0,1]$*, such that* $v_j(0) = 0$ *and for all* $\alpha \in [0,1]$*,*

$$\sum_{m=1}^{\lfloor \alpha K_N \rfloor} |\Delta_j^m|^2 \xrightarrow[N\to\infty]{\mathbb{P}} v_j(\alpha).$$

Then, when $N \to +\infty$*, the couple* $(M_0, M_T - M_0)$ *converges in distribution towards the tensor product between* μ_0 *and a centered Gaussian variable with variance*

$$\sigma_T^2 := v(T) + \sum_{j=1}^{j_{\max}} v_j(1).$$

Proof. For simplicity, we will consider the case where $M_0 = 0$. The general case can obtained by the same reasoning as in the proof of Theorem 3.22 of [8].

We first construct a new local martingale $\boldsymbol{M}$, which coincides with M at the terminal time T, and which fulfills the assumptions of Theorem 1.4 page 339 in [17]. The idea is to keep the same martingale between the $j_{\max}$ stopping times $(\tau_1, \dots, \tau_{j_{\max}})$, and to "stretch" the time at each of the latter by inserting a time interval of length 1. Each of the additional stretched time interval of length 1 is then divided into exactly K_N sub-intervals of length $\frac{1}{K_N}$; on the latter, the new, extended martingale, is piecewise constant and performs jumps with amplitudes Δ_j^m at the times

$$\tau_j + j - 1 + \frac{m}{K_N}, \quad (j, m) \in \{1, \ldots, j_{\max}\} \times \{1, \ldots, K_N\}.$$

In the present proof (and only here), we will use the convention

$$\tau_{j_{\max}+1} = t_{j_{\max}+1} = T.$$

We can now define the new martingale as

$$\boldsymbol{M}_s = \sum_{j=1}^{j_{\max}+1} \int_{(\tau_{j-1}+j-1)\wedge s}^{(\tau_j+j-1)\wedge s} dM^0_{s-j+1} + \sum_{j=1}^{j_{\max}} \sum_{m=1}^{K_N} \Delta_j^m \mathbf{1}_{t \geqslant \tau_j + j - 1 + \frac{m}{K_N}},$$

where $0 \leqslant s \leqslant T + j_{\max}$ denotes the new time index for the time stretched processes. Formally, we introduce the càdlàg integer-valued processes

$$s \mapsto j_s^N := \inf \{j \geqslant 0, \ \tau_{j+1} + j > s\}$$

which counts the number of stopping times $(\tau_j)_{j \geqslant 1}$ that are encountered before s on the stretched time interval. Then two cases are possible. Case (i): s belongs to an inserted stretching time interval, that is

$$\tau_{j_s^N} + j_s^N - 1 + \frac{m_s^N - 1}{K_N} \leqslant s < \tau_{j_s^N} + j_s^N - 1 + \frac{m_s^N}{K_N}$$

for some $1 \leqslant m_s^N \leqslant K_N$ which defines which of the K_N sub-intervals s belongs to. We can then naturally define the original (non-stretched) time as

$$t_s^N := \tau_{j_s^N}.$$

Case (ii): s does not belong to an inserted (i.e., due to stretching) time interval, that is

$$\tau_{j_s^N} + j_s^N \leqslant s < \tau_{j_s^N+1} + j_s^N,$$

in which case we naturally set $m_s^N = K_N$ as well as

$$t_s^N := s - j_s^N.$$

In any case, we are led to

$$t_s^N = (s - j_s^N) \wedge \tau_{j_s^N}.$$

The latter obviously defines two càdlàg processes $s \mapsto m_s^N$ and $s \mapsto t_s^N$. Note that t_s^N is a $(\mathcal{F}_t)_{t \geqslant 0}$ stopping time for each $s \geqslant 0$.

We also need do define the extended filtration naturally associated with this new time, and with respect to which $\boldsymbol{M}$ is indeed a martingale, and such that the processes $s \mapsto (j_s^N, m_s^N, t_s^N)$ are adapted. This can be done by setting

$$A \in \boldsymbol{\mathcal{F}}_s \Leftrightarrow A \cap \left\{ j_s^N \leqslant j, m_s^N \leqslant m, t_s^N \leqslant t \right\} \in \mathcal{F}_t \vee \mathcal{F}_j^m \qquad \forall j \geqslant 0, m \geqslant 1, t \geqslant 0$$

or, equivalently,

$$\boldsymbol{\mathcal{F}}_s = \left[\bigvee_{j=1}^{j_{\max}+1} \mathcal{F}_{(s-j+1)\wedge\tau_j} \right] \vee \sigma(A \cap \{j_s^N \leqslant j, m_s^N \leqslant m\}, j \geqslant 0, m \geqslant 1, A \in \mathcal{F}_j^m)$$

so that, by Doob's optional sampling theorem, $\boldsymbol{M}$ is an $\boldsymbol{\mathcal{F}}$–martingale. We also remark that, on the event $\{\tau_{j_{\max}} < T\}$, we have $\boldsymbol{M}_{T+j_{\max}} = M_T$.

For $0 \leqslant s \leqslant T + j_{\max}$, we next define the large N limit of the processes (j_s^N), (t_s^N) and $\frac{m_s^N}{K_N}$, which are respectively

$$j_s := \inf \{j \geqslant 0,\ t_{j+1} + j > s\} = \sum_{j=1}^{j_{\max}} \mathbf{1}_{s \geqslant t_j + j - 1},$$

$$t_s := (s - j_s) \wedge t_{j_s},$$

and

$$m_s := (s - j_s) \wedge 1,$$

as well as the asymptotic variance by

$$c(s) = v(t_s) + \sum_{j=1}^{j_s - 1} v_j(1) + v_{j_s}(m_s).$$

It is easily checked that the limit $c(s)$ is continuous. We finally define a quadratic variation $\boldsymbol{v}^N$ for $\boldsymbol{M}$ by

$$\boldsymbol{v}_s^N = v_{t_s^N}^N + \sum_{j=1}^{j_s^N - 1} \sum_{m=1}^{K_N} (\Delta_j^m)^2 + \sum_{m=1}^{m_s^N} (\Delta_{j_s^N}^m)^2.$$

It is clear that $(\boldsymbol{M}_s)^2 - \boldsymbol{v}_s^N$ is a local martingale.

From items 1, 2, 3, 6, we can check that for all $0 \leqslant s \leqslant T + j_{\max}$, $\boldsymbol{v}_s^N$ goes to $c(s)$ in probability. More precisely, because the processes are increasing, and $t_s^N \to t_s$ in probability, for any $\delta > 0$, for N large enough, we have with arbitrarily large probability that

$$v_{t_s - \delta}^N \leqslant v_{t_s^N}^N \leqslant v_{t_s + \delta}^N,$$

with $v^N_{t_s-\delta} \to v_{t_s-\delta}$ and $v^N_{t_s+\delta} \to v_{t_s+\delta}$. By taking δ small enough, we can have $v_{t_s+\delta} - v_{t_s-\delta}$ arbitrarily small by continuity of the limit, which proves the convergence for the first term. The third term can be treated similarly, and is not detailed.

Moreover, the assumptions on the jumps in Theorem 1.4 page 339 in [17] are verified by items 4 and 5. Therefore the process $\boldsymbol{M}$ converges in distribution to a Gaussian process with variance given by $c(s)$. In particular, we have the convergence in distribution of the final time marginal $\boldsymbol{M}_{T+j_{\max}}$. To finally transfer the convergence to M_T, we write

$$|M_T - \boldsymbol{M}_{T+j_{\max}}| = \mathbf{1}_{\tau_{j_{\max}}>T}|M_T - \boldsymbol{M}_{T+j_{\max}}| \leqslant \mathbf{1}_{\tau_{j_{\max}}>T}\ 2\|M\|_\infty,$$

which converges in probability to 0 from item 2. □

3.7 Proof of the main result

In this section we use Proposition 3.15 to prove our main result Theorem 2.3. Recall that $1 - K/N$ goes to θ when N goes to infinity.

Lemma 3.16. *At each resampling time τ_j, the local martingale*

$$\int_0^t \rho^N_{u^-} d\mathcal{M}_u,$$

jumps, and each jump can be decomposed as a sum of K martingale increments Δ^m_j, for $1 \leqslant m \leqslant K$,

$$\rho^N_{\tau_j^-} \Delta\mathcal{M}_{\tau_j} = \sum_{m=1}^K \Delta^m_j.$$

Moreover, we have, for any sequence $\alpha_N \to \alpha \in (0,1]$,

$$\sum_{m=1}^{\lfloor \alpha_N K \rfloor} (\Delta^m_j)^2 \xrightarrow[N\to\infty]{\mathbb{P}} \alpha\theta^{2j}(1-\theta)\mathbb{V}_{\eta_{t_j}}(Q).$$

Proof. We first detail the proof for the simpler case $\alpha_N = 1$. We have (see (3.8) in the proof of Lemma 3.5)

$$\begin{aligned}
\rho^N_{\tau_j^-} \Delta\mathcal{M}_{\tau_j} &= \frac{1}{\sqrt{N}}(1-K/N)^{j-1}\left(\sum_{n\notin \mathrm{Alive}_j} \mathbb{L}^n_{\tau_j} - \frac{K}{N-K}\sum_{n\in \mathrm{Alive}_j} \mathbb{L}^n_{\tau_j}\right) \\
&= (1-K/N)^j \sum_{n\notin \mathrm{Alive}_j} \frac{1}{\sqrt{N}}\left(\mathbb{L}^n_{\tau_j} - \frac{1}{N-K}\sum_{n\in \mathrm{Alive}_j} \mathbb{L}^n_{\tau_j}\right).
\end{aligned}$$

It is easy to check that this last expression is a sum of K martingale increments

$$\Delta_j^m = \frac{1}{\sqrt{N}}(1-K/N)^j \left(\mathbb{L}_{\tau_j}^m - \frac{1}{N-K} \sum_{n \in \mathrm{Alive}_j} \mathbb{L}_{\tau_j}^n \right).$$

Given $\mathcal{F}_{\tau_j}^-$, it is actually a sum of i.i.d. uniformly bounded variables, and so is the sum of their squares, so that it is concentrated around its mean (e.g., by Chebyshev's inequality)

$$\sum_{m=1}^{K} (\Delta_j^m)^2 = (1-K/N)^{2j-2} \left((1-K/N)^2 \frac{K}{N} \mathbb{V}_{\eta^N_{\mathrm{Alive}_j}} + O_p(1/\sqrt{N}) \right).$$

So, using Lemma 3.14, we finally get

$$\sum_{m=1}^{K} (\Delta_j^m)^2 \xrightarrow[N\to\infty]{\mathbb{P}} \theta^{2j}(1-\theta)\mathbb{V}_{\eta_{t_j}}(Q).$$

For a general α_N, the proof is similar, except that instead of all the particles in Alive_j, we take only the $\lfloor \alpha_N K \rfloor$ first ones. Note that the chosen ordering of Alive_j is irrelevant because the new particles are i.i.d. (given $\mathcal{F}_{\tau_j}^-$). ☐

Lemma 3.17. *For* $0 \leqslant j \leqslant j_{\max}$ *and* $0 \leqslant t \leqslant T$*, we have*

$$\int_{\tau_j \wedge t}^{\tau_{j+1}\wedge t} (\rho_{u^-}^N)^2 d\mathbb{A}_u \xrightarrow[N\to\infty]{\mathbb{P}} \theta^j(\gamma_{t_{j+1}\wedge t}(Q^2) - \gamma_{t_j \wedge t}(Q^2)).$$

Proof. The proof is quite similar to that of Lemma 3.14. For $\tau_j < u \leqslant \tau_{j+1}$, we have $\rho_{u^-}^N = (1-K/N)^j$, so that

$$\int_{\tau_j \wedge t}^{\tau_{j+1}\wedge t} (\rho_{u^-}^N)^2 d\mathbb{A}_u = (1-K/N)^j \int_{\tau_j \wedge t}^{\tau_{j+1}\wedge t} \rho_{u^-}^N d\mathbb{A}_u.$$

By assumption, the deterministic factor $(1-K/N)^j$ goes to θ^j when N goes to infinity. Additionally, since $\mathbb{A}$ is increasing and $\rho_{u^-}^N > 0$, we deduce that, for any $\delta > 0$,

$$\int_{(\tau_j \wedge t)+\delta}^{(\tau_{j+1}\wedge t)-\delta} \rho_{u^-}^N d\mathbb{A}_u \leqslant \int_{\tau_j \wedge t}^{\tau_{j+1}\wedge t} \rho_{u^-}^N d\mathbb{A}_u \leqslant \int_{(\tau_j \wedge t)-\delta}^{(\tau_{j+1}\wedge t)+\delta} \rho_{u^-}^N d\mathbb{A}_u.$$

From Lemma 3.9, we have

$$\begin{aligned}
&\gamma^N_{(\tau_{j+1}\wedge t)-\delta}(Q^2) - \gamma^N_{(\tau_j \wedge t)+\delta}(Q^2) + O_p(1/\sqrt{N}) \\
&\quad \leqslant \int_{\tau_j \wedge t}^{\tau_{j+1}\wedge t} \rho_{u^-}^N d\mathbb{A}_u \leqslant \gamma^N_{(\tau_{j+1}\wedge t)+\delta}(Q^2) - \gamma^N_{(\tau_j \wedge t)-\delta}(Q^2) + O_p(1/\sqrt{N}).
\end{aligned}$$

From Proposition 2.2, for N large enough, with large probability, it then comes

$$\gamma^N_{(t_{j+1}\wedge t)-\delta}(Q^2) - \gamma^N_{(t_j\wedge t)+\delta}(Q^2) + O_p(1/\sqrt{N})$$
$$\leqslant \int_{\tau_j\wedge t}^{\tau_{j+1}\wedge t} \rho^N_{u^-}\, d\mathbb{A}_u \leqslant \gamma^N_{(t_{j+1}\wedge t)+\delta}(Q^2) - \gamma^N_{(t_j\wedge t)-\delta}(Q^2) + O_p(1/\sqrt{N}).$$

Proposition 3.11 implies

$$\gamma^N_{(t_{j+1}\wedge t)-\delta}(Q^2) - \gamma^N_{(t_j\wedge t)+\delta}(Q^2) \xrightarrow[N\to\infty]{\mathbb{P}} \gamma_{(t_{j+1}\wedge t)-\delta}(Q^2) - \gamma_{(t_j\wedge t)+\delta}(Q^2),$$

and

$$\gamma^N_{(t_{j+1}\wedge t)+\delta}(Q^2) - \gamma^N_{(t_j\wedge t)-\delta}(Q^2) \xrightarrow[N\to\infty]{\mathbb{P}} \gamma_{(t_{j+1}\wedge t)+\delta}(Q^2) - \gamma_{(t_j\wedge t)-\delta}(Q^2).$$

By continuity of the mapping $t \mapsto \gamma_t(Q^2)$, see Lemma 3.13, we can choose δ small enough such that the difference of the two limits is arbitrarily small, both being close to $\gamma_{t_{j+1}\wedge t}(Q^2) - \gamma_{t_j\wedge t}(Q^2)$. □

Proof of Theorem 2.3 Recall that

$$\gamma^N_T(\varphi) - \gamma_T(\varphi) = \Big(\gamma^N_T(Q) - \gamma^N_0(Q)\Big) + \Big(\eta^N_0(Q^T(\varphi)) - \eta_0(Q^T(\varphi))\Big),$$

where, by (3.7),

$$\sqrt{N}(\gamma^N_t(Q) - \gamma^N_0(Q)) = \int_0^t \rho^N_{u^-}(d\mathbb{M}_u + d\mathcal{M}_u).$$

It turns out that the martingale $\sqrt{N}\gamma^N_t(Q)$ does not satisfy the assumptions of Proposition 3.15 because the number of resamplings is not a priori bounded. We therefore define a new martingale by setting the initial condition $M_0 := \sqrt{N}(\eta^N_0(Q^T(\varphi)) - \eta_0(Q^T(\varphi))$ as well as

$$M_t - M_0 = \sum_{j=1}^{j_{\max}} \mathbf{1}_{\tau_j\leqslant t}\rho^N_{\tau_j^-}\Delta\mathcal{M}_{\tau_j} + \sum_{j=1}^{j_{\max}} \int_{\tau_{j-1}\wedge t}^{\tau_j\wedge t} \rho^N_{u^-} d\mathbb{M}_u + \int_{\tau_{j_{\max}}\wedge t}^{T\wedge t\wedge\tau_{j_{\max}+1}} \rho^N_{\tau_{j_{\max}}} d\mathbb{M}_u.$$

Simple algebra reveals that

$$\int_0^T \rho^N_{u^-}(d\mathbb{M}_u + d\mathcal{M}_u) - (M_T - M_0) \neq 0$$

implies that $\tau_{j_{\max}+1} \leqslant T$. But by Proposition 2.2, this happens with arbitrarily small probability, provided N is large enough, so that

$$\left| \int_0^T \rho^N_{u^-} (d\mathbb{M}_u + d\mathcal{M}_u) - (M_T - M_0) \right| \xrightarrow[N\to\infty]{\mathbb{P}} 0.$$

As a consequence, it suffices to show the CLT for $M_t = M_0 + M^0_t + M^1_t$.

The rest of the proof is devoted to show that M_t indeed satisfies the assumptions of Proposition 3.15, with

$$M^1_t = \sum_{j=1}^{j_{\max}} \mathbf{1}_{\tau_j \leqslant t} \rho^N_{\tau_j^-} \Delta\mathcal{M}_{\tau_j},$$

and

$$M^0_t = \sum_{j=1}^{j_{\max}} \int_{\tau_{j-1}\wedge t}^{\tau_j \wedge t} \rho^N_{u^-} d\mathbb{M}_u + \int_{\tau_{j_{\max}}\wedge t}^{T\wedge t\wedge \tau_{j_{\max}+1}} \rho^N_{j_{\max}} d\mathbb{M}_u.$$

For M^1_t, item 2 comes from Proposition 2.2, item 5 is from the construction of the particle system and the fact that each little jump is of order $1/\sqrt{N}$, and item 6 is from Lemma 3.16.

For M^0_t, we define the increasing process v^N_t as

$$v^N_t = \sum_{j=1}^{j_{\max}} \int_{\tau_{j-1}\wedge t}^{\tau_j \wedge t} (\rho^N_{u^-})^2 d\mathbb{A}_u + \int_{\tau_{j_{\max}}\wedge t}^{T\wedge t\wedge \tau_{j_{\max}+1}} (\rho^N_{j_{\max}})^2 d\mathbb{A}_u.$$

The fact that $(M^0_t)^2 - v^N_t$ is a local martingale is from Lemma 3.6. Item 3 is from Lemma 3.17, item 4 from Lemma 3.4 (ii) and (iii) (we should not forget the $1/\sqrt{N}$ factor in the definition of $\mathbb{M}$ in equation (3.5)), and Lemma 3.6. The orthogonality between M^0 and M^1 is from Lemma 3.6.

The convergence of the initial condition $M_0 := \sqrt{N}(\eta^N_0(Q^T(\varphi)) - \eta_0(Q^T(\varphi))$ is the usual CLT.

We can then apply Proposition 3.15. Lemmas 3.16 and 3.17 imply that the total asymptotic variance of $\sqrt{N}\big(\gamma^N_T(\varphi) - \gamma_T(\varphi)\big)$ is given by

$$\begin{aligned}\sigma^2_T(\varphi) = \mathbb{V}_{\eta_0}(Q) &+ \sum_{j=1}^{j_{\max}} \theta^{2j}(1-\theta)\mathbb{V}_{\eta_{t_j}}(Q) \\ &+ \sum_{j=0}^{j_{\max}-1} \theta^j(\gamma_{t_{j+1}}(Q^2) - \gamma_{t_j}(Q^2)) + \theta^{j_{\max}}(\gamma_T(Q^2) - \gamma_{t_{j_{\max}}}(Q^2)).\end{aligned}$$

We recall that by definition $\eta_t = \gamma_t/\rho_t$ and that $\rho_t = \theta^j$ for $\theta^j \leqslant t < \theta^{j+1}$, so that we can rewrite the asymptotic variance as

$$\sigma_T^2(\varphi) = \eta_0(Q^2) - \eta_0(Q)^2 + \sum_{j=1}^{j_{\max}} \theta^{2j}(1-\theta)\big(\eta_{t_j}(Q^2) - \eta_{t_j}(Q)^2\big)$$
$$+ \sum_{j=0}^{j_{\max}-1} \theta^{2j}(\theta\eta_{t_{j+1}}(Q^2) - \eta_{t_j}(Q^2)) + \theta^{2j_{\max}}(\eta_T(Q^2) - \eta_{t_{j_{\max}}}(Q^2)).$$

This may be reformulated as

$$\sigma_T^2(\varphi) = \eta_0(Q^2) - \eta_0(Q)^2 + \sum_{j=0}^{j_{\max}-1} \theta^{2j+1}\eta_{t_{j+1}}(Q^2) + \theta^{2j_{\max}}\eta_T(Q^2) - \eta_0(Q^2)$$
$$- \sum_{j=1}^{j_{\max}} \theta^{2j+1}\eta_{t_j}(Q^2) - \sum_{j=1}^{j_{\max}}(1-\theta)\theta^{2j}\eta_{t_j}(Q)^2,$$
$$= \theta^{2j_{\max}}\mathbb{V}_{\eta_T}(Q) + \sum_{j=1}^{j_{\max}} \big(\theta^{2j-1} - \theta^{2j+1}\big)\eta_{t_j}(Q^2) - \sum_{j=1}^{j_{\max}} \theta^{2j}(1-\theta)\eta_{t_j}(Q)^2, \tag{3.16}$$

where, in the last line, we have used that by definition $\eta_T(Q^2) = \eta_T(\varphi^2)$ and $\eta_0(Q) = \gamma_T(\varphi) = \theta^{j_{\max}}\eta_T(\varphi)$. Finally, remarking that

$$\theta^{2j}(1-\theta) = (\theta^{2j-1} - \theta^{2j+1}) - \theta^{2j}(1/\theta - 1),$$

as well as

$$\eta_{t_j}(Q) = \gamma_{t_j}(Q)\theta^{-j} = \gamma_T(\varphi)\theta^{-j} = \eta_T(\varphi)\theta^{j_{\max}-j},$$

we conclude that

$$\sigma_T^2(\varphi) = \theta^{2j_{\max}}(\mathbb{V}_{\eta_T}(\varphi) + j_{\max}(1/\theta - 1)\eta_T(\varphi)^2) + \sum_{j=1}^{j_{\max}} \big(\theta^{2j-1} - \theta^{2j+1}\big)\mathbb{V}_{\eta_{t_j}}(Q).$$

Hence we have proved Theorem 2.3 for any test function φ in $\mathcal{D}$. To see that the result is still valid for any φ in $\overline{\mathcal{D}}$, it suffices to apply the same reasoning as in [8].

4 Supplementary material

4.1 Another formulation of the asymptotic variance

As mentioned in [7], it turns out that it is possible to make a connection between Fleming-Viot particle systems and interacting particle systems as exposed for example in the pair of books [15, 16]. Without going into details, we will just show that our asymptotic variance coincides with the one given in [15] page 452. As already noticed in [7], we need to use predicted measures instead of corrected ones. At each t_k, we denote by $\tilde{\eta}_{t_k}$ the predicted measure, that is $\eta_{t_{k-1}} Q^{t_k - t_{k-1}}$. We have $\tilde{\eta}_{t_k} = \theta \eta_{t_k} + (1-\theta)\delta_\partial$. For any test function φ such that $\varphi(\partial) = 0$, we have $\tilde{\eta}_{t_k}(\varphi) = \theta \eta_{t_k}(\varphi)$. Note also that $\eta_T = \tilde{\eta}_T$ since there is no resampling at the end.

We start from (3.16) and remark that

$$\sum_{j=1}^{j_{\max}} \theta^{2j} \eta_{t_j}(Q)^2 = \sum_{j=1}^{j_{\max}} \theta^{2(j-1)} \tilde{\eta}_{t_j}(Q)^2$$

to get, with the convention $t_{j_{\max}+1} = T$,

$$\begin{aligned}
\sigma_T^2(\varphi) &= \theta^{2j_{\max}} \mathbb{V}_{\tilde{\eta}_T}(Q) + \sum_{j=1}^{j_{\max}} \theta^{2j-2} \mathbb{V}_{\tilde{\eta}_{t_j}}(Q) - \sum_{j=1}^{j_{\max}} \theta^{2j} \theta \eta_{t_j}(Q^2) + \sum_{j=1}^{j_{\max}} \theta^{2j} \theta \eta_{t_j}(Q)^2 \\
&= \sum_{j=0}^{j_{\max}} \theta^{2j} \mathbb{V}_{\tilde{\eta}_{t_{j+1}}}(Q) - \sum_{j=1}^{j_{\max}} \theta^{2j+1} \mathbb{V}_{\eta_{t_j}}(Q).
\end{aligned}$$

Now, we observe that

$$\theta \mathbb{V}_{\eta_{t_j}}(Q) = \tilde{\eta}_{t_j}(Q^2) - \theta \tilde{\eta}_{t_{j+1}}(Q)^2 = \tilde{\eta}_{t_j}(\mathbf{1}_F (Q - \tilde{\eta}_{t_{j+1}}(Q))^2),$$

so that

$$\begin{aligned}
\sigma_T^2(\varphi) &= \sum_{j=0}^{j_{\max}} \theta^{2j} \mathbb{V}_{\tilde{\eta}_{t_{j+1}}}(Q) - \sum_{j=1}^{j_{\max}} \theta^{2j} \tilde{\eta}_{t_j}(\mathbf{1}_F (Q - \tilde{\eta}_{t_{j+1}}(Q))^2) \\
&= \sum_{j=1}^{j_{\max}+1} \theta^{2(j-1)} \mathbb{V}_{\tilde{\eta}_{t_j}}(Q) - \sum_{j=2}^{j_{\max}+1} \theta^{2(j-1)} \tilde{\eta}_{t_{j-1}}(\mathbf{1}_F (Q - \tilde{\eta}_{t_j}(Q))^2).
\end{aligned}$$

This is the result given in [15] page 452.

4.2 Stopping times and martingales

Lemma 4.1. *Let τ be a stopping time on a filtered probability space, and U an integrable and $\mathcal{F}_\tau$ measurable random variable such that $\mathbb{E}\left[\, U | \mathcal{F}_{\tau^-} \,\right] = 0$. 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

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

Next, by definition of $\mathcal{F}_{\tau^-}$, $\mathbb{E}\left[U\mathbf{1}_{t \geqslant \tau} | \mathcal{F}_s \right] \mathbf{1}_{s < \tau}$ and $\mathbf{1}_{t \geqslant \tau}$ are $\mathcal{F}_{\tau^-}$-measurable, hence the result follows from

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

□

Acknowledgements. We thank the referee for valuable comments and insightful suggestions, which led to a substantial improvement of the paper.

References

[1] S.K. Au and J.L. Beck. Estimation of small failure probabilities in high dimensions by subset simulation. *Probabilistic Engineering Mechanics*, 16(4):263–277, 2001.

[2] S.K. Au and J.L. Beck. Subset simulation and its application to seismic risk based on dynamic analysis. *Journal of Engineering Mechanics*, 129(8):901–917, 2003.

[3] 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.

[4] C.-E. Bréhier and T. Lelièvre. On a new class of score functions to estimate tail probabilities of some stochastic processes with adaptive multilevel splitting. *Chaos*, 29(3):033126, 13, 2019.

[5] 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.

[6] F. Cérou, P. Del Moral, T. Furon, and A. Guyader. Sequential Monte Carlo for Rare Event Estimation. *Stat. Comput.*, 22(3):795–808, 2012.

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

[8] F. Cérou, B. Delyon, A. Guyader, and M. Rousset. A Central Limit Theorem for Fleming-Viot Particle Systems with Hard Killing. *arXiv preprint arXiv:1709.06771*, 2017.

[9] F. Cérou, B. Delyon, A. Guyader, and M. Rousset. On the Asymptotic Normality of Adaptive Multilevel Splitting. *SIAM/ASA J. Uncertain. Quantif.*, 7(1):1–30, 2019.

[10] F. Cérou, T. Furon, and A. Guyader. Experimental assessment of the reliability for watermarking and fingerprinting schemes. *EURASIP J. Inf. Secur.*, 2008:6:1–6:12, January 2008.

[11] F. Cérou and A. Guyader. Adaptive Multilevel Splitting for Rare Event Analysis. *Stoch. Anal. Appl.*, 25(2):417–443, 2007.

[12] F. Cérou and A. Guyader. Fluctuation Analysis of Adaptive Multilevel Splitting. *Ann. Appl. Probab.*, 26(6):3319–3380, 2016.

[13] F. Cérou, A. Guyader, T. Lelièvre, and D. Pommier. A Multiple Replica Approach to Simulate Reactive Trajectories. *The Journal of Chemical Physics*, 134(5):054108, 2011.

[14] F. Cérou, A. Guyader, and M. Rousset. Adaptive Multilevel Splitting: Historical Perspective and Recent Results. *Chaos*, 29(4):043108, 12, 2019.

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

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

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

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

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

[20] A. Guyader, N. Hengartner, and E. Matzner-Løber. Simulation and estimation of extreme quantiles and extreme probabilities. *Applied Mathematics and Optimization*, 64:171–196, 2011.

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

[22] T. Lestang, F. Ragone, C.-E. Bréhier, C. Herbert, and F. Bouchet. Computing return times or return periods with rare event algorithms. *Journal of Statistical Mechanics: Theory and Experiment*, 2018(4):043213, 2018.

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

[24] L.J.S. Lopes, C.G. Mayne, C. Chipot, and T. Lelièvre. Adaptive Multilevel Splitting Method: Isomerization of the alanine dipeptide. *arXiv preprint arXiv:1707.00950*, 2017.

[25] P.E. Protter. *Stochastic integration and differential equations*. Springer-Verlag, Berlin, 2005.

[26] I. Teo, C.G. Mayne, K. Schulten, and T. Lelièvre. Adaptive multilevel splitting method for molecular dynamics calculation of benzamidine-trypsin dissociation time. *Journal of chemical theory and computation*, 12(6):2983–2989, 2016.

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