

NUMERICAL METHODS FOR STOCHASTIC DIFFERENTIAL EQUATIONS: TWO EXAMPLES

PAUL-ÉRIC CHAUDRU DE RAYNAL¹, GILLES PAGÈS² AND CLÉMENT REY³

Abstract. The goal of this paper is to present a series of recent contributions arising in numerical probability. First we present a contribution to a recently introduced problem: stochastic differential equations with constraints in law, investigated through various theoretical and numerical viewpoints. Such a problem may appear as an extension of the famous Skorokhod problem. Then a generic method to approximate in a weak way the invariant distribution of an ergodic Feller process by a Langevin Monte Carlo simulation. It is an extension of a method originally developed for diffusions and based on the weighted empirical measure of an Euler scheme with decreasing step. Finally, we mention without details a recent development of a multilevel Langevin Monte Carlo simulation method for this type of problem.

1. INTRODUCTION

Monte Carlo simulation-based probabilistic numerical methods, which were initially developed after the World War II to solve neutron simulation problems, received renewed interest in the 1990s with the emergence of tradable derivative products in finance (option pricing and hedging). This is largely due to the conceptual simplicity of the most popular of these methods: the Monte Carlo simulation method and the direct access to more and more powerful computing means via the development of microcomputers (PCs). Gradually, the analysis of these simulation-based numerical methods – not only the Monte Carlo method but also Metropolis-Hasting algorithm, simulated annealing, particle methods, etc – became a more autonomous field under the name of Numerical Probability.

In this paper, we would like to present three recent contributions on the topic or in close connection to it.

In Section 2, P.-É. Chaudru de Raynal, introduces a new problem in the field of diffusion processes by considering a constraint in law on the process which can be seen in some sense as an extension of the Skorokhod problem. After theoretical analysis numerical method of simulation are proposed and analyzed. This contribution is based on a work with P. Briand, A. Guillin and C. Labart [2]. In Section 3, C. Rey and G. Pagès propose a simulation method to compute the invariant distribution of a (non-simulatable) Feller continuous Markov process when the increments of this process can be approximated by simulatable Markov kernels. This work appears as an extension of a series of papers devoted to the approximation of ergodic diffusion processes by an Euler scheme with decreasing steps ([10, 11, 13, 16, 19]). The proposed algorithm based on the simulation of only one path of the approximating process appears for this reason as a Langevin Monte Carlo simulation

¹ LAMA, Université Savoie Mont Blanc, pe.deraynal@univ-savoie.fr

² LPSM, UPMC, gilles.pages@upmc.fr

³ LPSM, UPMC, clement.rey@upmc.fr

method (with a possible non-constant diffusion parameter). An extended version of the presented results, including detailed proofs, is available in [18]. In Section 4, is briefly presented an application of the multilevel (see [8]) paradigm (in its weighted version, see [14]), still for the recursive approximation of the invariant distribution in a Brownian diffusion framework. This is a joint work by G. Pagès and F. Panloup [17].

2. STOCHASTIC DIFFERENTIAL EQUATION WITH CONSTRAINT IN LAW: WELL-POSEDNESS, MEAN FIELD LIMIT AND NUMERICAL SCHEME

We shortly describe stochastic differential equations constrained in law: existence and uniqueness result as well as propagation of chaos for the corresponding interacting particles system constrained in mean field. Based on these results, we propose a numerical algorithm to approximate the solution of such a system. The present exposition is based on a work done with P. Briand, A. Guillin and C. Labart [2]. Here, we will only sketch the proofs and give the main ideas of how the results are obtained and we refer the reader to the aforementioned paper for details.

2.1. The system

Stochastic Differential Equations (SDE in short) constrained in law have been introduced recently in their backward form by Briand, Elie and Hu in [3]. In that framework, the path is reflected in order to constraint its own law to live in a given set. This somehow extends the usual Skorokhod problem for SDE (see *e.g.* the work of Tanaka [23] or Lions and Sznitman [15]). From a generic point of view, the issue consists in finding, for given $T > 0$, $b, \sigma, h : \mathbb{R} \rightarrow \mathbb{R}$, $(B_s)_{s \geq 0}$ a Brownian motion defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a couple (X, K) , K being deterministic, satisfying on $[0, T]$

$$\begin{cases} X_t = \xi + \int_0^t b(X_r)dr + \int_0^t \sigma(X_r)dB_r + K_t \\ \forall t \in [0, T], \quad \mathbb{E}[h(X_t)] \geq 0, \quad \int_0^T \mathbb{E}[h(X_s)]dK_s = 0, \end{cases} \quad (1)$$

where ξ is a square integrable random variable independent of the Brownian motion and such that $\mathbb{E}[h(\xi)] \geq 0$. Since this kind of dynamic depends on its own law, such a system is called non-linear or of McKean-Vlasov type (see [22]).

2.2. Formal derivation of the equation (1) from a financial issue

Let X denote the value of some portfolio of assets. Under the classical Black & Scholes assumptions, the basket of assets included in the portfolio can be modeled as a drifted multidimensional geometric Brownian motion $\{S_t\}_{t \geq 0}$. Hence, denoting by $\{\pi_t\}_{t \geq 0}$ the investment strategy on this basket of assets, the value of the portfolio at time t , denoted by X_t , is given by $\pi_t \cdot S_t$, *i.e.*

$$X_t = X_0 + \int_0^t \pi_s dS_s,$$

when we assume standard self-financing condition of Black & Scholes Theory. Therefore, assuming that π_t depends on X_t (which means that the strategy depends on the wealth), the dynamic of X_t is given (up to a change of variable) as the solution of the following SDE:

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t. \quad (2)$$

Suppose now that the owner of this portfolio wants to guarantee that the value of his investment cannot be below a given threshold $\alpha < 0$. The rule can be applied *stricto sensu* but it can also be exceeded with a (small)

probability β . In that case, the constraint on the portfolio is of the form $\mathbb{P}(X_t \geq \alpha) \geq 1 - \beta$ *i.e.* $\mathbb{E}[h(X_t)] \geq 0$ with $h = \mathbf{1}_{x \geq \alpha} - (1 - \beta)$.

Using the terminology of the measure at risk (see *e.g.* [1]), this means that the portfolio is constrained to live in acceptable sets of the form $\mathcal{A} = \{X : \mathbb{E}[h(X)] \geq 0\}$ for a given function $h : \mathbb{R} \ni x \mapsto u(x) - p \in \mathbb{R}$, where u is some utility function from \mathbb{R} to \mathbb{R} and p is a given level of risk:

$$\forall t \geq 0, \mathbb{E}[h(X_t)] \geq 0. \tag{3}$$

Obviously, in order to satisfy this constraint, the investor must have a way to add some cash in the portfolio through the time investment period. Let us denote by K_t the amount of cash that has to be added at time t in the portfolio in order to satisfy (3) (this can also be seen as the “price to pay” to balance the risk associated to X_t). Equation (2) then becomes

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t + dK_t. \tag{4}$$

Finally, we can assume that the investor does not want to add too many “extra cash” in the portfolio in order to optimize his investment, which means that the amount K_t should be minimal. This translates, mathematically, into the so-called Skorokhod condition:

$$\int_0^t \mathbb{E}[h(X_s)]dK_s = 0, \quad t \geq 0. \tag{5}$$

Hence, putting together all our conditions ((3) and (5)) on the dynamic (4), we end up with a dynamic of the form of (1) for the portfolio.

2.3. Heuristic derivation of the reflecting process.

In the framework considered here, the explicit representation of the reflecting process K plays a central role. We hence give the heuristic derivation of such a representation. Let us denote by U the (“non reflected part” of the) dynamic of the portfolio:

$$U_t = \int_0^t b(X_r)dr + \int_0^t \sigma(X_r)dB_r, \tag{6}$$

so that $X_t = U_t + K_t$ and define for all t in $[0, T]$ the mapping $\mathbb{R} \ni x \mapsto H_t(x) = \mathbb{E}[h(x + U_t)]$. Our problem consists in finding, at any time t , the minimal quantity of cash that has to be added in the portfolio X in order to ensure that the portfolio satisfies the risk constraint *i.e.* we are looking for the solution of the equation

$$x = \inf\{y \geq 0, \text{ s.t. } H_t(y) \geq 0\},$$

i.e.

$$x = (H_t^{-1}(0))^+,$$

provided the mapping H_t is invertible for any t . Since the amount of cash added cannot be negative we obtain, setting K_t the solution at time t , that:

$$K_t = \sup_{s \leq t} (H_s^{-1}(0))^+. \tag{7}$$

2.4. Existence and uniqueness result

Existence and uniqueness result for (1) relies on usual fixed point argument and so, on Lipschitz property of the coefficients. Therefore, we suppose that the following assumptions hold

(HC) The coefficients b and σ are supposed to be Lipschitz continuous

(Hh) The mapping $\mathbb{R} \ni x \mapsto h(x) \in \mathbb{R}$ satisfies:

- h is an increasing function;

- there exist two positive constants m and M such that

$$m|x - y| \leq |h(x) - h(y)| \leq M|x - y|.$$

Under these assumptions we can prove the following result:

Theorem 2.1. *Under (HC) and (Hh) the system (1) has a unique strong solution i.e. there exists a unique pair (X, K) satisfying (1) such that K is a deterministic increasing process.*

Sketch of proof. The proof heavily relies on an explicit representation of the process K heuristically derived in the previous paragraph. Let $\mathcal{P}(\mathbb{R})$ denotes the space of probability measures on \mathbb{R} and set

$$G_0 : \mathcal{P}(\mathbb{R}) \ni \mu \mapsto G_0(\mu) = \inf\{x \geq 0, \text{ s.t. } \int h(x+z)d\mu(z) \geq 0\} \in \mathbb{R}.$$

Hence, from the representation (7) of K we get that for any t ,

$$K_t = \sup_{s \leq t} G_0(\mu_s),$$

where μ_s is the marginal at time s of U defined by (6). Then, the bi-Lipschitz property of the function h propagates to H so that the mapping $G_0 : \mathcal{P}(\mathbb{R}) \ni \mu \mapsto G_0(\mu) \in \mathbb{R}$ is a Lipschitz function w.r.t. the Wasserstein-1 distance (recall that for any probability measures ν and ν' on \mathbb{R} the Wasserstein-1 distance between ν and ν' is given by $W(\nu, \nu') = \inf_{X \sim \nu; Y \sim \nu'} \mathbb{E}[|X - Y|]$). The result then follows from a usual fixed-point procedure. \square

One can then focus on the property of the Stieljes measure dK when h is a $\mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$ function. Namely, using Itô's formula it can be shown that the following result holds.

Corollary 2.1. *The Stieljes measure dK is absolutely continuous w.r.t. the Lebesgue measure with density*

$$k : \mathbb{R}^+ \ni t \mapsto \frac{(\mathbb{E}\mathcal{L}h(X_s))^-}{\mathbb{E}h'(X_s)} \mathbf{1}_{\mathbb{E}h(X_t)=0} \in \mathbb{R}.$$

This in particular ensures that, in that smooth setting, the SDE constrained in law is a usual McKean-Vlasov SDE with singular law dependence and this emphasizes as well the importance of the non negative assumption on h' (or equivalently the bi-Lipschitz property of h).

2.5. Constrained in law SDE as mean field limit of interacting particles system

The analogy of this kind of system with the usual McKean-Vlasov SDE suggests to consider the mean field limit of the corresponding particles system. Consider for instance the following particles constrained in mean field:

$$\begin{cases} X_t^i = \int_0^t b(X_s^i)ds + \int_0^t \sigma(X_s^i)dB_s^i + K_t^N, & 1 \leq i \leq N, \\ \frac{1}{N} \sum_{i=1}^N h(X_t^i) \geq 0, & \frac{1}{N} \sum_{i=1}^N \int_0^t h(X_s^i) dK_s^N = 0. \end{cases} \quad (8)$$

One can also use the analogous representation for the process K^N :

$$K_t^N = \sup_{s \leq t} G_0(\mu_s^N), \quad \mu_s^N = \frac{1}{N} \sum_{i=1}^N \delta_{U_s^i},$$

where

$$U_s^i = \int_0^s b(X_r^i) dr + \int_0^s \sigma(X_r^i) dB_r^i, \quad 1 \leq i \leq N,$$

and show that this system is well posed using fixed point argument as well.

Then, thanks to the explicit representation of K and this representation of K^N , we can implement a coupling argument as proposed in [22] and prove that the chaos propagates in (8) (the proof again relies on the Lipschitz properties of G_0) so that, denoting for any i in $\{1, \dots, N\}$ by \bar{X}^i the i.i.d. copies of (1) having the same Brownian motion as the i^{th} particle, we obtain that

Theorem 2.2. *Suppose that assumptions **(Hc)** and **(Hh)** are in force and suppose moreover that there exists $p > 4$ such that $\mathbb{E}|\xi|^p < \infty$. Then*

(1) *there exists a $C := C(\mathbf{(Hc)}, \mathbf{(Hh)}, \mathbb{E}|\xi|^p, T) > 0$ such that for all t in $[0, T]$ and any j in $\{1, \dots, N\}$:*

$$\mathbb{E}|X_t^j - \bar{X}_t^j|^2 \leq CN^{-1/2};$$

(2) *if, in addition, the function h is in $\mathcal{C}_b^2(\mathbb{R}, \mathbb{R})$, there exists a $C := C(\mathbf{(Hc)}, \mathbf{(Hh)}, \mathbb{E}|\xi|^p, T, \|h'\|_\infty, \|h''\|_\infty) > 0$ such that for all t in $[0, T]$ and for any j in $\{1, \dots, N\}$:*

$$\mathbb{E}|X_t^j - \bar{X}_t^j|^2 \leq CN^{-1}.$$

As suggested above, the proof is an adaptation to our framework of the coupling argument of [22]. Let us only focus on the two rates of convergence obtained. In the proof, the critical quantity which gives the rate is:

$$\sup_{s \leq t} W(\bar{\mu}_s^N, \mu_s),$$

where W is the Wasserstein-1 distance and where $\bar{\mu}_s^N$ is the empirical measure generated by the “non reflected part” of the $(\bar{X}^i)_{1 \leq i \leq N}$. We know from [21] and [7] that the rate of convergence of empirical measure of i.i.d. random variables to their law is of order $N^{-1/2}$. Moreover this rate is, in full generality, the optimal one. Especially, in our framework, it is not damaged by the supremum in time. In the second result, we then use the smoothness property of h and Itô’s formula to take benefit from the fact that the family of empirical measures $(\bar{\mu}_s^N)_{0 \leq s \leq T}$ is related to i.i.d. diffusion processes. This is how a better estimate is obtained.

2.6. On the numerical simulation of constrained in law SDE

As a direct application of that propagation of chaos result, one can design a numerical approximation of (1). The basic idea consists in using a classical Euler scheme for the numerical simulation of the particle system (8).

Let $0 = T_0 < T_1 < \dots < T_n = T$ be a subdivision of $[0, T]$. Given this subdivision, we denote by “ $\underline{\cdot}$ ” the mapping $s \mapsto \underline{s} = T_k$ if $s \in [T_k, T_{k+1})$, $k \in \{0, \dots, n-1\}$. For simplicity, we consider only the case of regular subdivisions: for a given integer n , $T_k = kT/n$, $k = 0, \dots, n$. We introduce the following discrete version of the particle system

$$\tilde{X}_t^i = \xi^i + \int_0^t b(\tilde{X}_{\underline{s}}^i) ds + \int_0^t \sigma(\tilde{X}_{\underline{s}}^i) dB_s^i + \sup_{s \leq t} G_0(\tilde{\mu}_{\underline{s}}^N), \quad 1 \leq i \leq N,$$

with the notation

$$\tilde{U}_t^i = \xi_0^i + \int_0^t b(\tilde{X}_{\underline{s}}^i) ds + \int_0^t \sigma(\tilde{X}_{\underline{s}}^i) dB_s^i, \quad 1 \leq i \leq N, \quad \tilde{\mu}_s^N = \frac{1}{N} \sum_{i=1}^N \delta_{\tilde{U}_s^i}.$$

Then, our numerical algorithm works as follows.

Algorithm [A]: particle approximation of (8)

- For $1 \leq j \leq N$ do

$$- \left(\left(\tilde{X}_0^{\tilde{\mu}^N} \right)^j, \left(\tilde{U}_0^{\tilde{\mu}^N} \right)^j, \tilde{\mu}_0^N \right) = (x, x, \delta_x)$$

end for

- For $1 \leq k \leq n$ do

- For $1 \leq j \leq N$ do

$$* G^j \sim \mathcal{N}(0, 1)$$

$$* \left(\tilde{U}_{T_k}^{\tilde{\mu}^N} \right)^j = \left(\tilde{U}_{T_{k-1}}^{\tilde{\mu}^N} \right) + (T/n)b \left(\left(\tilde{X}_{T_{k-1}}^{\tilde{\mu}^N} \right)^j \right) + \sqrt{(T/n)\sigma} \left(\left(\tilde{X}_{T_{k-1}}^{\tilde{\mu}^N} \right)^j \right) G^j$$

end for

$$- \tilde{\mu}_{T_k}^N = N^{-1} \sum_{j=1}^N \delta_{\left(\tilde{U}_{T_k}^{\tilde{\mu}^N} \right)^j}$$

$$- \Delta_k \hat{K}^N = \sup_{l \leq k} G_0(\tilde{\mu}_{T_l}^N) - \sup_{l \leq k-1} G_0(\tilde{\mu}_{T_l}^N)$$

- For $1 \leq j \leq N$ do

$$* \left(\tilde{X}_{T_k}^{\tilde{\mu}^N} \right)^j = \left(\tilde{X}_{T_{k-1}}^{\tilde{\mu}^N} \right)^j + (T/n)b \left(\left(\tilde{X}_{T_{k-1}}^{\tilde{\mu}^N} \right)^j \right) + \sqrt{(T/n)\sigma} \left(\left(\tilde{X}_{T_{k-1}}^{\tilde{\mu}^N} \right)^j \right) G^j + \Delta_k \hat{K}^N$$

end for

end for

We obtain the following strong error bounds on the resulting approximation process.

Theorem 2.3. *Suppose that assumptions **(Hc)** and **(Hh)** are in force and suppose moreover that there exists $p > 4$ such that $\mathbb{E}|\xi|^p < \infty$, then the algorithm **[A]** above leads to an approximation of order*

$$(1) (T/n)^{1/2} + N^{-1/2};$$

$$(2) (T/n)^{1/2} + N^{-1} \text{ if in addition } h \text{ is in } \mathcal{C}_b^2(\mathbb{R}, \mathbb{R}).$$

3. RECURSIVE COMPUTATION OF THE INVARIANT DISTRIBUTIONS OF A FELLER PROCESS

In this paper, we propose a method for the recursive computation of the invariant distribution (denoted ν) of a Feller processes $(X_t)_{t \geq 0}$. The starting idea is to consider a non-homogeneous discrete Markov process which can be simulated using a family of transitions kernels $(Q_\gamma)_{\gamma > 0}$ and approximating $(X_t)_{t \geq 0}$ in a sense made precise further on.

As suggested by the pointwise Birkhoff ergodic theorem, we then show that some sequence $(\nu_n)_{n \in \mathbb{N}^*}$ of random empirical measures *a.s.* weakly converges toward ν under some appropriate mean-reverting and moment assumptions. An abstract framework is developed which, among others, enables to extend this convergence to the L^p -Wasserstein distance. For a given f , $\nu_n(f)$ can be recursively defined making its computation straightforward.

3.1. Convergence to invariant distributions - A general approach

In this section, we build a sequence of empirical measures from an approximation $(\bar{X}_{\Gamma_n}^\gamma)_{n \in \mathbb{N}}$ of a Feller process $(X_t)_{t \geq 0}$ (which are not specified explicitly), where $\Gamma_n = \sum_{k=1}^n \gamma_k$ is the time grid such that the step sequence $(\gamma_n)_{n \in \mathbb{N}^*} \xrightarrow{n \rightarrow +\infty} 0$. We show how it *a.s.* weakly converges to the set \mathcal{V} , of the invariant distributions of $(X_t)_{t \geq 0}$. To this end, we will provide as weak as possible mean reverting assumptions on the pseudo-generator of $(\bar{X}_{\Gamma_n}^\gamma)_{n \in \mathbb{N}}$ on the one hand and appropriate rate conditions on the step sequence $(\gamma_n)_{n \in \mathbb{N}^*}$ on the other hand.

3.2. Presentation of the abstract framework

3.2.1. Construction of the random measures

Let $(\Omega, \mathcal{G}, \mathbb{P})$ be a probability space. We consider a Feller process $(X_t)_{t \geq 0}$ (see [6] for details) on $(\Omega, \mathcal{G}, \mathbb{P})$ taking values in a locally compact and separable metric space E . We denote by $(P_t)_{t \geq 0}$ the Feller semigroup

(see [20]) of this process. We recall that $(P_t)_{t \geq 0}$ is a family of linear operators from $\mathcal{C}_0(E)$ to itself such that $P_0 f = f$, $P_{t+s} f = P_t P_s f$, $t, s \geq 0$ (semigroup property) and $\lim_{t \rightarrow 0} \|P_t f - f\|_\infty = 0$ (Feller property). Using this semigroup, we can introduce the infinitesimal generator of $(X_t)_{t \geq 0}$ as a linear operator A defined on a subspace $\mathcal{D}(A)$ of $\mathcal{C}_0(E)$, satisfying: For every $f \in \mathcal{D}(A)$,

$$Af = \lim_{t \rightarrow 0} \frac{P_t f - f}{t} \quad (9)$$

exists for the $\|\cdot\|_\infty$ -norm. The operator $A : \mathcal{D}(A) \rightarrow \mathcal{C}_0(E)$ is thus well defined and $\mathcal{D}(A)$ is called the domain of A . From the Echeverria Weiss theorem (see Theorem 9.17 in [5]), the set of invariant distributions for $(X_t)_{t \geq 0}$ can be characterized in the following way:

$$\mathcal{V} = \{\nu \in \mathcal{P}(E), \forall t \geq 0, P_t \nu = \nu\} = \{\nu \in \mathcal{P}(E), \forall f \in \mathcal{D}(A), \nu(Af) = 0\}.$$

The starting point of our reasoning is thus to consider an approximation of A . First, we introduce the family of transition kernels $(\mathcal{Q}_\gamma)_{\gamma > 0}$ from $\mathcal{C}_0(E)$ to itself. Now, let us define the family of linear operators $\tilde{A} := (\tilde{A}_\gamma)_{\gamma > 0}$ from $\mathcal{C}_0(E)$ into itself, as follows

$$\forall f \in \mathcal{C}_0(E), \quad \gamma > 0, \quad \tilde{A}_\gamma f = \frac{\mathcal{Q}_\gamma f - f}{\gamma}. \quad (10)$$

The family \tilde{A} is usually called the pseudo-generator of the transition kernels $(\mathcal{Q}_\gamma)_{\gamma > 0}$ and is an approximation of A as γ tends to zero. From a practical viewpoint, the main interest of our approach is that we can consider that there exists $\bar{\gamma} > 0$ such that for every $x \in E$ and every $\gamma \in [0, \bar{\gamma}]$, $\mathcal{Q}_\gamma(x, dy)$ is simulable at a reasonable computational cost. We use the family $(\mathcal{Q}_\gamma)_{\gamma > 0}$, to build $(\bar{X}_{\Gamma_n})_{n \in \mathbb{N}}$ (this notation replaces $(\bar{X}_{\Gamma_n}^\gamma)_{n \in \mathbb{N}}$ from now on for clarity in the writing) as the non-homogeneous Markov approximation of the Feller process $(X_t)_{t \geq 0}$. It is defined on the time grid $\{\Gamma_n = \sum_{k=1}^n \gamma_k, n \in \mathbb{N}\}$ with the sequence $\gamma := (\gamma_n)_{n \in \mathbb{N}^*}$ of time step satisfying

$$\forall n \in \mathbb{N}^*, \quad 0 < \gamma_n \leq \bar{\gamma} := \sup_{n \in \mathbb{N}^*} \gamma_n < +\infty, \quad \lim_{n \rightarrow +\infty} \gamma_n = 0 \quad \text{and} \quad \lim_{n \rightarrow +\infty} \Gamma_n = +\infty. \quad (11)$$

Its transition probability distributions are given by $\mathcal{Q}_{\gamma_n}(x, dy)$, $n \in \mathbb{N}^*$, $x \in E$, *i.e.* :

$$\mathbb{P}(\bar{X}_{\Gamma_{n+1}} \in dy | \bar{X}_{\Gamma_n}) = \mathcal{Q}_{\gamma_{n+1}}(\bar{X}_{\Gamma_n}, dy), \quad n \in \mathbb{N}. \quad (12)$$

We can canonically extend $(\bar{X}_{\Gamma_n})_{n \in \mathbb{N}}$ into a *càdlàg* process by setting $\bar{X}(t, \omega) = \bar{X}_{\Gamma_{n(t)}}(\omega)$ with $n(t) = \inf\{n \in \mathbb{N}, \Gamma_{n+1} > t\}$. Then $(\bar{X}_{\Gamma_n})_{n \in \mathbb{N}}$ is a simulable (as soon as \bar{X}_0 is) non-homogeneous Markov chain with transitions

$$\forall m \leq n, \quad \bar{P}_{\Gamma_m, \Gamma_n}(x, dy) = \mathcal{Q}_{\gamma_{m+1}} \circ \dots \circ \mathcal{Q}_{\gamma_n}(x, dy),$$

and law

$$\mathcal{L}(\bar{X}_{\Gamma_n} | \bar{X}_0 = x) = \bar{P}_{\Gamma_n}(x, dy) = \mathcal{Q}_{\gamma_1} \circ \dots \circ \mathcal{Q}_{\gamma_n}(x, dy).$$

We use $(\bar{X}_{\Gamma_n})_{n \in \mathbb{N}}$ to design a Langevin Monte Carlo algorithm. Notice that this approach is generic since the approximation transition kernels $(\mathcal{Q}_\gamma)_{\gamma > 0}$ are not explicitly specified and then, it can be used in many different configurations including among others, weak numerical schemes or exact simulation *i.e.* $(\bar{X}_{\Gamma_n})_{n \in \mathbb{N}} = (X_{\Gamma_n})_{n \in \mathbb{N}}$. In particular, using high weak order schemes for $(X_t)_{t \geq 0}$ may lead to higher rates of convergence for the empirical measures. The approach we use to build the empirical measures is quite general than since we consider some

general weights which are not necessarily equal to the time steps. This generalization is motivated by technical reasons related to the analysis of the (weak) convergence rate not investigated in this note (see e.g. [10]). We define this weight sequence. Let $\eta := (\eta_n)_{n \in \mathbb{N}^*}$ be such that

$$\forall n \in \mathbb{N}^*, \quad \eta_n \geq 0, \quad \lim_{n \rightarrow +\infty} H_n = +\infty, \quad \text{with} \quad H_n = \sum_{k=1}^n \eta_k. \quad (13)$$

Now we present our algorithm adapted from the one introduced in [10] designed with a Euler scheme with decreasing step $(\bar{X}_{\Gamma_n})_{n \in \mathbb{N}}$ of a Brownian diffusion process $(X_t)_{t \geq 0}$. For $x \in E$, let δ_x denote the Dirac mass at point x . For every $n \in \mathbb{N}^*$, we define the random weighted empirical random measures as follows

$$\nu_n^\eta(dx) = \frac{1}{H_n} \sum_{k=1}^n \eta_k \delta_{\bar{X}_{\Gamma_{k-1}}}(dx). \quad (14)$$

This paper is dedicated to show that *a.s.* every weak limiting distribution of $(\nu_n^\eta)_{n \in \mathbb{N}^*}$ belongs to \mathcal{V} . In particular when the invariant measure of $(X_t)_{t \geq 0}$ is unique, *i.e.* $\mathcal{V} = \{\nu\}$, we show that \mathbb{P} -*a.s.* $\lim_{n \rightarrow +\infty} \nu_n^\eta f = \nu f$, for a generic class of continuous test functions f . The approach we develop can be decomposed into two steps. First, we establish a tightness property to obtain existence of at least one weak limiting distribution for $(\nu_n^\eta)_{n \in \mathbb{N}^*}$. Then, in a second step, we identify everyone of these limiting distributions with an invariant distributions of the Feller process $(X_t)_{t \geq 0}$ exploiting the Echeverria Weiss theorem (see [5] Theorem 9.17).

3.2.2. Assumptions on the random measures

In this part, we present the necessary assumptions on the pseudo-generator $\tilde{A} = (\tilde{A}_\gamma)_{\gamma > 0}$ in order to prove the convergence of the empirical measures $(\nu_n^\eta)_{n \in \mathbb{N}^*}$.

Recursive control. In our framework, we introduce a well suited assumption, referred to as the mean reverting recursive control of the pseudo-generator \tilde{A} , that leads to a tightness property on $(\nu_n^\eta)_{n \in \mathbb{N}^*}$ from which follows the existence (in weak sense) of a limiting distribution for $(\nu_n^\eta)_{n \in \mathbb{N}^*}$. A supplementary interest of our approach is that it is designed to obtain the *a.s.* convergence of $(\nu_n^\eta(f))_{n \in \mathbb{N}^*}$ for a generic class of continuous test functions f which is larger than $\mathcal{C}_b(E)$. To do so, we introduce a Lyapunov function V related to $(\bar{X}_{\Gamma_n})_{n \in \mathbb{N}}$. Assume that V a Borel function such that

$$\mathcal{L}_V \equiv V : (E \rightarrow [v_*, +\infty), v_* > 0 \quad \text{and} \quad \lim_{x \rightarrow \infty} V(x) = +\infty. \quad (15)$$

We now relate V to $(\bar{X}_{\Gamma_n})_{n \in \mathbb{N}}$ introducing its mean reversion Lyapunov property. Let $\psi, \phi : [v_*, \infty) \rightarrow (0, +\infty)$ some Borel functions such that $\tilde{A}_\gamma \psi \circ V$ exists for every $\gamma \in (0, \bar{\gamma}]$. Let $\alpha > 0$ and $\beta \in \mathbb{R}$. We assume

$$\mathcal{RC}_{Q,V}(\psi, \phi, \alpha, \beta) \equiv \begin{cases} (i) & \exists n_0 \in \mathbb{N}^*, \forall n \geq n_0, x \in E, \quad \tilde{A}_{\gamma_n} \psi \circ V(x) \leq \frac{\psi \circ V(x)}{V(x)} (\beta - \alpha \phi \circ V(x)). \\ (ii) & \liminf_{y \rightarrow +\infty} \phi(y) > \beta/\alpha. \end{cases} \quad (16)$$

Lyapunov functions are usually used to show the existence and sometimes the uniqueness of the invariant measure of Feller processes. We refer to the extensive literature on the topic for more details: See for instance [9], [5] or [16].

Notice that the condition $\mathcal{RC}_{Q,V}(I_d, \phi, \alpha, \beta)(i)$ with ϕ concave appears in [4] to prove sub-geometrical ergodicity of Markov chains. In [12], a similar hypothesis to $\mathcal{RC}_{Q,V}(I_d, \phi, \alpha, \beta)(i)$, with ϕ not necessarily concave, is also used (with \tilde{A}_{γ_n} replaced by A) to study the convergence of the weighted empirical measures (14) for the Euler scheme of a Brownian diffusion. The function ϕ controls the mean reverting property. In particular, we call strongly mean reverting property when $\phi = I_d$ and weakly mean reverting property when $\lim_{y \rightarrow +\infty} \phi(y)/y = 0$,

for instance $\phi(y) = y^a$, $a \in (0, 1)$ for every $y \in [v_*, \infty)$. The function ψ is closely related to the identification of the set of test functions f for which we have $\lim_{n \rightarrow +\infty} \nu_n^\eta(f) = \nu(f)$ *a.s.*, when ν is the unique invariant distribution of the underlying Feller process. To this end, for $s \geq 1$, which is related to step weight assumption, we introduce the sets of test functions for which we will show the *a.s.* convergence of the weighted empirical measures (14):

$$\mathcal{C}_{\tilde{V}_{\psi, \phi, s}}(E) = \left\{ f \in \mathcal{C}(E), |f(x)| = o_{x \rightarrow \infty}(\tilde{V}_{\psi, \phi, s}(x)) \right\}, \quad (17)$$

$$\text{with } \tilde{V}_{\psi, \phi, s} : E \rightarrow \mathbb{R}_+, x \mapsto \tilde{V}_{\psi, \phi, s}(x) := \frac{\phi \circ V(x) \psi \circ V(x)^{1/s}}{V(x)}.$$

Notice that our approach benefits from providing generic results because we consider general Feller processes and approximations but also because the functions ϕ and ψ are not specified explicitly.

Remark 3.1. *The above assumption is devised to be checked on time discretization schemes of a Feller process that will be used in practice for practical simulation, typically some schemes of Euler type with decreasing steps. However it always corresponds to an “underlying” – and more intuitive – mean-reverting assumption satisfied by the Feller process may appear somewhat hidden by this formulation. In fact, in seminal studies such as in [10] (see also [13], [19] for related works), the results are established (for the Euler scheme of a diffusion process, possible with jumps or locally Lipschitz coefficients) under the natural mean-reverting assumption*

$$AV(x) \leq \beta - \alpha V^a(x), \quad a \in (0, 1]. \quad (18)$$

Its version with the pseudo-infinitesimal generator A_{γ_n} corresponds to Assumption $\mathcal{RC}_{Q,V}(\psi, \phi, \alpha, \beta)$ (see 16)) when $\Psi = I_d$ and $\phi(y) = |y|^a$. In fact, under the infinitesimal generator approximation assumption (19) thereafter and suitable hypothesis on the regularity of $(X_t)_{t \geq 0}$ (related to square integrability of its increments for the empirical measures), (16) follows from (18), where ψ is a polynomial or an exponential function and $\phi(y) = y^a$, $a \in (0, 1]$. This connection is detailed in [18].

Infinitesimal generator approximation. This section presents the assumption that enables to characterize the limiting distributions of the *a.s.* tight sequence $(\nu_n^\eta(dx, \omega))_{n \in \mathbb{N}^*}$. We aim to estimate the distance between \mathcal{V} and ν_n^η (see (14)) for n large enough. We thus introduce an hypothesis concerning the distance between $(\tilde{A}_\gamma)_{\gamma > 0}$, the pseudo-generator of $(Q_\gamma)_{\gamma > 0}$, and A , the infinitesimal generator of $(P_t)_{t \geq 0}$. We assume that there exists $\mathcal{D}(A)_0 \subset \mathcal{D}(A)$ with $\mathcal{D}(A)_0$ dense in $\mathcal{C}_0(E)$ such that:

$$\mathcal{E}(\tilde{A}, A, \mathcal{D}(A)_0) \equiv \quad \forall \gamma \in (0, \bar{\gamma}], \forall f \in \mathcal{D}(A)_0, \forall x \in E, \quad |\tilde{A}_\gamma f(x) - Af(x)| \leq \Lambda_f(x, \gamma), \quad (19)$$

where $\Lambda_f : E \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$ can be represented in the following way: Let $(\tilde{\Omega}, \tilde{\mathcal{G}}, \tilde{\mathbb{P}})$ be a probability space. Let $g : E \rightarrow \mathbb{R}_+^q$, $q \in \mathbb{N}$, be a locally bounded Borel measurable function and let $\tilde{\Lambda}_f : (E \times \mathbb{R}_+ \times \tilde{\Omega}, \mathcal{B}(E) \otimes \mathcal{B}(\mathbb{R}_+) \otimes \tilde{\mathcal{G}}) \rightarrow \mathbb{R}_+^q$ be a measurable function such that $\sup_{i \in \{1, \dots, q\}} \tilde{\mathbb{E}}[\sup_{x \in E} \sup_{\gamma \in (0, \bar{\gamma}]} \tilde{\Lambda}_{f,i}(x, \gamma, \tilde{\omega})] < +\infty$ and

$$\forall x \in E, \forall \gamma \in (0, \bar{\gamma}], \quad \Lambda_f(x, \gamma) = \langle g(x), \tilde{\mathbb{E}}[\tilde{\Lambda}_f(x, \gamma, \tilde{\omega})] \rangle_{\mathbb{R}^q}$$

Moreover, we assume that for every $i \in \{1, \dots, q\}$, $\sup_{n \in \mathbb{N}^*} \nu_n^\eta(g_i, \omega) < +\infty$, $\mathbb{P}(d\omega) - a.s.$, and that $\tilde{\Lambda}_{f,i}$ satisfies one of the following two properties:

There exists a measurable function $\underline{\gamma} : (\tilde{\Omega}, \tilde{\mathcal{G}}) \rightarrow ((0, \bar{\gamma}], \mathcal{B}((0, \bar{\gamma}]))$ such that:

$$\tilde{\mathbb{P}}(d\tilde{\omega})\text{-a.s.} \quad \left\{ \begin{array}{l} (i) \quad \forall K \in \mathcal{K}_E, \quad \lim_{\gamma \rightarrow 0} \sup_{x \in K} \tilde{\Lambda}_{f,i}(x, \gamma, \tilde{\omega}) = 0, \\ (ii) \quad \lim_{x \rightarrow \infty} \sup_{\gamma \in (0, \underline{\gamma}(\tilde{\omega})]} \tilde{\Lambda}_{f,i}(x, \gamma, \tilde{\omega}) = 0, \end{array} \right.$$

or, for the uniform convergence setting,

$$\tilde{\mathbb{P}}(d\tilde{\omega}) - a.s. \quad \limsup_{\gamma \rightarrow 0} \sup_{x \in E} \tilde{\Lambda}_{f,i}(x, \gamma, \tilde{\omega}) g_i(x) = 0.$$

with \mathcal{K}_E the set of compact subsets of E .

Remark 3.2. Let $(F, \mathcal{F}, \lambda)$ be a measurable space. Using the exact same approach, the results we obtain hold when we replace the probability space $(\tilde{\Omega}, \tilde{\mathcal{G}}, \tilde{\mathbb{P}})$ by the product measurable space $(\tilde{\Omega} \times F, \tilde{\mathcal{G}} \otimes \mathcal{F}, \tilde{\mathbb{P}} \otimes \lambda)$ in the above representations of Λ_f but we restrict to this case for sake of clarity. This observation can be useful when we study jump process where λ stands for the jump intensity.

This representation assumption benefits from the fact that the transition functions $(\mathcal{Q}_\gamma(x, dy))_{\gamma \in (0, \bar{\gamma}]}$, $x \in E$, can be represented using distributions of random variables which are involved in the computation of $(\bar{X}_{\Gamma_n})_{n \in \mathbb{N}^*}$. In particular, this approach is well adapted to numerical schemes associated to a time grid (implemented here with a decreasing timestep) such as Euler, Milstein, Ninomiya-Victoir... for stochastic differential equations with a Brownian part or/and a jump component.

Growth control and Step Weight assumptions. We conclude with hypothesis concerning the control of the martingale part of one step of our approximation. Let $\rho \in [1, 2]$ and let $\epsilon_{\mathcal{I}} : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ be an increasing function. For $F \subset \{f, f : (E, \mathcal{B}(E)) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))\}$ and $g : E \rightarrow \mathbb{R}_+$ a Borel function, we assume that, for every $n \in \mathbb{N}$,

$$\mathcal{GC}_Q(F, g, \rho, \epsilon_{\mathcal{I}}) \equiv \mathbb{P} - a.s. \quad \forall f \in F, \quad \mathbb{E}[|f(\bar{X}_{\Gamma_{n+1}}) - \mathcal{Q}_{\gamma_{n+1}} f(\bar{X}_{\Gamma_n})|^\rho | \bar{X}_{\Gamma_n}] \leq C_f \epsilon_{\mathcal{I}}(\gamma_{n+1}) g(\bar{X}_{\Gamma_n}), \quad (20)$$

with $C_f > 0$ a finite constant which may depend on f . We will combine this assumption with the following step weight related ones:

$$SW_{\mathcal{I}, \gamma, \eta}(g, \rho, \epsilon_{\mathcal{I}}) \equiv \mathbb{P} - a.s. \quad \sum_{n=1}^{\infty} \left| \frac{\eta_n}{H_n \gamma_n} \right|^\rho \epsilon_{\mathcal{I}}(\gamma_n) g(\bar{X}_{\Gamma_n}) < +\infty. \quad (21)$$

Remark 3.3. The reader may notice that $\mathcal{GC}_Q(F, g, \rho, \epsilon_{\mathcal{I}})$ holds as soon as (20) is satisfied with $\mathcal{Q}_{\gamma_{n+1}} f(\bar{X}_{\Gamma_n})$, $n \in \mathbb{N}^*$, replaced by a $\mathcal{F}_n^{\bar{X}} := \sigma(\bar{X}_{\Gamma_k}, k \leq n)$ - progressively measurable process $(\bar{\mathfrak{X}}_n)_{n \in \mathbb{N}^*}$ since we have $\mathcal{Q}_{\gamma_{n+1}} f(\bar{X}_{\Gamma_n}) = \mathbb{E}[f(\bar{X}_{\Gamma_{n+1}}) | \bar{X}_{\Gamma_n}]$ and $\mathbb{E}[|f(\bar{X}_{\Gamma_{n+1}}) - \mathcal{Q}_{\gamma_{n+1}} f(\bar{X}_{\Gamma_n})|^\rho | \bar{X}_{\Gamma_n}] \leq 2^\rho \mathbb{E}[|f(\bar{X}_{\Gamma_{n+1}}) - \bar{\mathfrak{X}}_n|^\rho | \bar{X}_{\Gamma_n}]$ for every $\bar{\mathfrak{X}}_n \in L^2(\mathcal{F}_n^{\bar{X}})$.

We will also use the hypothesis

$$SW_{II, \gamma, \eta}(F) \equiv \mathbb{P} - a.s. \quad \forall f \in F, \quad \sum_{n=0}^{\infty} \frac{(\eta_{n+1}/\gamma_{n+1} - \eta_n/\gamma_n)_+}{H_{n+1}} |f(\bar{X}_{\Gamma_n})| < +\infty, \quad (22)$$

with the convention $\eta_0/\gamma_0 = 1$. Notice that this last assumption holds as soon as the sequence $(\eta_n/\gamma_n)_{n \in \mathbb{N}^*}$ is non-increasing.

At this point we can focus now on the main results concerning this general approach.

3.3. Convergence

3.3.1. Almost sure tightness

From the recursive control assumption, the following Theorem establish the *a.s.* tightness of the sequence $(\nu_n^\eta)_{n \in \mathbb{N}^*}$ and also provides a uniform control of $(\nu_n^\eta)_{n \in \mathbb{N}^*}$ on a generic class of test functions.

Theorem 3.1. Let $s \geq 1$, $\rho \in [1, 2]$, $v_* > 0$, and let us consider the Borel functions $V : E \rightarrow [v_*, \infty)$, $g : E \rightarrow \mathbb{R}_+$, $\psi : [v_*, \infty) \rightarrow \mathbb{R}_+$ and $\epsilon_{\mathcal{I}} : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ an increasing function. We have the following properties:

A. Assume that $\tilde{A}_{\gamma_n}(\psi \circ V)^{1/s}$ exists for every $n \in \mathbb{N}^*$, and that $\mathcal{GC}_Q(\{(\psi \circ V)^{1/s}, g, \rho, \epsilon_{\mathcal{I}}\})$ (see (20)), $\mathcal{SW}_{\mathcal{I}, \gamma, \eta}(g, \rho, \epsilon_{\mathcal{I}})$ (see (21)) and $\mathcal{SW}_{\mathcal{I}, \gamma, \eta}(\{(\psi \circ V)^{1/s}\})$ (see (22)) hold. Then

$$\mathbb{P}\text{-a.s.} \quad \sup_{n \in \mathbb{N}^*} -\frac{1}{H_n} \sum_{k=1}^n \eta_k \tilde{A}_{\gamma_k}(\psi \circ V)^{1/s}(\bar{X}_{\Gamma_{k-1}}) < +\infty. \quad (23)$$

B. Let $\alpha > 0$ and $\beta \in \mathbb{R}$. Let $\phi : [v_*, \infty) \rightarrow \mathbb{R}_+^*$ be a continuous function such that $C_\phi := \sup_{y \in [v_*, \infty)} \phi(y)/y < +\infty$. Assume that (23) holds and

i. $\mathcal{RC}_{Q,V}(\psi, \phi, \alpha, \beta)$ (see (16)) holds.

ii. L_V (see (15)) holds and $\lim_{y \rightarrow +\infty} \frac{\phi(y)\psi(y)^{1/s}}{y} = +\infty$.

Then,

$$\mathbb{P}\text{-a.s.} \quad \sup_{n \in \mathbb{N}^*} \nu_n^\eta(\tilde{V}_{\psi, \phi, s}) < +\infty. \quad (24)$$

with $\tilde{V}_{\psi, \phi, s}$ defined in (17). Therefore, the sequence $(\nu_n^\eta)_{n \in \mathbb{N}^*}$ is \mathbb{P} -a.s. tight.

3.3.2. Identification of the limit

In Theorem 3.1, we obtained the tightness of $(\nu_n^\eta)_{n \in \mathbb{N}^*}$. It remains to show that every limiting point of this sequence is an invariant distribution of the Feller process with infinitesimal generator A . This is the interest of the following Theorem which relies on the infinitesimal generator approximation and the Echeverria Weiss theorem (see [5] Theorem 9.17).

Theorem 3.2. Let $\rho \in [1, 2]$. We have the following properties:

A. Let $\mathcal{D}(A)_0 \subset \mathcal{D}(A)$, with $\mathcal{D}(A)_0$ dense in $\mathcal{C}_0(E)$. We assume that $\tilde{A}_{\gamma_n} f$ exists for every $f \in \mathcal{D}(A)_0$ and every $n \in \mathbb{N}^*$. Also assume that there exists $g : E \rightarrow \mathbb{R}_+$ a Borel function and $\epsilon_{\mathcal{I}} : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ an increasing function such that $\mathcal{GC}_Q(\mathcal{D}(A)_0, g, \rho, \epsilon_{\mathcal{I}})$ (see (20)) and $\mathcal{SW}_{\mathcal{I}, \gamma, \eta}(g, \rho, \epsilon_{\mathcal{I}})$ (see (21)) hold and that

$$\lim_{n \rightarrow +\infty} \frac{1}{H_n} \sum_{k=1}^n |\eta_{k+1}/\gamma_{k+1} - \eta_k/\gamma_k| = 0.$$

Then

$$\mathbb{P}\text{-a.s.} \quad \forall f \in \mathcal{D}(A)_0, \quad \lim_{n \rightarrow +\infty} \frac{1}{H_n} \sum_{k=1}^n \eta_k \tilde{A}_{\gamma_k} f(\bar{X}_{\Gamma_{k-1}}) = 0. \quad (25)$$

B. We assume that (25) and $\mathcal{E}(\tilde{A}, A, \mathcal{D}(A)_0)$ (see (19)) hold. Then

$$\mathbb{P}\text{-a.s.} \quad \forall f \in \mathcal{D}(A)_0, \quad \lim_{n \rightarrow +\infty} \nu_n^\eta(Af) = 0.$$

It follows that, \mathbb{P} -a.s., every weak limiting distribution ν_∞^η of the sequence $(\nu_n^\eta)_{n \in \mathbb{N}^*}$ belongs to \mathcal{V} , the set of the invariant distributions of $(X_t)_{t \geq 0}$. Finally, if the hypothesis from Theorem 3.1 point **B.** hold and $(X_t)_{t \geq 0}$ has a unique invariant distribution, i.e. $\mathcal{V} = \{\nu\}$, then

$$\mathbb{P}\text{-a.s.} \quad \forall f \in \mathcal{C}_{\tilde{V}_{\psi, \phi, s}}(E), \quad \lim_{n \rightarrow +\infty} \nu_n^\eta(f) = \nu(f), \quad (26)$$

with $\mathcal{C}_{\tilde{V}_{\psi, \phi, s}}(E)$ defined in (17).

In the particular case where the function ψ is polynomial, (26) also reads as the a.s. convergence of the empirical measures for some L^p -Wasserstein distances, $p > 0$. From the liberty granted by the choice of ψ in this abstract framework, where only a recursive control with mean reverting is required, it is also possible to obtain convergence for functions ψ with exponential growth.

4. TOWARD A MULTILEVEL LANGEVIN MONTE CARLO

Multilevel Monte Carlo simulation methods have been introduced by M. Giles in [8] in 2008 to speed up simulations when only a biased approximations of the random variable of interest can be simulated at a reasonable computational cost. It relies on the combination of a coarse approximation corrected by several levels made of the difference of two refined levels at different scales. Refined levels need few simulated paths to produce good correctors which provide the efficiency of these methods inspired by multi-grid method developed in Numerical Analysis. Under assumptions of weak and strong quadratic convergence rates of the approximators, it is possible to optimize the *effort* (product of the variance by the complexity) on each level, including the coarse one, to produce an estimator which behaves almost like an unbiased one. Weighted versions were introduced in [14].

In [17], the weighted version of the multilevel paradigm is adapted to Langevin simulation. Langevin for invariant distributions of diffusions originally relies on the simulation of only one path of an Euler scheme with decreasing step (see *e.g.* [10]). The Langevin version of the multilevel paradigm consists in simulating paths of various lengths on each refined level to control optimally the effort. The new method is applied to various problems like double-well oscillator and sparse regression learning with a huge gain in terms performances, in fact quite comparable to what is observed with regular Monte Carlo simulation.

REFERENCES

- [1] P. ARTZNER, F. DELBAEN, J.-M. EBER, AND D. HEATH, *Coherent measures of risk*, Math. Finance, 9 (1999), pp. 203–228.
- [2] P. BRIAND, P.-E. CHAUDRU DE RAYNAL, AND C. LABART, *Particles systems and numerical schemes for mean reflected stochastic differential equations*, arXiv:1612.06886, (2016).
- [3] P. BRIAND, R. ÉLIE, AND Y. HU, *BSDEs with mean reflexion*, Ann. Appl. Probab. (to appear), (2016).
- [4] R. DOUC, G. FORT, E. MOULINES, AND P. SOULIER, *Practical drift conditions for subgeometric rates of convergence*, Ann. Appl. Probab., 14 (2004), pp. 1353–1377.
- [5] S. N. ETHIER AND T. G. KURTZ, *Markov processes*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York, 1986. Characterization and convergence.
- [6] W. FELLER, *The parabolic differential equations and the associated semi-groups of transformations*, Annals of Mathematics, 55 (1952), pp. 468–519.
- [7] N. FOURNIER AND A. GUILLIN, *On the rate of convergence in Wasserstein distance of the empirical measure*, Probab. Theory Related Fields, 162 (2015), pp. 707–738.
- [8] M. B. GILES, *Multilevel Monte Carlo path simulation*, Oper. Res., 56 (2008), pp. 607–617.
- [9] R. HAS’MINSKII, *Stochastic stability of differential equations*, vol. 7 of Monographs and Textbooks on Mechanics of Solids and Fluids : Mechanics and Analysis, Sijthoff & Noordhoff, Alphen aan den Rijn, 1980.
- [10] D. LAMBERTON AND G. PAGÈS, *Recursive computation of the invariant distribution of a diffusion*, Bernoulli, 8 (2002), pp. 367–405.
- [11] ———, *Recursive computation of the invariant distribution of a diffusion: The case of a weakly mean reverting drift*, Stochastics and Dynamics, 03 (2003), pp. 435–451.
- [12] V. LEMAIRE, *Estimation récursive de la mesure invariante d’un processus de diffusion*, PhD thesis, 2005. Thèse de doctorat dirigée par D. Lamberton et G. Pagès, Mathématiques appliquées, Université de Marne-la-Vallée.
- [13] V. LEMAIRE, *An adaptive scheme for the approximation of dissipative systems*, Stochastic Processes and their Applications, 117 (2007), pp. 1491 – 1518.
- [14] V. LEMAIRE AND G. PAGÈS, *Multilevel Richardson-Romberg extrapolation*, Bernoulli, 23 (2017), pp. 2643–2692.
- [15] P.-L. LIONS AND A.-S. SZNITMAN, *Stochastic differential equations with reflecting boundary conditions*, Comm. Pure Appl. Math., 37 (1984), pp. 511–537.
- [16] G. PAGÈS, *Sur quelques algorithmes récursifs pour les probabilités numériques*, ESAIM Probab. Statist., 5 (2001), pp. 141–170 (electronic).
- [17] G. PAGÈS AND F. PANLOUP, *Weighted multilevel langevin simulation of invariant measures*, Ann. Appl. Probab. (to appear), (2016).
- [18] G. PAGÈS AND C. REY, *Recursive computation of the invariant distribution of Feller Markov processes*, ArXiv e-prints, 1703.04557v4, (2017).
- [19] F. PANLOUP, *Recursive computation of the invariant measure of a stochastic differential equation driven by a Lévy process*, Ann. Appl. Probab., 18 (2008), pp. 379–426.
- [20] A. PAZY, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Applied Mathematical Sciences, Springer New York, 1992.
- [21] S. T. RACHEV AND L. RÜSCHENDORF, *Mass transportation problems. Vol. I*, Probability and its Applications (New York), Springer-Verlag, New York, 1998. Theory.

- [22] A.-S. SZNITMAN, *Topics in propagation of chaos*, in *École d'Été de Probabilités de Saint-Flour XIX—1989*, vol. 1464 of *Lecture Notes in Math.*, Springer, Berlin, 1991, pp. 165–251.
- [23] H. TANAKA, *Stochastic differential equations with reflecting boundary condition in convex regions*, *Hiroshima Math. J.*, 9 (1979), pp. 163–177.