SIAM/ASA J. Uncertainty Quantification
Vol. 7, No. 1, pp. 1–30

On the Asymptotic Normality of Adaptive Multilevel Splitting*

Frédéric Cérou†, Bernard Delyon‡, Arnaud Guyader§, and Mathias Rousset†

Abstract. Adaptive multilevel splitting (AMS) is a generic Monte Carlo method for Markov processes that simulates rare events and estimates associated probabilities. Despite its practical efficiency, there are almost no theoretical results on the convergence of this algorithm. The purpose of this paper is to prove both consistency and asymptotic normality results in a general setting. This is done by associating to the original Markov process a level-indexed process, also called a stochastic wave, and by showing that AMS can then be seen as a Fleming–Viot type particle system. This being done, we can finally apply general results on Fleming–Viot particle systems that we have recently obtained.

Key words. sequential Monte Carlo, Fleming–Viot particle systems, rare events simulation

AMS subject classifications. 82C22, 65C05, 60K35, 60J60

DOI. 10.1137/18M1187477

1. Introduction.

In this article, we prove asymptotic results for the adaptive multilevel splitting (AMS) algorithm used to estimate rare events or to simulate conditionally on rare events. This method belongs to the family of importance splitting algorithms, a set of techniques that date back to Kahn and Harris [17] and Rosenbluth and Rosenbluth [23] to analyze particle transmission energies and molecular polymer conformations. The adaptive version of this method was proposed in [8]. Here we consider the last particle version of this algorithm, introduced in [14] and presented in [9] in the context of molecular dynamics. Recently, this algorithm was successfully applied to real-world chemical computations in [25] as well as to Monte Carlo particle transport problems [19].

To the best of our knowledge, there are almost no theoretical results on the convergence of this algorithm, with the notable exception of the idealized case [1, 3, 4]. Note, however, that estimators of unnormalized averages are known to be unbiased in wide generality (see [2]). Under regularity assumptions discussed below, we give an L^2-estimate as well as a central limit theorem (CLT). In both cases, we consider the real algorithm and do not consider the idealized case. We also discuss the asymptotic variance given by the CLT.

The general framework is as follows. Given a stopped Markov process $(Y_s)_{s\geq 0}$ in a space

*Received by the editors May 15, 2018; accepted for publication (in revised form) October 8, 2018; published electronically January 2, 2019.

http://www.siam.org/journals/juq/7-1/M118747.html

Funding: This work was partially supported by the French Agence Nationale de la Recherche, under grant ANR-14-CE23-0012, and by the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement number 614492.

†INRIA–Rennes & Université de Rennes, CNRS, IRMAR - UMR 6625, F-35000 Rennes, France (frederic.cerou@inria.fr, mathias.rousset@inria.fr).

‡Université de Rennes, CNRS, IRMAR - UMR 6625, F-35000 Rennes, France (bernard.delyon@univ-rennes1.fr).

§LPSM, Sorbonne Université, 75005 Paris, France, and CERMICS, École des Ponts ParisTech, 77455 Marne la Vallée, France (arnaud.guyader@upmc.fr).

E and a function $\xi : E \to \mathbb{R}$ such that $\xi(Y_0) = 0$ almost surely, the goal is to compute the probability that $\sup_s \xi(Y_s) > 1$ (the rare event), and compute the distribution of Y given that $\sup_s \xi(Y_s) > 1$. In this context, AMS is constructed as an interacting particle system consisting of N particles/trajectories $(Y^n)_{n=1\dots N}$ simulated according to the distribution of the underlying process Y. Define the score of a particle/trajectory as $\sup_s \xi(Y_s)$. At each iteration, the particle with minimal score, say τ, is killed. Another particle, uniformly and randomly chosen, is then cloned so that the number of particles/trajectories remains constant and equal to N. Its trajectory is also resampled from the first entrance time of the level set $\{\xi > \tau\}$. Finally, the algorithm is stopped as soon as all particles have reached the level set $\{\xi > 1\}$. The probability is estimated by $(1 - 1/N)^J$, where J denotes the total number of iterations, and the final empirical distribution of particles is an estimator of the law of Y conditioned by the event $\{\sup_s \xi(Y_s) > 1\}$.

The main result of this paper is a CLT on the latter estimators in the large population limit, that is, when N goes to infinity. This CLT heavily relies on a CLT for Fleming–Viot particle systems that we have recently obtained [6]. The key point here is that the AMS algorithm can be recast as a Fleming–Viot particle system by introducing a *level-indexed process*, also called a stochastic wave in [11], associated to the pair (Y, ξ). The latter is obtained through a discontinuous *time change*, where the levels induced by ξ play the role of a new time parameter, and the associated particle state is given by the first entrance in successive level sets. Details can be found in section 3.

The CLT is obtained for diffusions under three main assumptions (referred to as Assumptions 1, 2, and 3) on the pair (Y, ξ). These assumptions include the case where Y is a diffusion in $\mathbb{R}^d$ satisfying a stochastic differential equation (SDE) of the form

$$\mathrm{d}Y_s = b(Y_s)\,\mathrm{d}s + \sigma(Y_s)\,\mathrm{d}W_s, \tag{1.1}$$

with smooth coefficients (b, σ), and ξ is a smooth function with compact level sets satisfying everywhere some nondegeneracy condition of the form $(\nabla\xi)^T\sigma \neq 0$.

In particular, as explained in [25], this algorithm is applied in computational chemistry to simulate so-called "reactive trajectories," including those for real-world chemical applications. To fix ideas, consider an overdamped system solution to the SDE

$$\mathrm{d}Y_s = -\nabla V(Y_s)\,\mathrm{d}s + \sqrt{2\beta^{-1}}\,\mathrm{d}W_s,$$

where V is the interaction energy of the system and β^{-1} is the temperature. Then, let $A \subset \{\xi < 0\}$ denote a set of initial configurations, typically a "metastable" state defined by a sufficiently thin energy level set around a local minimum of V. In this chemical context, ξ is called a "reaction coordinate" and parametrizes a chemical reaction starting from an initial configuration modeled by A up to a final configuration modeled by the level set $\{\xi > 1\}$ (see Figure 1). The associated "reactive trajectory" is then defined by an initial Y_0 close to but distinct from A and by the conditional distribution $\mathcal{L}\{(Y_s)_{s\geq 0} | S_1 < S_A\}$ of trajectories reaching the final configuration set $\{\xi > 1\}$ (at a time denoted S_1) before reaching the initial configuration set A (at a time denoted S_A). The associated event $\{S_1 < S_A\}$ represents the rare event of interest. In particular, the associated mean time $\mathbb{E}[S_1 | S_1 < S_A]$ and the probability $\mathbb{P}(S_1 < S_A)$ are crucial for the estimation of the underlying chemical kinetics. It

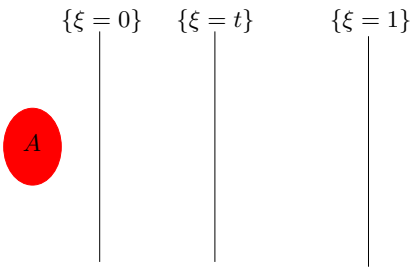

Figure 1. *The metastable state A and the reaction coordinate $\xi(y_1, y_2) = y_1$.*

turns out that AMS is particularly efficient at providing estimators of such quantities, and as stated above, the aim of this paper is to prove a large population CLT for the latter. The interested reader can find details and simulations in [25].

As mentioned above, our core idea consists of reformulating the AMS algorithm as a Fleming–Viot particle system and then applying the CLT for Fleming–Viot particle systems of [6]. This is in fact a quite generic method that may be applied to other types of underlying processes Y, for instance, to diffusions with degenerate condition $(\nabla\xi)^T\sigma = 0$, or to piecewise deterministic Markov processes. However, in these cases, defining the associated level-indexed process and checking the assumptions of the CLT for the Fleming–Viot particle system requires extensive, specific analysis that is left for future work.

The paper is organized as follows. In section 2, we introduce the AMS algorithm and state the main result of the article, namely Theorem 2.7. As mentioned above, the assumptions for the latter are illustrated in the diffusive case. In section 3, we reformulate the AMS algorithm as a Fleming–Viot particle system built from a so-called level-indexed process. Finally, applying [6], this allows us to establish the desired result. Most of the proofs and technical results are gathered in the appendices.

2. Setting, algorithm, and main result.

2.1. Setting. Let E denote a Polish state space. If $\xi : E \to \mathbb{R}$ is a measurable function and I is a subset of $\mathbb{R}$, we denote

$$\{\xi \in I\} = \xi^{-1}(I) = \{y \in E,\ \xi(y) \in I\}\,.$$

Additionally, if μ is a Borel probability measure on E, Y is an E-valued random variable with law μ, and $\varphi : E \to \mathbb{R}$ is a test function, we write

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

Let $(Y_s)_{s\geq 0}$ denote a time homogeneous Markov process with continuous trajectories in E that may be defined from any initial condition $y_0 \in E$. We also assume that the mapping $\xi : E \to \mathbb{R}$, called a level function, is continuous. In what follows, we suppose for simplicity that the law $\eta_0 := \mathcal{L}(Y_0)$ is supported by the level set $\{\xi = 0\}$, meaning that

$$\eta_0(\xi = 0) = 1. \tag{2.1}$$

For each $t \geq 0$, we denote the first entrance time in levels strictly greater than t by

$$S_t := \inf \{s \geq 0,\ \xi(Y_s) > t\} \in [0, +\infty].$$

Note that by continuity of ξ and Y, for all $t \geq 0$ with $S_t < +\infty$, we have

$$\xi(Y_{S_t}) = t. \tag{2.2}$$

Let A denote a Borel set in E. By convention, in what follows, the process $(Y_s)_{s\geq 0}$ is stopped at the random time $S_1 \wedge S_A$, where

$$S_A := \inf \{s \geq 0,\ Y_s \in A\} \in [0, +\infty].$$

Assuming that

$$p_1 := \mathbb{P}(S_1 < S_A) > 0,$$

the goal of the algorithm is to estimate p_1 as well as the conditional distribution $\mathcal{L}(Y_{S_1}|S_1 < S_A)$.

Specific algorithms have been developed in order to efficiently simulate such events, especially when they are rare. The upcoming section recalls the last particle version of the AMS algorithm as introduced in [9]. The main goal of this paper is to prove the consistency and the asymptotic normality of this algorithm.

For simplicity, we will assume that almost surely,

$$S_1 \wedge S_A < +\infty, \tag{2.3}$$

which implies that the particle trajectories defined in the AMS algorithm are all defined on finite time intervals. While removing or modifying condition (2.1) and, especially, condition (2.2) requires substantial changes in the definition of the level-indexed process in section 3, the condition (2.3) is merely technical and can be simply removed up to dealing with infinite length trajectories (see Appendix E).

2.2. Adaptive multilevel splitting. From here on, the integer N denotes the number of trajectories, also called particles. This sample size will stay unchanged all along the algorithm. Additionally, random variables denoted with the superscripts n, j, for instance, $Z^{n,j}$, means that it concerns the trajectory with index n at iteration $j \geq 0$. AMS is given by Algorithm 2.1. Figure 2 illustrates the first two steps of Algorithm 2.1 in the case where $N = 3$.

Assumption 2 below will ensure that almost surely, in the last step of Algorithm 2.1,

$$\forall h > 0, \qquad \sup_{s\in[\sigma_j,\sigma_j+h]} \xi(Y_s^{N_j,j}) > \xi(Y_{\sigma_j}^{N_j,j}).$$

In particular, this implies that the sequence $(\tau_j)_{j\geq 0}$ is strictly increasing. Moreover, Assumption 3 below will imply that this algorithm stops after a finite number of iterations almost surely (see Proposition 2.5).

For any $t \in [0, 1]$, let us denote by J_t the number of branchings of this algorithm between level 0 and level t, that is,

$$J_t := \sup \{j,\ \tau_j \leq t\},$$

which by definition satisfies

$$\tau_{J_t} \leq t < \tau_{1+J_t}.$$

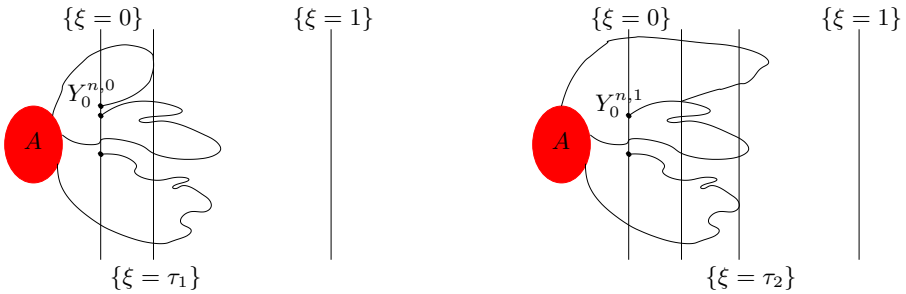

Figure 2. *The first two steps of AMS with $N = 3$ trajectories.*

Algorithm 2.1. Adaptive multilevel splitting.

We start with a sample of the initial condition of the process Y, which means that

$$Y_0^{1,0}, \dots, Y_0^{N,0} \overset{\text{i.i.d.}}{\sim} \eta_0. \tag{2.4}$$

From each initial condition $Y_0^{n,0}$, we simulate a trajectory $(Y_s^{n,0})_{s\geq 0}$. We recall that the latter is stopped when hitting A or level set $\{\xi > 1\}$. Set $\tau_0 = 0$ and then iterate on $j \geq 1$:

(i) For $1 \leq n \leq N$, compute the score of each particle, meaning the supremum of the level ξ along each particle's trajectory:

$$\sup_{0\leq s\leq S_A^{n,j-1}\wedge S_1^{n,j-1}} \xi(Y_s^{n,j-1}).$$

Find the particle with the smallest score:

$$\begin{cases} N_j := \arg\min_{n=1,\dots,N} \sup_{0\leq s\leq S_A^{n,j-1}\wedge S_1^{n,j-1}} \xi(Y_s^{n,j-1}), \\ \tau_j := \sup_{0\leq s\leq S_A^{N_j,j-1}\wedge S_1^{N_j,j-1}} \xi(Y_s^{N_j,j-1}). \end{cases} \tag{2.5}$$

Under Assumptions 1 and 2 below, a unique particle satisfies (2.5) (see Proposition 2.5).

(ii) Stop the algorithm if $\tau_j = 1$.

(iii) For $n \neq N_j$, set $(Y_s^{n,j})_{s\geq 0} = (Y_s^{n,j-1})_{s\geq 0}$.

(iv) Pick an index M_j uniformly at random in $\{1, \dots, N\} \setminus \{N_j\}$. Replace the trajectory with index N_j by a resampled version of the trajectory with index M_j, starting from the hitting time of level τ_j, that is

- set $\sigma_j := \inf\{s \geq 0,\ \xi(Y_s^{M_j,j}) > \tau_j\} < +\infty$;
- for $s < \sigma_j$, set $Y_s^{N_j,j} = Y_s^{M_j,j}$;
- for $s \geq \sigma_j$, simulate a new piece of trajectory $(Y_s^{N_j,j})_{s\geq\sigma_j}$ according to the law of the underlying process $(Y_s)_{s\geq 0}$ with initial condition $Y_{\sigma_j}^{M_j,j}$.

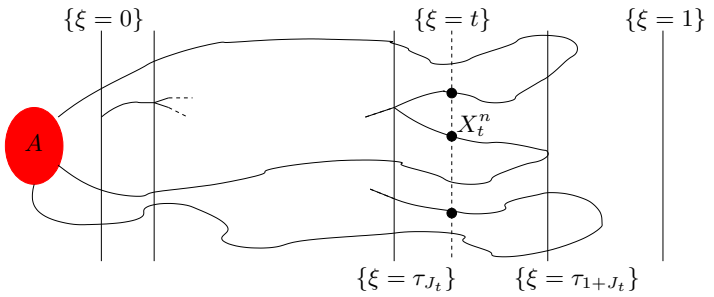

Figure 3. *The entrance points X_t^n of the level set $\{\xi = t\}$.*

Accordingly, the value of j at the end of the algorithm is J_1, and the final particle system is given by the N trajectories $(Y_s^{n,J_1})_{s\geq 0}$, $1 \leq n \leq N$. By construction, and by Proposition 2.5 below, all of these trajectories reach the level set $\{\xi = 1\}$ and are stopped at this specific time.

Similarly, for a given level $t \in [0,1]$, the particle trajectories after J_t iterations are given by $(Y_s^{n,J_t})_{s\geq 0}$, and the associated entrance times are

$$S_t^n := \inf\{s \geq 0,\ \xi(Y_s^{n,J_t}) > t\},$$

with entrance states $Y_{S_t^n}^{n,J_t}$. To lighten the notation and to prepare the definition of the level-indexed process in section 3, we denote the latter states (see Figure 3) by

$$X_t^n := Y_{S_t^n}^{n,J_t}.$$

Thus, by (2.2) one has $\xi(X_t^n) = t$. Then, for any test function φ, we estimate the law of Y_{S_t} given that $Y_{S_t} < Y_{S_A}$ by the empirical distribution

$$\eta_t^N(\varphi) := \frac{1}{N}\sum_{n=1}^N \varphi(Y_{S_t^n}^{n,J_t}) = \frac{1}{N}\sum_{n=1}^N \varphi(X_t^n).$$

In the same vein, since exactly one trajectory is resampled at each step of the algorithm, our estimator for $p_t = \mathbb{P}(Y_{S_t} < Y_{S_A})$ is

$$p_t^N := \left(1 - \frac{1}{N}\right)^{J_t}.$$

It was already established in [2] (see also [6] and the discussion in section 3 of the present article) that $p_t^N \times \eta_t^N(\varphi)$ is in fact an unbiased estimator:

$$\mathbb{E}\left[p_t^N \times \eta_t^N(\varphi)\right] = \mathbb{E}\left[\varphi(Y_{S_t})\mathbf{1}_{S_t<S_A}\right].$$

The unbiasedness of the probability estimate is important if we want to use parallel computing. Although the algorithm is intrinsically sequential, it is possible to run several realizations in parallel and then take the mean of the estimates. More complete discussions can be found in [26] about various parallel versions of AMS, and in [2] about unbiased generalizations. All of these versions are not included in the result of the present paper, which is restricted to the last particle variant. Obtaining a CLT for them may be possible but yet requires to extend accordingly the CLT for Fleming–Viot particle systems of [6].

2.3. Assumptions. In this section, we gather some sufficient conditions to ensure the well-posedness of the previous algorithm and to obtain the main results of section 2.4. We illustrate these assumptions in the case of a strong solution of an SDE with smooth coefficients.

Let us begin with some topological and regularity conditions.

Assumption 1 (Feller regularity). *E is a locally compact state space, ξ is continuous, $\eta_0(\xi = 0) = 1$, $A \subset \{\xi < 0\}$, and $(Y_s)_{s\geq 0}$ is a Feller diffusion process, i.e., a Feller process with continuous trajectories.*

Remark 2.2. We recall that Feller processes are strong Markov with respect to their natural filtration denoted $\left(\mathcal{F}_s^Y = \sigma\left(Y_{s'}, 0 \leq s' \leq s\right)\right)_{s\geq 0}$, the latter being necessarily right-continuous (see, for example, [12, Thm. 2.7, p. 169]).

For the next assumption, we recall the notation $S_t := \inf\{s \geq 0,\ \xi(Y_s) > t\}$ as well as $S_B := \inf\{s \geq 0,\ Y_s \in B\}$ for any set $B \subset E$. Besides, $\mathring{A}$ and $\bar{A}$ denote respectively the interior and the closure of the set A.

Assumption 2 (almost sure strict entrance). *For any $t \in [0,1]$ and $y \in E$ such that $\xi(y) = t$,*

$$\mathbb{P}_y\left(S_t = 0\right) = 1. \tag{2.6}$$

In the same way, for all $y \in \{\xi = 0\}$,

$$\mathbb{P}_y\left(S_{\bar{A}} = S_{\mathring{A}}\right) = 1. \tag{2.7}$$

By the strong Markov property, (2.6) ensures that S_t, defined as the first entrance time in levels strictly greater than t, is in fact equal to the hitting time of level t, that is,

$$S_t = \inf\left\{s \geq 0,\ \xi(Y_s) = t\right\}.$$

Additionally, since the process Y has continuous trajectories, (2.7) obviously implies that for all $y \in \{0 \leq \xi \leq 1\}$,

$$\mathbb{P}_y\left(S_{\bar{A}} = S_{\mathring{A}}\right) = 1. \tag{2.8}$$

Moreover, we will show in Lemma D.1 that (2.7) and the strong Markov property imply that the jump times of the càdlàg process $t \mapsto Y_{S_t}$ have atomless distributions. This property is indeed required in [6] in order to get the CLT for Fleming–Viot particle systems.

Let us now define the integral operator

$$q(\varphi)(y) := \mathbb{E}_y\left[\varphi(Y_{S_1})\mathbf{1}_{S_1<S_A}\right], \qquad y \in \{0 \leq \xi \leq 1\}. \tag{2.9}$$

Denoting by $C_b(\{\xi = 1\})$ the set of continuous and bounded functions on the level set $\{\xi = 1\}$, we will prove in Lemma D.3 that under Assumptions 1 and 2, if $\varphi \in C_b(\{\xi = 1\})$, then $q(\varphi)$ is bounded and continuous on $\{0 \leq \xi \leq 1\}$. The proof is based on a general result given in Appendix A, namely Lemma A.4. The integral operator q will prove crucial in the remainder of the paper as it will appear in the asymptotic variance of the CLT.

Our next assumption ensures a uniform control on the probabilities of success, namely $\mathbb{P}_y\left(S_1 < S_A\right)$, with respect to the initial condition.

Assumption 3 (uniform positive probability of reaching the last level). *We assume that almost surely,*

$$S_A \wedge S_1 < +\infty,$$

and that

$$\inf_{y \in \{\xi = 0\}} \mathbb{P}_y \left(S_1 < S_A \right) > 0.$$

First, note that under Assumption 1, by the strong Markov property, one has

$$\inf_{y \in \{0 \leq \xi \leq 1\}} \mathbb{P}_y \left(S_1 < S_A \right) = \inf_{y \in \{\xi = 0\}} \mathbb{P}_y \left(S_1 < S_A \right).$$

As mentioned previously, the condition $S_A \wedge S_1 < +\infty$ is a technical simplification of minor significance that may in fact be removed; see Appendix E.

In Appendix F, a stronger but easier-to-check variant of the infimum condition in Assumption 3 is presented.

Let us illustrate the previous conditions in a more specific framework. It turns out that Assumptions 1, 2, and 3 are satisfied for elliptic diffusions in a bounded domain. More precisely, let $Y_s \in \mathbb{R}^d$ be a solution to the SDE

$$\mathrm{d}Y_s = b(Y_s)\,\mathrm{d}s + \sigma(Y_s)\,\mathrm{d}W_s, \tag{2.10}$$

where b and σ are functions from $\mathbb{R}^d$ to, respectively, $\mathbb{R}^d$ and $\mathbb{R}^{d \times n}$, with $n \geq 1$. We denote as usual $a = \sigma\sigma^T$. Then we have the following result, whose proof is detailed in Appendix B.

Lemma 2.3. *Let Y be the solution to (2.10). Assume that*

(a) *σ, b are in $C^2(\mathbb{R}^d)$ with bounded derivatives of order $i = 0, 1, 2$ on $\{-1 \leq \xi \leq 1\}$;*

(b) *ξ is in $C^2(\mathbb{R}^d)$ with bounded derivatives of order $i = 1, 2$ on $\{-1 \leq \xi \leq 1\}$, and also that*

$$A = \bar{A} = \{\xi \leq -1\};$$

(c) *there exists $\delta > 0$ such that $(\nabla \xi)^T a \nabla \xi \geq \delta$ on $\{-1 \leq \xi \leq 1\}$.*

Then Assumptions 1, 2, and 3 hold true.

Remark 2.4. Condition (c) ensures that the martingale part of the process $t \mapsto \xi(Y_t)$ has a strictly positive quadratic variation. It may happen that $(Y_s)_{s \geq 0}$ is Feller and ξ is smooth, but (2.6) does not hold without the addition of condition (c). Consider, for example, the case where $(Y_s)_{s \geq 0}$ is the solution to an ordinary differential equation (ODE). As a consequence, Assumption 2 has to be modified without condition (c); for instance, one may need to resort to an ad hoc restriction of the state space for which (2.6) is still satisfied.

2.4. Main result. For any test function φ and any $t \in [0, 1]$, let us define the unnormalized measure γ_t by

$$\gamma_t(\varphi) := \mathbb{E}\left[\varphi(Y_{S_t}) \mathbf{1}_{S_t < S_A} \right],$$

so that $\gamma_0 = \eta_0$. Accordingly, the probability that the process $(Y_s)_{s \geq 0}$ reaches level t is $p_t := \gamma_t(\mathbf{1}) = \mathbb{P}(S_t < S_A)$, and the law of Y_{S_t} given that $S_t < S_A$ is denoted η_t and satisfies $\eta_t(\varphi) := \gamma_t(\varphi)/\gamma_t(\mathbf{1})$.

The purpose of AMS Algorithm 2.1 is to approximate the previous quantities. Namely, for any $t \in [0,1]$, let us recall that the probability p_t is estimated by

$$p_t^N := \left(1 - \frac{1}{N}\right)^{J_t},$$

where J_t denotes the number of iterations necessary to reach level t. The measures η_t and γ_t are respectively estimated by

$$\eta_t^N(\varphi) := \frac{1}{N}\sum_{n=1}^{N} \varphi(Y_{S_t^n}^{n,J_t}) = \frac{1}{N}\sum_{n=1}^{N} \varphi(X_t^n) \qquad \text{and} \qquad \gamma_t^N(\varphi) := p_t^N \eta_t^N(\varphi).$$

Our first statement is a well-posedness result. As will be explained in section 3.3, it is connected to the first point of Theorem 3.7 and to Lemma 3.8.

Proposition 2.5. *Under Assumptions 1 and 2, AMS Algorithm 2.1 is well-posed in the sense that there is only one particle with minimal score in (2.5). Additionally, under Assumption 3, AMS Algorithm 2.1 is nonexplosive in the sense that the algorithm stops after a finite number of iterations almost surely.*

Our second statement is a consistency result in the L^2 sense and coincides with the second point of Theorem 3.7.

Proposition 2.6. *Under Assumptions 1, 2, and 3, for any $\varphi \in C_b(\{\xi = 1\})$, one has*

$$\mathbb{E}\left[\left(\gamma_1^N(\varphi) - \gamma_1(\varphi)\right)^2\right] \leq \frac{6\,\|\varphi\|_\infty^2}{N}.$$

Let us now come to the central limit result, which corresponds to the last point of Theorem 3.7. The asymptotic variance is described through the integral operator (2.9), namely $q(\varphi)(y) := \mathbb{E}_y\left[\varphi(Y_{S_1})\mathbf{1}_{S_1<S_A}\right]$ defined for any $y \in \{0 \leq \xi \leq 1\}$.

Theorem 2.7. *Under Assumptions 1, 2, and 3, for any $\varphi \in C_b(\{\xi = 1\})$, one has*

$$\sqrt{N}\left(\gamma_1^N(\varphi) - \gamma_1(\varphi)\right) \xrightarrow[N\to\infty]{\mathcal{D}} \mathcal{N}(0, \sigma_1^2(\varphi)),$$

where

$$\sigma_1^2(\varphi) = p_1^2 \mathbb{V}_{\eta_1}(\varphi) - p_1^2 \log(p_1)\,\eta_1(\varphi)^2 - 2\int_0^1 \mathbb{V}_{\eta_t}(q(\varphi))p_t dp_t.$$

Then it is easy to see that Slutsky's lemma and the decomposition

$$\eta_T^N(\varphi) - \eta_T(\varphi) = \frac{1}{\gamma_T^N(\mathbf{1})}\left(\gamma_T^N(\varphi - \eta_T(\varphi)) - \gamma_T(\varphi - \eta_T(\varphi))\right) \tag{2.11}$$

lead to the upcoming result.

Corollary 2.8. *Under Assumptions 1, 2, and 3, for any $\varphi \in C_b(\{\xi = 1\})$, one has*

$$\sqrt{N}\left(\eta_1^N(\varphi) - \eta_1(\varphi)\right) \xrightarrow[N\to\infty]{\mathcal{D}} \mathcal{N}(0, \sigma_1^2(\varphi - \eta_1(\varphi))/p_1^2).$$

Additionally,

$$\sqrt{N}\left(p_1^N - p_1\right) \xrightarrow[N\to\infty]{\mathcal{D}} \mathcal{N}(0,\sigma^2),$$

where

$$\sigma^2 = \sigma_1^2(\mathbf{1}) = -p_1^2 \log(p_1) - 2\int_0^1 \mathbb{V}_{\eta_t}(q(\mathbf{1}))p_t dp_t.$$

First, we can remark that all the terms in the expressions for the asymptotic variances are nonnegative. Indeed, $\log p_1 < 0$, and $t \mapsto p_t$ is nonincreasing, making dp_t nonpositive.

Another important remark is that all of these asymptotic variances can be viewed as the limit of the asymptotic variances for the algorithm with a finite number of levels $0 < t_1 < \cdots < t_K = 1$ as in [5, 10], when the number K of levels tends to infinity. Details and explanations are provided in [7, section 2.4].

In the rest of this section, we discuss some consequences of the previous results. We begin with the number of steps of the algorithm. Essentially, this number grows logarithmically with the rarity of the event and linearly with the number of particles. Remember that one step requires simulation of only one new trajectory, computation of its score, and comparison with the $(N-1)$ scores already evaluated. Hence we can conclude that the total complexity of the algorithm scales like $\mathcal{O}_P(-N \log N \log p_1)$. A similar remark was already given in [14] but was restricted to the so-called idealized setting.

Corollary 2.9. *Under Assumptions 1, 2, and 3, the number of steps of the AMS algorithm satisfies*

$$J_1 = -N \log(p_1) + \mathcal{O}_P(\sqrt{N}).$$

Proof. Indeed, Proposition 2.6 with $\varphi = \mathbf{1}$ gives $p_1^N = \gamma_1^N(\mathbf{1})$ and $p_1 = \gamma_1(\mathbf{1})$ so that

$$p_1^N = p_1 + \mathcal{O}_P(1/\sqrt{N})$$

and

$$\log(p_1^N) = \log(p_1) + \mathcal{O}_P(1/\sqrt{N}).$$

Additionally,

$$\log(p_1^N) = J_1 \log(1 - 1/N) = J_1(-1/N + o(1/N)).$$

Therefore, by using both expressions we get

$$J_1 = -(\log(p_1) + \mathcal{O}_P(1/\sqrt{N}))(1/N + o(1/N))^{-1} = -N \log(p_1) + \mathcal{O}_P(\sqrt{N}).$$ ∎

Now we can focus our attention on the asymptotic variance of the probability estimate, that is,

$$\sigma^2 = \sigma_1(\mathbf{1})^2 = -p_1^2 \log(p_1) - 2\int_0^1 \mathbb{V}_{\eta_t}(q(\mathbf{1}))p_t dp_t. \tag{2.12}$$

By choosing the level function $\xi^\star(y) = q(\mathbf{1})(y) = \mathbb{P}_y(S_1 < S_A)$, it turns out that η_t is supported on the level set t of $\xi^\star$. Additionally, for every y on this level set, we have $q(\mathbf{1})(y) = \mathbb{P}_y(S_1 < S_A) = t$ so that

$$\mathbb{V}_{\eta_t}(q(\mathbf{1})) = \mathbb{V}_{\eta_t}(\xi^\star) = 0.$$

Hence the integral term vanishes, and σ^2 reduces to

$$\sigma^2 = -p_1^2 \log(p_1).$$

Function $y \mapsto \xi^\star(y)$ is called the committor function in molecular dynamics, where its prominent role is well known (see, e.g., [15, 20]). In fact, the knowledge of the committor function typically requires solving a PDE, which in turn is much more involved than the problem of estimating rare event probabilities. However, it is important to note that the committor function gives the best possible asymptotic variance.

This phenomenon also arises when considering the idealized case, where we assume that at each branching, we can generate a new trajectory reaching at least the current level, and independent of the other particles' trajectories (see [3, 14]). Note also that in the one-dimensional case, if ξ is strictly increasing, then the level sets are reduced to one point, η_t is a Dirac measure, and the variance is minimal.

In contrast, we can also exhibit the worst value for σ^2. For that, consider the variance term in the integrand of (2.12), that is, $\mathbb{V}_{\eta_t}(q(\mathbf{1}))$. It corresponds to the variance of the random variable $Z = q(\mathbf{1})(X)$, with X drawn according to η_t. Hence Z is between 0 and 1, and its mean value is p_1/p_t. Under those constraints, the largest variance is that of a Bernoulli variable with parameter p_1/p_t and is given by $p_1/p_t(1 - p_1/p_t)$. In this situation, we have

$$-2\int_0^1 \mathbb{V}_{\eta_t}(q(\mathbf{1}))p_t dp_t = -2\int_0^1 \frac{p_1}{p_t}\left(1 - \frac{p_1}{p_t}\right) p_t \, dp_t = 2p_1^2 \log(p_1) - 2p_1^2 + 2p_1,$$

which yields

$$\sigma^2 \leq 2p_1(1 - p_1) + p_1^2 \log(p_1) \leq 2p_1(1 - p_1).$$

Notice that this upper bound is exactly twice the variance of a naive Monte Carlo method, which simply consists of simulating N i.i.d. (independent and identically distributed) replicates of the original process Y and counting the proportion of trajectories that reach the level set $\{\xi = 1\}$ before A. If we take into account the computational cost, we see that naive Monte Carlo costs N trajectories, whereas the cost of AMS is of order $-N \log p_1$ so that the worst case is worse than naive Monte Carlo by a factor of $-2 \log p_1$.

Consequently, we see that if we make a very bad choice for ξ, things can get pretty bad—even worse than naive Monte Carlo. Nonetheless, this has to be compared to importance sampling, where one is not even guaranteed to have a finite variance (see, for example, [13]). We summarize the previous results in the following corollary.

Corollary 2.10. *Under Assumptions 1, 2, and 3, the asymptotic variance for the probability estimator satisfies*

$$-p_1^2 \log(p_1) \leq \sigma^2 \leq 2p_1(1 - p_1).$$

Let us conclude this section with some comments on the asymptotic variance for $\eta_1^N(\varphi)$, namely $\sigma_1^2(\varphi - \eta_1(\varphi))/p_1^2$. First, notice that for all $t \in [0, 1]$,

$$\gamma_t(q(\varphi)) = \mathbb{E}[q(\varphi(Y_{S_t}))\mathbf{1}_{S_t < S_A}] = \mathbb{E}[\mathbb{E}[\varphi(Y_{S_1})\mathbf{1}_{S_1 < S_A}|Y_{S_t}]\mathbf{1}_{S_t < S_A}],$$

which yields

$$\gamma_t(q(\varphi)) = \mathbb{E}[\varphi(Y_{S_1})\mathbf{1}_{S_1 < S_A}] = \gamma_1(\varphi).$$

As a consequence, by linearity of both q and η_t, one has

$$\eta_t(q(\varphi - \eta_1(\varphi))) = \eta_t(q(\varphi)) - \eta_t(q(\mathbf{1}))\eta_1(\varphi) = (\gamma_t(q(\varphi)) - \gamma_1(\varphi))/p_t = 0.$$

Hence, denoting $r_t = p_t/p_1$, we are led to the alternative expression

$$\sigma_1^2(\varphi - \eta_1(\varphi))/p_1^2 = \mathbb{V}_{\eta_1}(\varphi) - 2\int_0^1 \eta_t((q(\varphi - \eta_1(\varphi)))^2)\, r_t\, dr_t.$$

It is readily seen that $|q(\varphi - \eta_1(\varphi))| \leq \|\varphi - \eta_1(\varphi)\|_\infty q(\mathbf{1})$ so that

$$-2\int_0^1 \eta_t((q(\varphi - \eta_1(\varphi)))^2)\, r_t\, dr_t \leq -2\|\varphi - \eta_1(\varphi)\|_\infty^2 \int_0^1 \eta_t((q(\mathbf{1}))^2)\, r_t\, dr_t.$$

Taking into account that $\eta_t(q(\mathbf{1})) = 1/r_t$, we get

$$-2\int_0^1 \eta_t((q(\mathbf{1}))^2)\, r_t\, dr_t = -2\int_0^1 \mathbb{V}_{\eta_t}(q(\mathbf{1}))\, r_t\, dr_t - 2\log(p_1).$$

So we have the bound

$$\mathbb{V}_{\eta_1}(\varphi) \leq \sigma_1^2(\varphi - \eta_1(\varphi))/p_1^2 \leq \mathbb{V}_{\eta_1}(\varphi) + \|\varphi - \eta_1(\varphi)\|_\infty^2 \left(\frac{\sigma^2}{p_1^2} - \log p_1\right), \tag{2.13}$$

with σ^2 as in Corollary 2.8. The lower bound is the variance we would get with an i.i.d. sample from η_1. As noticed in Corollary 2.10, at best the second term in the right-hand side of (2.13) reduces to $-2\|\varphi - \eta_1(\varphi)\|_\infty^2 \log(p_1)$.

Remark 2.11. In AMS Algorithm 2.1, we assume that the initial condition (2.4) consists of N i.i.d. random variables $Y_0^{n,0}$, $1 \leq n \leq N$, with common law η_0. In fact, this assumption can be relaxed to any exchangeable initial condition satisfying a bound of the form

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

for some constant $c > 0$, as well as satisfying the following CLT:

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

In that case, all the results of this section still hold true (see Remark 2.8 in [6]).

2.5. Extension to path observables and entrance times. This section deals with an extension of the CLT to richer observables. For this purpose, we can consider the following Polish space.

Definition 2.12. *Let $C_{\mathrm{stop}}(\mathbb{R}_+, E)$ denote the space of continuous paths, with possibly a given terminal time s. We will use the notation*

$$y_{[0,s]} := \begin{cases} (y_{s'})_{0\leq s'\leq s} & \text{if } s < +\infty, \\ (y_{s'})_{s'\geq 0} & \text{if } s = +\infty. \end{cases}$$

We say that a sequence $y^n_{[0,s^n]}$ *in* $C_{\mathrm{stop}}(\mathbb{R}_+, E)$ *converges towards* $y_{[0,s]}$ *if* $\lim_n s^n = s \in \overline{\mathbb{R}}_+$ *and* $\lim_n (y^n_{s'\wedge s^n})_{s'\geq 0} = (y_{s'\wedge s})_{s'\geq 0}$ *in* $C(\mathbb{R}_+, E)$ *endowed with uniform convergence on compacts. This defines a Polish topology on* $C_{\mathrm{stop}}(\mathbb{R}_+, E)$ *that we will always use in what follows unless otherwise specified.*

Let us illustrate how we understand the convergence on compacts in the previous definition. Consider, for example, that for all n, $s_n < s < +\infty$. Then $y^n_{[0,s^n]}$ converges to $y_{[0,s]}$ if and only if $s_n \to s$, and both $\sup_{0\leq s'<s_n} d_E(y^n_{s'}, y_{s'})$ and $\sup_{s_n\leq s'<s_n} d_E(y^n_{s_n}, y_{s'})$ converge to 0 when $n \to +\infty$. The reader can easily infer the other cases from that example.

The main message is that, mutatis mutandis, the central limit result of Theorem 2.7 is still valid in this new context. More precisely, we have the extended following CLT (see Appendix G for the proof).

Theorem 2.13. *Let* $\psi : C_{\mathrm{stop}}(\mathbb{R}_+, E) \to \mathbb{R}$ *denote a given continuous and bounded functional. Set* $\mathcal{X}^n_t := Y^{n,J_t}_{[0,S^n_t]}$ *and*

$$\eta^N_t := \frac{1}{N}\sum_{n=1}^{N} \delta_{\mathcal{X}^n_t}.$$

Additionally, consider

$$\gamma_t(\psi) := \mathbb{E}\left[\psi(Y_{[0,S_t]})\mathbf{1}_{S_t<S_A}\right]$$

and

$$q(\psi)(y_{[0,s]}) := \mathbb{E}\left[\psi(Y_{[0,S_1]})\mathbf{1}_{S_1<S_A}|Y_{[0,s]} = y_{[0,s]}\right].$$

Denote, as before, $\eta_t := \gamma_t/\gamma_t(\mathbf{1})$ *and* $\gamma^N_t := p^N_t\eta^N_t$. *If Assumptions 1, 2, and 3 are satisfied, then Proposition 2.6, Theorem 2.7, and Corollary 2.8 hold true when replacing* φ *with* ψ.

Remark 2.14. In the special case of entrance times, i.e., observables of the form $\psi(Y_{S_1}, S_1)$, the CLT is in fact a direct consequence of Theorem 2.7. Indeed, consider the time homogeneous Markov Feller process

$$\widetilde{Y}_h := (Y_h, s_0 + h), \qquad h \geq 0,$$

defined for each initial condition $\widetilde{Y}_0 = (y_0, s_0) \in \{0 \leq \xi \leq 1\} \times [0, 1]$. By construction, it can be easily checked that if Y satisfies Assumptions 1, 2, and 3, then it is also true for $\widetilde{Y}$ so that this case is included in Theorem 2.7.

3. Level-indexed processes and Fleming–Viot particle systems.

In this section, we introduce a càdlàg process X based on the couple (Y, ξ) and called the level-indexed process. In a different framework, it was introduced by Dynkin and Vanderbei in [11] and called a stochastic wave. They mainly studied it in the case where Y is a diffusion—but apparently without any specific application in mind. In our framework, thanks to a slight modification of this object, we can interpret AMS Algorithm 2.1 as a Fleming–Viot particle system. The results of [6] on Fleming–Viot particle systems can then be applied in order to prove Proposition 2.6 and Theorem 2.7.

3.1. Level-indexed processes.

Let us denote by ∂ a cemetery point. Recall that $S_A := \inf\{s \geq 0,\ Y_s \in A\}$ and $S_t := \inf\{s \geq 0,\ \xi(Y_s) > t\}$.

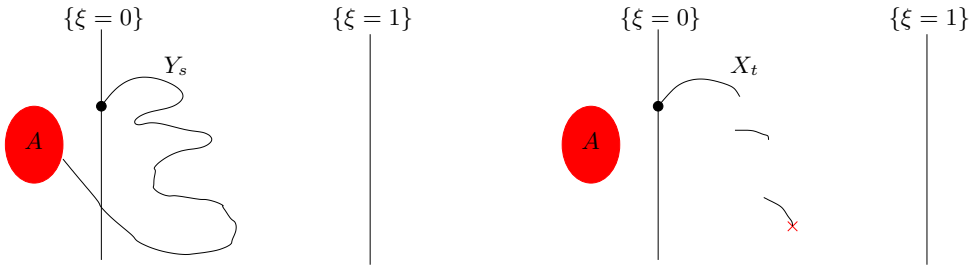

Figure 4. *The level-indexed process X_t associated to the pair (Y, ξ).*

Definition 3.1 (level-indexed process). *Let the first condition of Assumption 2 be satisfied, namely that for each $t \in [0,1]$ and for each $y \in \{\xi = t\}$, we have $\mathbb{P}_y(S_t = 0) = 1$. The level-indexed process, or stochastic wave, $(X_h)_{h\geq 0}$ with state space $F \cup \{\partial\}$, where*

$$F := \{0 \leq \xi \leq 1\},$$

and associated with the pair (Y, ξ) and initial condition $Y_0 = x$, is defined by its initial condition $X_0 = x$, and for any $h \geq 0$,

$$X_h := \begin{cases} Y_{S_{(\xi(x)+h)\wedge 1}} & \text{if } S_{(\xi(x)+h)\wedge 1} < S_A, \\ \partial & \text{if } S_{(\xi(x)+h)\wedge 1} \geq S_A. \end{cases}$$

Remark 3.2.

- The first condition of Assumption 2 is necessary to ensure the consistency of the definition of the level-indexed process. For instance, it is necessary to ensure that $X_0 = Y_{S_{\xi(x)}} = x$ given the initial condition $X_0 = Y_0 = x$.
- If Y has continuous trajectories, ξ is continuous, and $S_1 \wedge S_A < +\infty$ as is the case here, then X_h is càdlàg and $\xi(X_h) = (\xi(x) + h) \wedge 1$ for all $h \geq 0$. See also Appendix E for the case where $S_1 \wedge S_A = +\infty$ with nonzero probability.
- If the initial level is $\xi(x) = 0$ and if $t \in [0,1]$ is such that $X_t \neq \partial$, then $\xi(X_t) = t$ (see Figure 4). In particular, if $X_1 \neq \partial$, then $X_t = Y_{S_t}$ for all $t \in [0,1]$.
- If Y is Feller, as is the case here, then Y is strong Markov with respect to its right-continuous natural filtration $\mathcal{F}^Y$. By construction of the level-indexed process, it implies that X is—at least—a time homogeneous Markov process with respect to the filtration $(\mathcal{F}^Y_{S_{(\xi(X_0)+h)\wedge 1}})_{h\geq 0}$, and thus a fortiori with respect to its smaller natural filtration.
- If $F = \{0 \leq \xi \leq 1\}$ is compact, then continuity Lemma A.4 implies that the level-indexed process X is itself Feller.

In the case where Y is not stopped at S_A, the level-indexed process has been introduced in [11] and called a stochastic wave. If, for example, $Y_s = (Y_s^1, Y_s^2)$ is a two-dimensional Brownian motion with $Y_0 = 0$ and $\xi(y^1, y^2) = y^1$, then $X_t^1 = t$ and $X_t^2 = Y_{S_t}^2$, where $S_t = \inf\{s \geq 0, Y_s^1 = t\}$ is a symmetric Cauchy process with a dense set of discontinuity points. As pointed out in [11], this representation of the symmetric Cauchy process is due to [24]. This is illustrated in Figure 5.

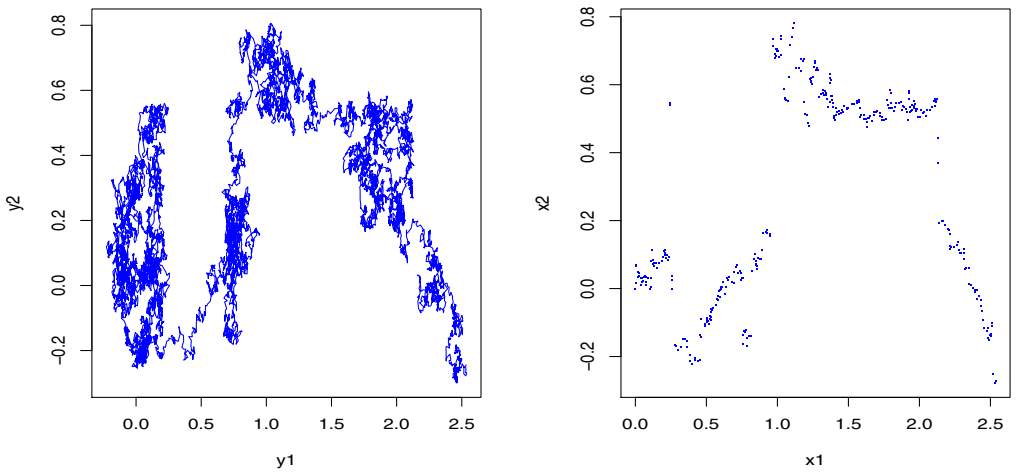

Figure 5. *2D Brownian trajectory $Y_s = (Y_s^1, Y_s^2)$ and associated stochastic wave $(X_t^1, X_t^2) = (t, Y_{S_t}^2)$ when $\xi(y^1, y^2) = y^1$.*

Remark 3.3 (soft versus hard killing times). It turns out that under Assumption 2, the killing time of the level-indexed process is typically "soft" in the sense that it is a totally inaccessible stopping time, i.e., a stopping time that cannot be predicted (see [16] for a precise definition, as well as the discussion in [6]). This is, for instance, a consequence of the Feller property when X is Feller. Note that this is a stronger property than having an atomless distribution. Interestingly, the CLT in [6] also holds true for "hard" killing times so that it may be used to treat cases beyond Assumption 2.

The Markov semigroup of the level-indexed process, defined by

$$Q^h\varphi(x) := \mathbb{E}[\varphi(X_h)|X_0 = x],$$

can be easily related to the integral operator $q(\varphi)(y) = \mathbb{E}_y\left[\varphi(Y_{S_1})\mathbf{1}_{S_1<S_A}\right]$ as follows.

Lemma 3.4. *For any $x \in F$ and any $\varphi : F \to \mathbb{R}$ extended to $F \cup \{\partial\}$ with the convention $\varphi(\partial) = 0$, one has*

$$Q^{1-\xi(x)}\varphi(x) = q(\varphi)(x).$$

3.2. AMS as a Fleming–Viot particle system. AMS Algorithm 2.1 can be recast in the form of a Fleming–Viot algorithm as studied in [6]. For this purpose, let us consider a time homogeneous càdlàg Markov process $(X_h)_{h\geq 0}$ in $F \cup \{\partial\}$, constructible from any initial condition in F. We assume that ∂ is an absorbing state, meaning that $X_{h'} \in \partial$ whenever $X_h \in \partial$ and $h' \geq h$. Let us first recall what we mean by a Fleming–Viot particle system.

Definition 3.5 (Fleming–Viot particle system). *An exchangeable particle system denoted by $(X_t^1, \ldots, X_t^N)_{t\geq 0}$ in F^N is called the Fleming–Viot particle system associated with $(X_h)_{h\geq 0}$ if the following hold:*

- *Initialization: the particles are initially i.i.d. with distribution η_0,*

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

- *Evolution and killing: between branching times, each particle evolves independently according to the law of the underlying Markov process X until one of them hits ∂.*
- *Branching (or rebirth): the killed particle is taken from ∂ and is instantaneously given the state of one of the $(N-1)$ other particles—the choice being uniformly random.*
- *The system continues until time 1.*

Note that in order to be well defined, a Fleming–Viot particle system should almost surely satisfy the following two properties: (i) particles die at different times, and (ii) there is a finite number of branchings in the time interval $[0,1]$. Some conditions ensuring (i) and (ii) are given and discussed below.

The next result makes explicit the connection among the AMS algorithm, Fleming–Viot particle systems, and the level-index process.

Lemma 3.6. *Let Assumptions 1 and 2 hold true for the pair (Y,ξ). Recall that the particles have initial level 0, i.e., $\eta_0(\xi=0)=1$. Consider AMS Algorithm 2.1. For each $n=1,\dots,N$ and each $t\in[0,1]$, set as before*

$$S_t^n := \inf\left\{s\geq 0,\, \xi(Y_s^{n,J_t})>t\right\} = \inf\left\{s\geq 0,\, \xi(Y_s^{n,J_t})=t\right\},$$

and set as well

$$X_t^n := Y_{S_t^n}^{n,J_t}.$$

Then $(X_t^1,\dots,X_t^N)_{t\geq 0}$ is the Fleming–Viot particle system in $F=\{0\leq\xi\leq 1\}$ associated with the level-indexed process X of the pair (Y,ξ) in the sense of Definitions 3.1 and 3.5.

Proof. For $t\in[0,1]$, $j\geq 0$, and $n=1,\dots,N$, let us first define

$$S_t^{n,j} := \inf\{s\geq 0,\ \xi(Y_s^{n,j})\in A\cup\{\xi=t\}\}$$

and

$$X_t^{n,j} := \begin{cases} Y_{S_t^{n,j}}^{n,j} & \text{if } \xi(Y_{S_t^{n,j}}^{n,j})=t,\\ \partial & \text{otherwise.}\end{cases}$$

By Assumption 2 and Remark 3.2, the initial condition satisfies

$$(X_0^{1,0},\dots,X_0^{N,0}) = (Y_0^{1,0},\dots,Y_0^{N,0})\in\{\xi=0\}^N,$$

so that

$$\xi(X_0^{1,0})=\cdots=\xi(X_0^{N,0})=0.$$

Note also that for all $t\in[0,1]$, if $X_t^{n,j}\neq\partial$, then

$$\xi(X_t^{n,j}) = \xi(Y_{S_t^{n,j}}^{n,j}) = t,$$

so that $X^{n,j}$ is indeed the level-indexed process associated with $Y^{n,j}$ in the sense of Definition 3.1.

Set $\tau_0=0$. By construction of AMS Algorithm 2.1, the processes $(X_t^{1,j},\dots,X_t^{N,j})_{0\leq t\leq 1}$ can thus be iteratively constructed for $j\geq 1$ as follows:

1. We can reformulate N_j and τ_j defined in AMS Algorithm 2.1 as

$$\text{(3.1)} \qquad \begin{cases} N_j := \arg\min_{n=1,\dots,N} \sup_{0\le s\le S_A^{n,j-1}\wedge S_1^{n,j-1}} \xi(Y_s^{n,j-1}), \\ \tau_j := \sup_{0\le s\le S_A^{N_j,j-1}\wedge S_1^{N_j,j-1}} \xi(Y_s^{N_j,j-1}). \end{cases}$$

2. Stop if $\tau_j = 1$; i.e., all trajectories are still alive at time 1.
3. Set $X^{n,j} := X^{n,j-1}$ for $n \neq N_j$. Pick a number M_j uniformly at random in $\{1,\dots,N\}\setminus\{N_j\}$.
4. Replace the trajectory on $[0,\tau_j]$ of the particle with index N_j by the trajectory of the particle with index M_j; that is, set $(X_t^{N_j,j})_{0\le t\le\tau_j} := (X_t^{M_j,j})_{0\le t\le\tau_j}$. Let particle N_j evolve independently starting from state $X_{\tau_j}^{N_j,j}$ at time τ_j until time 1 or until it is killed.

If we now set

$$X_t^n := X_t^{n,J_t} = X_t^{n,j-1} \quad \text{for} \quad \tau_{j-1} \le t < \tau_j$$

for $n = 1,\dots,N$ and $j \ge 1$, we thus obtain by definition the Fleming–Viot particle system associated with the level-indexed Markov process X of Definition 3.5. ∎

3.3. L^2-estimate and CLT for Fleming–Viot particle systems. Building on [6], we can now present two sufficient assumptions to obtain the desired L^2-estimate and CLT for Fleming–Viot particle systems based on the level-indexed processes.

The first assumption is the following.

Assumption (Ã). *This assumption has two parts.*

(i) *For any initial condition $x \in F = \{0 \le \xi \le 1\}$, the jump "times" of the level-indexed process $(X_h)_{h\ge0}$ have an atomless distribution:*

$$\mathbb{P}(X_{h^-} \neq X_h \mid X_0 = x) = 0 \quad \forall x \in F,\ \forall h \ge 0.$$

(ii) *If $\varphi \in C_b(\{\xi = 1\})$, then the mapping $x \mapsto q(\varphi)(x) = Q^{1-\xi(x)}\varphi(x)$ is continuous on F.*

The second key assumption is simply the following.

Assumption (B). *The Fleming–Viot particle system is well defined in the sense that $\mathbb{P}(J_1 < +\infty) = 1$, where J_1 denotes the number of branchings until final time 1.*

Under these assumptions, [6] implies the following (see Appendix C for details on how to rigorously import the content of [6]).

Theorem 3.7. *Under Assumptions (Ã) and (B), one has the following:*

- *The Fleming–Viot particle system is well-posed in the sense that only one particle is killed at each branching time.*
- *L^2-estimate: for any $\varphi \in C_b(\{\xi = 1\})$,*

$$\mathbb{E}\left[\left(\gamma_1^N(\varphi) - \gamma_1(\varphi)\right)^2\right] \le \frac{6\,\|\varphi\|_\infty^2}{N}.$$

- *Central limit theorem: for any $\varphi \in C_b(\{\xi = 1\})$,*

$$\sqrt{N}\left(\gamma_1^N(\varphi) - \gamma_1(\varphi)\right) \xrightarrow[N\to\infty]{\mathcal{D}} \mathcal{N}(0,\sigma_1^2(\varphi)),$$

where

$$\sigma_1^2(\varphi) = p_T^2 \mathbb{V}_{\eta_1}(\varphi) - p_1^2 \log(p_1)\, \eta_1(\varphi)^2 - 2 \int_0^1 \mathbb{V}_{\eta_t}(Q^{1-t}(\varphi)) p_t dp_t.$$

Proposition 2.6, Theorem 2.7, and Corollary 2.8 are then consequences of the following lemma.

Lemma 3.8. *Assumptions 1 and 2 imply Assumption (Ã). With the addition of Assumption 3, they also imply Assumption (B).*

The proof of Lemma 3.8 is given in Appendix D. For now, let us just give some intuition behind this result. First, assume for simplicity that $\xi(X_0) = 0$ and that a jump of the level-indexed process occurs, i.e., $X_{t^-} \neq X_t$ for a given $t \in [0, 1]$. Then, since by left continuity $X_{t^-} = Y_{S_{t^-}}$, this jump means that $s \mapsto \xi(Y_s)$ has a local maximum with value t. However, under Assumption 2, this is prohibited since the Y-hitting times of levels $> t$ and $\geq t$ are equal almost surely.

Second, the continuity of $x \mapsto q(\varphi)(x) = \mathbb{E}_x\left[\varphi(Y_{S_1})\mathbf{1}_{S_1 < S_A}\right]$ is a consequence of the fact that Y is Feller (Assumption 1) and that for Feller processes, hitting times of the interior or the closure of, respectively, A and $\{\xi \geq 1\}$ are the same.

Finally, the fact that the algorithm has almost surely a finite number of branchings (non-explosion) comes directly from the uniform lower bound of Assumption 3 through a comparison with a geometric random sequence.

Appendix A. Preliminaries on Feller processes. In this appendix, we recall some standard properties of continuous Feller processes. Most proofs are detailed in [6] in the case of càdlàg processes.

Definition A.1. *Let E be a locally compact Polish space. Let $C_0(E)$ denote the space of continuous functions that vanish at infinity. A continuous process $(Y_s)_{s\geq 0}$ in E is Feller if each of its probability transition maps $C_0(E)$ into itself: for all $\varphi \in C_0(E)$ and $s \geq 0$, $(y \mapsto \mathbb{E}_y[\varphi(Y_s)]) \in C_0(E)$.*

Feller processes have the following useful standard properties: (i) The associated natural filtration $\mathcal{F}_s^Y := \sigma\left(Y_{s'},\, 0 \leq s' \leq s\right)$ is right-continuous, and (ii) Y is strong Markov with respect to $\mathcal{F}^Y$.

We will need the following slightly less standard pathwise continuity of Feller processes.

Lemma A.2. *Let $C(\mathbb{R}_+, E)$ denote the space of continuous trajectories endowed with uniform convergence on compacts. Let $(Y_s^y)_{s\geq 0} \in C(\mathbb{R}_+, E)$ denote a given Feller process with initial condition $Y_0 = y$. Then the mapping $y \mapsto \mathcal{L}((Y_s^y)_{s\geq 0})$ from E to probabilities on $C(\mathbb{R}_+, E)$, endowed with convergence in distribution, is continuous.*

Proof. In [6, Lem. 4.3], the convergence is shown in the Skorokhod space instead of $C(\mathbb{R}_+, E)$ using Theorem 17.25 of [18]. The Skorokhod topology and the topology of uniform convergence on compacts on $C(\mathbb{R}_+, E)$ are known to be the same on continuous trajectories; see Lemma 10.1 in Chapter 3 of [12] (see also Problem 7 in Chapter VI of [21]).

Hence we have the result. ∎

We then recall some lower and upper semicontinuity of hitting times with respect to the locally uniform topology.

Lemma A.3. *Let $B \subset E$ be a Borel set. For each $y \in C(\mathbb{R}_+, E)$, define $s_{\mathring{B}}(y) := \inf\{s \geq 0,\ y_s \in \mathring{B}\}$ as well as $s_{\bar{B}}(y) := \inf\{s \geq 0,\ y_s \in \bar{B}\}$. Then $s_{\mathring{B}}$ is upper semicontinuous in $C(\mathbb{R}_+, E)$ and $s_{\bar{B}}$ is lower semicontinuous in $C(\mathbb{R}_+, E)$: for any sequence $(y^n)_{n\geq 1}$ converging to $y \in C(\mathbb{R}_+, E)$,*

$$\limsup_n s_{\mathring{B}}(y^n) \leq s_{\mathring{B}}(y),$$
$$s_{\bar{B}}(y) \leq \liminf_n s_{\bar{B}}(y^n).$$

Proof. See Lemma 4.4 in [6]. ∎

We can then conclude with the general property used to prove the continuity of the integral operator q defined in (2.9) (see Lemma D.3). We denote as before $S_{\bar{B}} := \inf\{s \geq 0,\ Y_s \in \bar{B}\}$ as well as $S_{\mathring{B}} := \inf\{s \geq 0,\ Y_s \in \mathring{B}\}$.

Lemma A.4. *Let $B \subset E$ be a Borel set, let Y be a continuous Feller process, and let $\lim_n y^n = y$ be a converging sequence of initial conditions. If*

$$\mathbb{P}_y\left(S_{\bar{B}} = S_{\mathring{B}}\right) = 1, \tag{A.1}$$

then the distribution of S_B under $\mathbb{P}_{y^n}$ converges when $n \to +\infty$ towards its distribution under $\mathbb{P}_y$. If, moreover, $\mathbb{P}_y(S_B < +\infty) > 0$, then the joint distribution of (S_B, Y_{S_B}) in $\mathbb{R}_+ \times E$ under $\mathbb{P}_{y^n}(\ \ |S_B < +\infty)$ converges when $n \to +\infty$ towards the joint distribution under $\mathbb{P}_y(\ \ |S_B < +\infty)$.

Proof. Using Lemma A.2 and a Skorokhod embedding argument, a sequence $(Y^n_s)_{s\geq 0}$ of Feller processes with initial conditions $(y^n)_{n\geq 0}$ can be constructed on a single probability space so that $\lim_n Y^n = Y$ in $C(\mathbb{R}_+, E)$ almost surely, where Y denotes the Feller process with initial condition y. Then Lemma A.3 with (A.1) implies that $\lim_n S^n_B = S_B$, and hence the first result. The second result follows by continuity of Y. ∎

In order to obtain a pathwise version of the main CLT of the present paper, we will need a pathwise version of the latter continuity result. For this purpose, let us recall that the Polish space $C_{\text{stop}}(\mathbb{R}_+, E)$ of continuous paths with a possibly given end time (see Definition 2.12) is equipped with a topology defined by the convergence of end times and of processes stopped at the end time uniformly on any finite time intervals.

The following technical lemma about the continuity of the extension of paths will prove useful.

Lemma A.5. *The extension map*

$$\begin{array}{rcccl} T & : & C_{\text{stop}}(\mathbb{R}_+, E) \times C(E, \mathbb{R}_+) & \longrightarrow & C(E, \mathbb{R}_+), \\ & & y = ((y_{s'})_{0\leq s'\leq s}, (\tilde{y}_h)_{h\geq 0}) & \longmapsto & Ty = \left\{ \begin{array}{ll} y_{s'}, & s' \leq s, \\ \tilde{y}_{s'-s}, & s' \geq s \end{array} \right. \end{array}$$

defined for paths satisfying $y_s = \tilde{y}_0$ is continuous.

Proof. Let d stand for the distance on E. Denote by $(y^n_{[0,s^n]})$ a sequence of paths in $C_{\text{stop}}(\mathbb{R}_+, E)$ converging to $y^\infty_{[0,s^\infty]}$ (for the topology given in Definition 2.12), and let $(\tilde{y}^n)$

be a sequence of paths in $C(E,\mathbb{R}_+)$ converging to $\tilde{y}^\infty$ (uniformly on compact sets). We assume that $y^n_{s^n} = \tilde{y}^n_0$ for all n, and assume as well that $y^\infty_{s^\infty} = \tilde{y}^\infty_0$. We have to prove the convergence of the extended function sequence (Ty^n) to Ty^∞. Note that for any s, one can control $d(Ty^\infty_s, Ty^n_s)$ by considering all cases as follows:

- If $s \le s^\infty \wedge s^n$,
$$d(Ty^\infty_s, Ty^n_s) \le d(y^\infty_{s\wedge s^\infty}, y^n_{s\wedge s^n}).$$
- If $s \ge s^\infty \vee s^n$,
$$d(Ty^\infty_s, Ty^n_s) \le d(\tilde{y}^\infty_{s-s^\infty}, \tilde{y}^n_{s-s^n}) \le d(\tilde{y}^\infty_{s-s^\infty}, \tilde{y}^\infty_{s-s^n}) + d(\tilde{y}^n_{s-s^n}, \tilde{y}^\infty_{s-s^n}).$$
- If $s^n \le s \le s^\infty$,
$$\begin{aligned} d(Ty^\infty_s, Ty^n_s) &= d(y^\infty_s, \tilde{y}^n_{s-s^n}) \\ &\le d(y^\infty_s, \tilde{y}^\infty_{s-s^n}) + d(\tilde{y}^n_{s-s^n}, \tilde{y}^\infty_{s-s^n}) \\ &\le d(y^\infty_s, y^\infty_{s^\infty}) + d(\tilde{y}^\infty_0, \tilde{y}^\infty_{s-s^n}) + d(\tilde{y}^n_{s-s^n}, \tilde{y}^\infty_{s-s^n}). \end{aligned}$$
- If $s^\infty \le s \le s^n$,
$$d(Ty^\infty_s, Ty^n_s) = d(\tilde{y}^\infty_{s-s^\infty}, y^n_s) \le d(y^\infty_{s\wedge s^\infty}, y^n_{s\wedge s^n}) + d(\tilde{y}^\infty_{s-s^\infty}, \tilde{y}^\infty_0).$$

Let $s_0 \ge 0$ be given. The convergence assumptions, together with the uniform continuity of Ty^∞ on compacts, imply that, when n goes to infinity, all the right-hand sides converge uniformly to 0 with respect to $s \in [0, s_0]$. Hence we have the result. ∎

We can then safely prove the following pathwise continuity of stopped Feller processes.

Lemma A.6. *Let $\lim_n y^n_{[0,s^n]} = y_{[0,s]}$ denote a converging sequence of initial path conditions in $C_{\mathrm{stop}}(\mathbb{R}_+, E)$. Let $B \subset E$ denote a Borel set, and let Y denote a continuous Feller process. Assume that*
$$\mathbb{P}_{y_s}\left(S_{\bar{B}} = S_{\mathring{B}}\right) = 1.$$
Then the distribution of $(Y_{0\le s\le S_B})$ in $C_{\mathrm{stop}}(\mathbb{R}_+, E)$ under $\mathbb{P}(\ \cdot\ | Y_{[0,s^n]} = y^n_{[0,s^n]})$ converges when $n \to +\infty$ towards its distribution under $\mathbb{P}(\ \cdot\ | Y_{[0,s]} = y_{[0,s]})$.

Remark A.7. Recall that, by Definition 2.12, if $S_B = +\infty$, then $(Y_{0\le s\le S_B})$ is actually $(Y_{0\le s<+\infty})$.

Proof. It is an extension of the proof of Lemma A.4 using Lemma A.5. Indeed, the latter and a Skorokhod embedding argument allow us to construct a sequence Y^n converging almost surely to Y in $C(E,\mathbb{R}_+)$ such that (i) $Y^n_{s'} = y^n_{s'}$ for $s' \le s^n$ and $Y_{s'} = y_{s'}$ for $s' \le s$, and (ii) all processes are distributed according to Y with initial conditions prescribed by (i). Then Lemma A.3 with (A.1) implies that $\lim_n S^n_B = S_B$ almost surely, and hence the result. ∎

Appendix B. Proof of Assumptions 1, 2, and 3 for diffusions in $\mathbb{R}^d$. We can now establish Lemma 2.3 by checking successively that Assumptions 1, 2, and 3 hold true under the conditions (a), (b), and (c) stated in Lemma 2.3.

Step 1. Assumption 1 holds true.

Indeed, condition (a) implies that the diffusion is a strong solution of the SDE (2.10) and is Feller; see, for example, [12, Thm. 2.4, p. 373].

Step 2. Assumption 2 holds true.

By definition of the stopping times S_t, condition (2.6) of Assumption 2 will follow from

$$\mathbb{P}_y\Big(\overline{\lim_{s\downarrow 0}}\, s^{-\frac{1}{2}}(\xi(Y_s)-\xi(y)) = +\infty\Big) = 1 \tag{B.1}$$

for any $y \in \{0 \le \xi \le 1\}$. On the other hand, recalling that $A := \{\xi \le -1\}$, condition (2.7) of Assumption 2 follows similarly from the strong Markov property for Feller processes and from the fact that, for any $y \in \{\xi = -1\}$,

$$\mathbb{P}_y\Big(\underline{\lim_{s\downarrow 0}}\, s^{-\frac{1}{2}}(\xi(Y_s)-\xi(y)) = -\infty\Big) = 1. \tag{B.2}$$

We claim that both (B.1) and (B.2) hold true for any $y \in \{-1 \le \xi \le 1\}$, which will conclude the proof of Step 2.

Since ξ is C^2, Itô's formula gives

$$\mathrm{d}\xi(Y_s) = \Big(\nabla\xi(Y_s)^T b(Y_s) + \tfrac{1}{2}Tr(\nabla^2\xi(Y_s)a(Y_s))\Big)\,\mathrm{d}s + \nabla\xi(Y_s)^T\sigma(Y_s)\,\mathrm{d}W_s$$

so that if we denote by $U_s\,\mathrm{d}s$ the finite variation part of this decomposition and set

$$\Sigma_s = \sqrt{\big((\nabla\xi)^T a\nabla\xi\big)(Y_s)} \ge \sqrt{\delta} > 0,$$

then for the one-dimensional Brownian motion $\mathrm{d}\widetilde{W}_s = \Sigma_s^{-1}\nabla\xi(Y_s)^T\sigma(Y_s)\,\mathrm{d}W_s$, we obtain

$$\mathrm{d}\xi(Y_s) = U_s\,\mathrm{d}s + \Sigma_s\,\mathrm{d}\widetilde{W}_s. \tag{B.3}$$

Since $s \mapsto U_s$ is continuous, it remains to prove that

$$\mathbb{P}_y\Big(\underline{\lim_{s\downarrow 0}}\, s^{-\frac{1}{2}}\int_0^s \Sigma_r\,\mathrm{d}\widetilde{W}_r = -\infty\Big) = \mathbb{P}_y\Big(\overline{\lim_{s\downarrow 0}}\, s^{-\frac{1}{2}}\int_0^s \Sigma_r\,\mathrm{d}\widetilde{W}_r = +\infty\Big) = 1. \tag{B.4}$$

For this, note that the process

$$s \mapsto B_{\int_0^s \Sigma_r^2\,\mathrm{d}r} := \int_0^s \Sigma_r\,\mathrm{d}\widetilde{W}_r \tag{B.5}$$

is a time-changed Brownian motion B (see Chapter V, section 1 of [22]). The law of the iterated logarithm for the Brownian motion (see Chapter II, section 1 of [22]) now implies (B.4) since almost surely we have

$$\begin{aligned}
\overline{\lim_{s\to 0}}\, s^{-\frac{1}{2}}\int_0^s \Sigma_r\,\mathrm{d}\widetilde{W}_r &= \overline{\lim_{s\to 0}}\Big(\frac{1}{s}\int_0^s \Sigma_r^2\,\mathrm{d}r\Big)^{\frac{1}{2}}\Big(\int_0^s \Sigma_r^2\,\mathrm{d}r\Big)^{-\frac{1}{2}} B_{\int_0^s \Sigma_r^2\,\mathrm{d}r}\\
&= \Sigma_0\,\overline{\lim_{s\to 0}}\Big(\int_0^s \Sigma_r^2\,\mathrm{d}r\Big)^{-\frac{1}{2}} B_{\int_0^s \Sigma_r^2\,\mathrm{d}r}\\
&= +\infty.
\end{aligned}$$

The same reasoning applies for the other limit in (B.4).

Step 3. Assumption 3 holds true.

Consider the differential equation (B.3) above and recall that condition (a) on the coefficients implies that U_s is bounded, while the positive lower bound in condition (c) implies that $\Sigma_s > 0$ is bounded from above and from below.

We then consider the continuous process $s \mapsto Z_s$ defined by $Z_0 = \xi(Y_0)$ and

$$dZ_s := -\lambda_0 \Sigma_s^2 \, ds + \Sigma_s \, d\widetilde{W}_s,$$

where λ_0 is such that

$$U_s \Sigma_s^{-2} \leq \lambda_0$$

almost surely and for all $s \geq 0$. By construction, (i) the process $s \mapsto Z_s - \xi(Y_s)$ is decreasing and thus negative; and (ii) $s \mapsto Z_s$ is the time-changed Brownian motion (B.5) but drifted with constant drift $-\lambda_0$, that is,

$$\widetilde{Z}_{l=\int_0^s \Sigma_r^2} := Z_s$$

satisfies $d\widetilde{Z}_l = -\lambda_0 dl + dB_l$. We will denote by $S_{\pm 1}^Z$ the first hitting time of ± 1 by Z and denote by $L_{\pm 1}^{\tilde{Z}}$ the first hitting time of ± 1 of $\tilde{Z}$ so that

$$\int_0^{S_{\pm 1}^Z} \Sigma_r^2 \, dr = L_{\pm 1}^{\tilde{Z}}.$$

Consider also the stopping time σ defined by

$$\int_0^{\sigma} \Sigma_r^2 \, dr = 1.$$

Notice that $\sigma \leq 1/\delta$ almost surely. Now, let us first prove that

$$p_0 := \sup_{y \in \{-1 \leq \xi \leq 1\}} \mathbb{P}_y \left(S_1 = S_{-1} = +\infty \right) = 0.$$

Conditioning and applying the strong Markov property yields

$$\begin{aligned} \mathbb{P}_y \left(S_1 = S_{-1} = +\infty \right) &= \mathbb{E}_y \left[\mathbb{E}_y [\mathbf{1}_{S_1 = S_{-1} = +\infty} | \mathcal{F}_\sigma^Y] \mathbf{1}_{S_1 \wedge S_{-1} > \sigma} \right] \\ &= \mathbb{E}_y \left[\mathbb{E} [\mathbf{1}_{S_1 = S_{-1} = +\infty} | Y_\sigma] \mathbf{1}_{S_1 \wedge S_{-1} > \sigma} \right] \\ &\leq p_0 \mathbb{P}_y \left(S_1 \wedge S_{-1} > \sigma \right) \leq p_0 \mathbb{P}_y \left(S_1 > \sigma \right). \end{aligned}$$

Since $\xi(Y_s) \geq Z_s$, we have $S_1 \leq S_1^Z$ so that

$$\mathbb{P}_y(S_1 > \sigma) \leq \mathbb{P}_y(S_1^Z > \sigma) = \mathbb{P}\left(\int_0^{S_1^Z} \Sigma_r^2 dr > \int_0^{\sigma} \Sigma_r^2 dr \right) = \mathbb{P}_y(L_1^{\tilde{Z}} > 1).$$

Since $\tilde{Z}$ starting from 0 is stochastically smaller than $\tilde{Z}$ starting from $\xi(y) \geq 0$, it yields $\mathbb{P}_y(L_1^{\tilde{Z}} > 1) \leq \mathbb{P}(L_1^{\tilde{Z}} > 1 | \tilde{Z} = 0) < 1$ so that

$$p_0 \leq p_0 \times \mathbb{P}(L_1^{\tilde{Z}} > 1 | \tilde{Z} = 0),$$

which shows that $p_0 = 0$.

Finally, let us prove that

$$p_1 := \inf_{y \in \{\xi = 0\}} \mathbb{P}_y \left(S_1 < S_{-1} \right) > 0.$$

Obviously, since $Z_s \leq \xi(Y_s)$ and $\xi(y) = 0$,

$$\mathbb{P}_y(S_1 < S_{-1}) \geq \mathbb{P}_y(S_1^Z < S_{-1}^Z) = \mathbb{P}(L_1^{\tilde{Z}} < L_{-1}^{\tilde{Z}} | \tilde{Z} = 0) > 0,$$

with the last term being independent of the choice of $y \in \{\xi = 0\}$.

Appendix C. Remarks on the main result of [6]. Let us now explain the connection between Assumptions (Ã) and (B) and the set of assumptions in [6]. Theorem 3.7 of the present paper corresponds exactly to Propositions 3.3 and 3.13 in [6], where they are established under Assumption (B), also called Assumption (B) in [6], as well as a weaker variant of Assumption (Ã), called Assumption (A′) and recalled below.

Definition C.1. *First, let us fix a measurable bounded function $\varphi : F \to \mathbb{R}$, and denote for each $1 \leq n \leq N$ and any $t \in [0, 1]$,*

$$\mathbb{L}_t^n := Q^{1-t}(\varphi)(X_t^n) = q(\varphi)(X_t^n),$$

where φ is omitted in order to lighten the notation. For any $n \in \{1, \ldots, N\}$ and any $k \geq 0$, we denote by $\tau_{n,k}$ the kth branching time of particle n, with the convention $\tau_{n,0} = 0$. Moreover, for any $j \geq 0$, we denote by τ_j the jth branching time of the whole system of particles, with the convention $\tau_0 = 0$.

Remark C.2. The identity in the definition of $\mathbb{L}_t^n$ comes from Lemma 3.4.

A key assumption imposed on the Fleming–Viot particle system in [6] is the following.

Assumption (A′). *We assume that the particle system is such that for the bounded test function φ, $t \mapsto \mathbb{L}_t^n$ is càdlàg for each $1 \leq n \leq N$, and that*

(i) *only one particle is killed at each branching time: if $n \neq m$, then $\tau_{n,k} \neq \tau_{m,j}$ almost surely for any $j, k \geq 1$. In other words, the particle system is well defined;*

(ii) *the processes $\mathbb{L}_t^n$ and $\mathbb{L}_t^m$ do not jump at the same time: if $n \neq m$, then*

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

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

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

In order to obtain precisely Theorem 3.7, it remains to show that Assumption (Ã) implies Assumption (A′), that is,

$$(\tilde{A}) \Rightarrow (A').$$

In fact, this can be proven using exactly the same arguments as those used to prove Lemma 3.1 in [6]. In the latter, it is shown that a slightly stronger but very similar assumption

(denoted Assumption (A) and not specific to the AMS context) implies Assumption (A′), that is,

$$(A) \Rightarrow (A').$$

However, the proof of $(\tilde{A}) \Rightarrow (A')$ is very similar to that of $(A) \Rightarrow (A')$, so we will not go into more detail on this point. In summary, following the arguments of the proof of Lemma 3.1 in [6], we can check the chain of implications

$$(A) \Rightarrow (\tilde{A}) \Rightarrow (A').$$

Appendix D. Assumptions 1, 2, 3 imply (Ã) and (B).

Lemma D.1. *Under Assumptions 1 and 2, for any $y \in F = \{0 \leq \xi \leq 1\}$ and any $t \in]0,1]$ satisfying $\xi(y) \leq t$, one has $\mathbb{P}_y(S_t = S_{t^-}) = 1$, meaning that the jump times of $t \mapsto S_t$ have an atomless distribution.*

Proof. Let us recall that

$$S_t := \inf\{s \geq 0,\ \xi(Y_s) > t\} = \inf\{s \geq 0,\ \xi(Y_s) = t\}$$

is a stopping time with respect to the natural filtration of Y for all $t \in [0,1]$, and that by continuity of (Y, ξ), the process $t \mapsto S_t$ is càdlàg. By construction, for $t > 0$, S_{t^-} is the supremum of the increasing sequence of stopping times $(S_{t-1/k})_{k \geq 1}$, and thus is itself a stopping time.

Y is Feller by Assumption 1, so in particular it is strong Markov: for the stopping time S_{t^-}, this gives

$$\mathbb{P}_y(S_t = S_{t^-}) = \mathbb{P}_y\left(S_t = S_{t^-} = +\infty\right) + \mathbb{E}_y\left[\mathbf{1}_{S_{t^-} < +\infty} \mathbb{P}_{Y_{S_{t^-}}}\left(S_t = S_{t^-}\right)\right].$$

But (2.6) in Assumption 2 directly implies that $\mathbb{P}_{Y_{S_{t^-}}}\left(S_t = S_{t^-} = 0\right) = 1$ almost surely so that $\mathbb{P}_y(S_t = S_{t^-}) = 1$. ∎

Lemma D.2. *Assumptions 1 and 2 imply Assumption (Ã)(i). In other words, for any $h \geq 0$ and any initial condition $X_0 = x \in F$, one has $X_h = X_{h^-}$ almost surely.*

Proof. Consider the level $t_h := 1 \wedge (\xi(x) + h)$. Then $X_h = Y_{S_{t_h}}$, where $Y_0 = x$, with the convention $Y_{+\infty} = \partial$. Since $h \mapsto t_h$ and $s \mapsto Y_s$ are almost surely continuous, the result is then a consequence of the fact that $S_{t_h} = S_{t_h^-}$, which is precisely the result of Lemma D.1. ∎

Lemma D.3. *Assumptions 1 and 2 imply Assumption (Ã)(ii), that is, if $\varphi : \{\xi = 1\} \to \mathbb{R}$ is continuous and bounded, then the integral operator*

$$y \mapsto q(\varphi)(y) := \mathbb{E}_y\left[\varphi(Y_{S_1}) \mathbf{1}_{S_1 < S_A}\right]$$

is continuous on the set $\{0 \leq \xi \leq 1\}$.

Proof. Consider Lemma A.4. Letting $B := A \cup \{\xi > 1\}$, we may write

$$q(\varphi)(y) = \mathbb{E}_y\left[\varphi(Y_{S_B}) \mathbf{1}_{\xi(Y_{S_B}) \geq 1} \mathbf{1}_{S_B < \infty}\right].$$

The result is now a direct consequence of Lemma A.4, because Assumption 2 guarantees that $S_{\{\xi \geq 1\}} = S_{\{\xi > 1\}}$ and $S_{\bar{A}} = S_{\mathring{A}}$. Since $\bar{B} \subset \bar{A} \cup \{\xi \geq 1\}$, we deduce that $S_{\bar{B}} = S_{\mathring{B}}$. ∎

Lemma D.4. *Assumption 3 implies Assumption (B), meaning that the number of branchings on the time interval* $[0,1]$ *is almost surely finite.*

Proof. Define

$$\varepsilon := \inf_{y:\, 0\le\xi(y)\le 1} \mathbb{P}_y\left(S_1 < S_A\right) > 0.$$

Denote by $\mathcal{J}_n$ the total number of branchings of particle n during the algorithm and, as before, denote by $J_1 = \sum_{n=1}^N \mathcal{J}_n$ the total number of branchings. Clearly, we have that $\mathbb{P}(\mathcal{J}_n \ge j) \le (1-\varepsilon)^j$, so

$$\mathbb{E}[\mathcal{J}_n] = \sum_{j=1}^{\infty} \mathbb{P}(\mathcal{J}_n \ge j) \le 1/\varepsilon.$$

We conclude that $\mathbb{E}[J_1] \le N/\varepsilon < +\infty$, as desired. ∎

Appendix E. Removing the condition $S_A \wedge S_1 < +\infty$. The following property enables us to deal with transient cases where the condition $S_A \wedge S_1 < +\infty$ is not satisfied almost surely, which means that the event $S_t = S_A = +\infty$ may happen with positive probability.

Lemma E.1. *Under Assumptions 1, 2, and 3, then almost surely, either* $S_1 < +\infty$ *or* $\xi(Y_s) < 0$ *for* s *large enough, that is,* $\sup\{s \ge 0,\ \xi(Y_s) \ge 0\} < +\infty$.

Proof. Let $y = Y_0 \in \{0 \le \xi \le 1\}$ be any initial condition. By Assumption 3, we have

$$2\varepsilon := \inf_{y\in\{\xi=0\}} \mathbb{P}_y\left(S_1 < S_A\right) = \inf_{y\in\{0\le\xi\le 1\}} \mathbb{P}_y\left(S_1 < S_A\right) > 0.$$

In particular, this implies that

$$\sup_{y\in\{0\le\xi\le 1\}} \mathbb{P}_y(S_1 = +\infty) \le 1 - 2\varepsilon.$$

For each y, a simple dominated convergence argument shows that we can construct a measurable function $s : \{0 \le \xi \le 1\} \to \mathbb{N}_+$ such that

$$\sup_{y\in\{0\le\xi\le 1\}} \mathbb{P}_y\left(S_1 < s(y)\right) \le 1 - \varepsilon.$$

Consider the increasing double sequence of stopping times

$$\sigma_0^1 := 0 < \sigma_1^1 \le \sigma_0^2 < \sigma_1^2 \le \dots$$

defined for each $n \ge 1$ by

$$\begin{cases} \sigma_0^n = \inf\left\{s \ge \sigma_1^{n-1},\ Y_s \in \{\xi \ge 0\}\right\}, \\ \sigma_1^n = \sigma_0^n + s(Y_{\sigma_0^n}). \end{cases}$$

By construction, we have the implication

$$\sup\{s \ge 0,\ \xi(Y_s) \ge 0\} = +\infty \ \Rightarrow\ \sup_n \sigma_0^n < +\infty,$$

so it remains to prove that

$$\mathbb{P}(\{S_1 = +\infty\} \cap \{\sup_n \sigma_0^n < +\infty\}) = 0.$$

The strong Markov property and the definition of $s(y)$ imply that

$$\begin{aligned}\mathbb{P}(\{S_1 \geq \sigma_0^{n+1}\} \cap \{\sigma_0^{n+1} < +\infty\} | Y_{\sigma_0^n}, S_1 \geq \sigma_0^n, \sigma_0^n < +\infty) \\ \leq \mathbb{P}(S_1 \geq \sigma_1^n | Y_{\sigma_0^n}, S_1 \geq \sigma_0^n, \sigma_0^n < +\infty) \leq 1 - \varepsilon.\end{aligned}$$

Iterating the conditioning yields

$$\mathbb{P}(\{S_1 \geq \sigma_0^{n+1}\} \cap \{\sigma_0^{n+1} < +\infty\}) \leq (1-\varepsilon)^{n+1}$$

so that by σ-additivity,

$$\mathbb{P}(\{S_1 \geq \lim_n \sigma_0^n\} \cap \{\sup_n \sigma_0^n < +\infty\}) = 0.$$

The result follows since $\lim_n \sigma_0^n = +\infty$. ∎

Appendix F. A variant of Assumption 3. The following variant of Assumption 3 may be useful in practice.

Assumption 3′. *There exists $t_0 \leq 0$ such that $A \subset \{\xi < t_0\}$, the level set $\{\xi = t_0\}$ is compact, and*

$$\forall y \in \{\xi = t_0\}\,,\ \mathbb{P}_y(S_1 < S_{\bar{A}}) > 0. \tag{F.1}$$

Indeed, one has the following implication of assumptions.

Lemma F.1. *If Assumptions 1, 2, and 3′ are satisfied, then so is Assumption 3.*

Proof. First, suppose that Assumptions 1 and 2 are satisfied. We claim that the mapping

$$y \mapsto \mathbb{P}_y(S_1 < S_{\bar{A}})$$

is lower semicontinuous on $\{0 \leq \xi \leq t_0\}$, in the sense that if $y^n \to y$, then

$$\mathbb{P}_y(S_1 < S_{\bar{A}}) \leq \liminf_n \mathbb{P}_{y^n}(S_1 < S_{\bar{A}}).$$

Note that Lemma D.3 already implies that this mapping is continuous on $\{0 \leq \xi \leq 1\}$.

The proof of the claim is similar to that of Lemma A.4. Indeed, using Lemma A.2 and a Skorokhod embedding argument, a sequence $(Y_s^n)_{s\geq 0}$ of Feller processes with initial conditions $(y^n)_{n\geq 0}$ can be constructed on a single probability space so that $\lim_n Y^n = Y$ in $C(\mathbb{R}_+, E)$ almost surely, where Y denotes the Feller process with initial condition y. Then Lemma A.3 with (A.1) implies that $\lim_n S_1^n = S_1$ as well as $\liminf_n S_{\bar{A}}^n \geq S_{\bar{A}}$. But obviously,

$$\{S_1 < S_{\bar{A}}\} \subset \left\{S_1 < \liminf_n S_{\bar{A}}^n\right\} \subset \bigcup_N \bigcap_{n\geq N} \{S_1^n < S_{\bar{A}}^n\}$$

so that $\mathbb{P}\left(S_1 < S_{\bar{A}}\right) \leq \liminf_n \mathbb{P}\left(S_1^n < S_{\bar{A}}^n\right)$, and hence the claim.

Next, suppose that Assumptions 1, 2, and 3′ hold true. For any initial condition $y \in \{0 \le \xi \le 1\}$, denote $\sigma_0 := \inf\{s \ge 0,\ \xi(Y_s) = t_0\}$. By the strong Markov property for Feller processes, we may write

$$\begin{aligned}\mathbb{P}_y(S_1 < S_{\bar{A}}) &= \mathbb{E}_y\left[\mathbf{1}_{\sigma_0 < S_1}\mathbb{P}_{Y_{\sigma_0}}(S_1 < S_{\bar{A}})\right] + \mathbb{P}_y\left(S_1 < S_{\bar{A}}, S_1 \le \sigma_0\right)\\ &\ge \mathbb{P}_y(\sigma_0 < S_1) \inf_{z \in \{\xi = t_0\}} \mathbb{P}_z(S_1 < S_{\bar{A}}) + (1 - \mathbb{P}_y(\sigma_0 < S_1))\\ &\ge \inf_{z \in \{\xi = t_0\}} \mathbb{P}_z(S_1 < S_{\bar{A}}).\end{aligned}$$

Because a lower semicontinuous function on a compact reaches its infimum, and using Assumption 3′, we get that the latter infimum is > 0, and hence Assumption 3. ∎

Appendix G. Proof of Theorem 2.13. Theorem 2.13 is a pathwise extension of Theorem 2.7 and is stated under the same set of assumptions, namely Assumptions 1, 2, and 3. The proof follows along the same lines as the latter. The main difference consists of the definition of the level-indexed process and its state space, which is augmented in order to include pathwise information.

Once the appropriate definition of the level-indexed objects is set up, the interpretation of the AMS algorithm as a Fleming–Viot particle system is strictly identical to the specific case described in section 3. From there, it is then sufficient to check Assumptions (Ã) and (B) once again, thanks to Assumptions 1, 2, and 3, but in a more general pathwise context.

First, we define the extended level-indexed process as well as its state space. To do so, Definition 3.1 is generalized as follows. The extended state space $F \cup \{\partial\}$ is now defined by

$$\begin{aligned}F := \big\{&y_{[0,s]} \in C_{\mathrm{stop}}(\mathbb{R}_+, \{0 \le \xi \le 1\})\\ &\text{such that } s < +\infty,\ \xi(y_0) = 0,\ \text{ and } \forall s' \in [0, s)\,,\ \xi(y_{s'}) < \xi(y_s)\big\}\,,\end{aligned} \tag{G.1}$$

which is the set of trajectories where the maximum value of ξ is reached only at the endpoint. This specific choice of the state space F is adapted to the following construction of the level-indexed process. With an initial condition $\mathcal{X}_0 := y_{[0,s]} \in F$ being given, we define the level-indexed process as

$$\mathcal{X}_h := \begin{cases} Y_{[0, S_{(\xi(y_s)+h)\wedge 1}]} & \text{if } S_{(\xi(y_s)+h)\wedge 1} < S_A,\\ \partial & \text{if } S_{(\xi(y_s)+h)\wedge 1} \ge S_A.\end{cases}$$

In the above, we have taken as initial condition $Y_{[0,s]} = y_{[0,s]}$ in order to define the underlying Feller process Y. Note that in the simpler, usual case, where the initial condition is $\mathcal{X}_0 = y_0 \in \{\xi = 0\}$, we have

$$\mathcal{X}_t = \begin{cases} Y_{[0,S_t]} & \text{if } S_t < S_A,\\ \partial & \text{if } S_t \ge S_A.\end{cases}$$

As before, this rather complicated definition of $\mathcal{X}$ is required in order to interpret it as a time homogeneous Markov process.

Lemma G.1. *The set F defined by (G.1) is a Borel subset of the Polish space $C_{\mathrm{stop}}(\mathbb{R}_+, \{0 \le \xi \le 1\})$. $(\mathcal{X}_h)_{h \ge 0}$ is a càdlàg process taking values in $F \cup \partial$, which is time homogeneous Markov with respect to its natural filtration.*

Proof. First, F can be constructed using the countable intersection of open subsets of the form

$$\left\{ y_{[0,s]} \in C_{\text{stop}}\left(\mathbb{R}_+, \{0 \leq \xi \leq 1\}\right), \;\; \xi(y_{s'}) < \xi(y_s) \; \forall s' \leq 0 \vee (s - 1/q) \right\},$$

where $q \in \mathbb{N}^\star$. As a consequence, F is a Borel subset.

Second, as in section 3, the time homogeneous Markov property is a direct consequence of the strong Markov property of Y. ∎

We now wish to check Assumptions (Ã) and (B) in order to prove pathwise CLT Theorem 2.13.

Lemma G.2. *Under Assumptions 1, 2, and 3, Assumptions (Ã) and (B) hold true for the pathwise level-indexed process $\mathcal{X}$.*

Proof. The fact that Assumption (B) follows from Assumption 3 has already been established in Lemma D.4.

Assumption (Ã)(i) in the pathwise case is similar to the proof of Lemma D.2, which follows from Lemma D.1, namely the fact that $S_{t^-} = S_t$ almost surely. Let us give some details. Let $\mathcal{X}_0 = y_{[0,s_0]}$ be a given initial condition with initial level $t_0 = \xi(y_{s_0})$. The topology of the space $C_{\text{stop}}(\mathbb{R}_+, \{0 \leq \xi \leq 1\})$ implies that the mapping $s \mapsto y_{[0,s]} \in F$ which spans the same trajectory with different end times is continuous at $s = s_1$ if $s \mapsto y_s$ is. As a consequence, as in Lemma D.2, since Y is a continuous trajectory, and $t \mapsto S_t$ is a càdlàg increasing process, $\mathcal{X}$ is also càdlàg and has a jump at h only if $S_{(t_0+h)\wedge 1}$ has one. The proof then follows from Lemma D.1.

The only new technical point is to check Assumption (Ã)(ii), that is, the continuity of

$$y_{[0,s]} \mapsto \mathbb{E}\left[\psi(Y_{[0,S_1]})\mathbf{1}_{S_1<S_A} | Y_{[0,s]} = y_{[0,s]}\right],$$

where $y_{[0,s]} \in F$, and ψ is continuous and bounded on $C_{\text{stop}}\left(\mathbb{R}_+, \{0 \leq \xi \leq 1\}\right)$. This is a pathwise version of Lemma D.3 and, in fact, is a consequence of the pathwise continuity property stated in Lemma A.6 which follows from Assumption 2. Indeed, the latter states that if $\lim_n y^n_{[0,s^n]} = y_{[0,s]}$ is a converging sequence of initial conditions in F, and $\tilde{\psi}$ is a continuous functional on $C_{\text{stop}}(\mathbb{R}_+, \{0 \leq \xi \leq 1\})$, then $\mathbb{E}[\tilde{\psi}(Y_{[0,S_1\wedge S_A]})|Y_{[0,s]} = y^n_{[0,s^n]}]$ is converging to the corresponding limit. It remains to note that Assumption (Ã)(ii) is precisely this continuity property for the functional

$$\begin{cases} \tilde{\psi}(y_{[0,s]}) = \psi(y_{[0,s]}) & \text{if} \quad \xi(y_s) = 1, \\ \tilde{\psi}(y_{[0,s]}) = 0 & \text{if} \quad y_s \in \bar{A}, \end{cases}$$

which is indeed continuous under the C_{stop} topology.

Note that we have assumed that $S_1 \wedge S_A < +\infty$ according to Assumption 3. Otherwise, $\tilde{\psi}$ must be extended by 0, with the continuity of the extension following from Lemma E.1. ∎

Acknowledgments. The authors thank Ismael Bailleul for the reference on stochastic waves [11], and Laurent Miclo for a fruitful discussion at the beginning of this work.

REFERENCES

[1] C.-E. Bréhier, *Large deviations principle for the adaptive multilevel splitting algorithm in an idealized setting*, ALEA Lat. Am. J. Probab. Math. Stat., 12 (2015), pp. 717–742, http://alea.impa.br/articles/v12/12-27.pdf.

[2] C.-E. Bréhier, M. Gazeau, L. Goudenège, T. Lelièvre, and M. Rousset, *Unbiasedness of some generalized adaptive multilevel splitting algorithms*, Ann. Appl. Probab., 26 (2016), pp. 3559–3601, https://doi.org/10.1214/16-AAP1185.

[3] C.-E. Bréhier, L. Goudenège, and L. Tuleda, *Central limit theorem for adaptive multilevel splitting estimators in an idealized setting*, in Monte Carlo and Quasi-Monte Carlo Methods, Springer Proc. Math. Statist. 163, Springer, 2014, pp. 245–260, https://doi.org/10.1007/978-3-319-33507-0_10.

[4] C.-E. Bréhier, T. Lelièvre, and M. Rousset, *Analysis of adaptive multilevel splitting algorithms in an idealized case*, ESAIM Probab. Stat., 19 (2015), pp. 361–394, https://doi.org/10.1051/ps/2014029.

[5] F. Cérou, P. Del Moral, F. Le Gland, and P. Lezaud, *Genetic genealogical models in rare event analysis*, ALEA Lat. Am. J. Probab. Math. Stat., 1 (2006), pp. 181–203, http://alea.impa.br/articles/v1/01-08.pdf.

[6] F. Cérou, B. Delyon, A. Guyader, and M. Rousset, *A Central Limit Theorem for Fleming-Viot Particle Systems with Hard Killing*, preprint, 2017, https://arxiv.org/abs/1709.06771.

[7] F. Cérou, B. Delyon, A. Guyader, and M. Rousset, *A Central Limit Theorem for Fleming-Viot Particle Systems with Soft Killing*, preprint, 2017, https://arxiv.org/abs/1611.00515.

[8] F. Cérou and A. Guyader, *Adaptive multilevel splitting for rare event analysis*, Stoch. Anal. Appl., 25 (2007), pp. 417–443, https://doi.org/10.1080/07362990601139628.

[9] F. Cérou, A. Guyader, T. Lelièvre, and D. Pommier, *A multiple replica approach to simulate reactive trajectories*, J. Chem. Phys., 134 (2011), 054108, https://doi.org/10.1063/1.3518708.

[10] P. Del Moral, *Feynman-Kac Formulae. Genealogical and Interacting Particle Systems with Applications*, Probab. Appl., Springer-Verlag, New York, 2004, https://doi.org/10.1007/978-1-4684-9393-1.

[11] E. Dynkin and R. Vanderbei, *Stochastic waves*, Trans. Amer. Math. Soc., 275 (1983), pp. 771–779, https://doi.org/10.2307/1999052.

[12] S. Ethier and T. Kurtz, *Markov Processes: Characterization and Convergence*, John Wiley & Sons, New York, 2005, https://doi.org/10.1002/9780470316658.

[13] P. Glasserman and Y. Wang, *Counterexamples in importance sampling for large deviations probabilities*, Ann. Appl. Probab., 7 (1997), pp. 731–746, https://doi.org/10.1214/aoap/1034801251.

[14] A. Guyader, N. Hengartner, and E. Matzner-Løber, *Simulation and estimation of extreme quantiles and extreme probabilities*, Appl. Math. Optim., 64 (2011), pp. 171–196, https://doi.org/10.1007/s00245-011-9135-z.

[15] G. Hummer, *From transition paths to transition states and rate coefficients*, J. Chem. Phys., 120 (2004), pp. 516–523, https://doi.org/10.1063/1.1630572.

[16] J. Jacod and A. Shiryaev, *Limit Theorems for Stochastic Processes*, 2nd ed., Grundlehren Math. Wiss. 288, Springer, Berlin, 2003, https://doi.org/10.1007/978-3-662-05265-5.

[17] H. Kahn and T. Harris, *Estimation of particle transmission by random sampling*, National Bureau of Standards Appl. Math. Series, 12 (1951), pp. 27–30.

[18] O. Kallenberg, *Foundations of Modern Probability*, Probability and Its Applications, Springer-Verlag, New York, 2002, https://doi.org/10.1007/b98838.

[19] H. Louvin, E. Dumonteil, T. Lelièvre, M. Rousset, and C. Diop, *Adaptive multilevel splitting for Monte Carlo particle transport*, EPJ Nuclear Sci. Technol., 3 (2017), 29, https://doi.org/10.1051/epjn/2017022.

[20] P. Metzner, C. Schütte, and E. Vanden-Eijnden, *Illustration of transition path theory on a collection of simple examples*, J. Chem. Phys., 125 (2006), 084110, https://doi.org/10.1063/1.2335447.

[21] D. Pollard, *Convergence of Stochastic Processes*, Springer Ser. Statist., Springer, New York, 1984, https://doi.org/10.1007/978-1-4612-5254-2.

[22] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*, 3rd ed., Grundlehren Math. Wiss. 293, Springer, Berlin, 1999, https://doi.org/10.1007/978-3-662-06400-9.

[23] M. Rosenbluth and A. Rosenbluth, *Monte Carlo calculation of the average extension of molecular chains*, J. Chem. Phys., 23 (1955), pp. 356–359, https://doi.org/10.1063/1.1741967.

[24] F. Spitzer, *Some theorems concerning* 2*-dimensional Brownian motion*, Trans. Amer. Math. Soc., 87 (1958), pp. 187–197, https://doi.org/10.2307/1993096.

[25] I. Teo, C. Mayne, K. Schulten, and T. Lelièvre, *Adaptive multilevel splitting method for molecular dynamics calculation of benzamidine-trypsin dissociation time*, J. Chem. Theory Comput., 12 (2016), pp. 2983–2989, https://doi.org/10.1021/acs.jctc.6b00277.

[26] C. Walter, *Using Poisson Processes for Rare Event Simulation*, theses, Université Paris Diderot/Sorbonne Paris Cité, 2016, https://tel.archives-ouvertes.fr/tel-01575418.