ESAIM: M2AN 59 (2025) 1333–1363
https://doi.org/10.1051/m2an/2024086

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

CONVERGENCE OF PARTICLES AND TREE BASED SCHEME FOR SINGULAR FBSDES

Jean-François Chassagneux[1,*] and Mohan Yang[2]

Abstract. We study an implementation of the theoretical splitting scheme introduced in Chassagneux and Yang (*J. Comput. Phys.* **468** (2022) 111459) for singular FBSDEs (Carmona and Delarue, *Probab. Theory Relat. Fields* **157** (2013) 333–388) and their associated quasi-linear degenerate PDEs. The fully implementable algorithm is based on particles approximation of the transport operator and tree like approximation of the diffusion operator appearing in the theoretical splitting. We prove the convergence with a rate of our numerical method under some reasonable conditions on the coefficients functions. This validates *a posteriori* some numerical results obtained in Chassagneux and Yang (*J. Comput. Phys.* **468** (2022) 111459). We conclude the paper with a numerical section presenting various implementations of the algorithm and discussing their efficiency in practice.

Mathematics Subject Classification. 65C30, 65C20, 65M12, 60H35.

Received March 6, 2023. Accepted December 30, 2024.

1. Introduction

In this work, we study the theoretical and numerical convergence of a particles and tree based scheme for singular FBSDEs. The class of singular Forward Backward Stochastic Differential Equations (FBSDE) is motivated by application in the modeling of carbon markets [4, 8, 10]. In these models, the FBSDE solution represents the equilibrium price of an allowance (to emit one ton of CO_2) in a perfect market. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a stochastic basis supporting a d-dimensional Brownian motion W and $T > 0$ a terminal time. We denote by $\mathbb{F} := (\mathcal{F}_t)_{t\geqslant 0}$ the filtration generated by the Brownian motion (augmented). The singular FBSDE system, with solution $(P_t, E_t, Y_t, Z_t)_{0\leqslant t\leqslant T}$, has the following form:

$$\begin{cases} \mathrm{d}P_t = b(P_t)\,\mathrm{d}t + \sigma(P_t)\,\mathrm{d}W_t \\ \mathrm{d}E_t = \mu(P_t, Y_t)\,\mathrm{d}t \\ \mathrm{d}Y_t = Z_t\,\mathrm{d}W_t \end{cases}. \tag{1.1}$$

The terminal condition is typically given by

$$Y_T = \phi(E_T) \text{ where } \phi(e) = \mathbf{1}_{e\geqslant\Lambda}, \tag{1.2}$$

Keywords and phrases. Singular FBSDEs, probabilistic methods, splitting scheme, particle method.

[1] ENSAE-CREST and Institut Polytechnique de Paris, Palaiseau, France.

[2] UFR de Mathématiques & LPSM, Université Paris Cité, Bâtiment Sophie Germain, 8 place Aurélie Nemours, 75013 Paris, France.

*Corresponding author: chassagneux@ensae.fr

for some $\Lambda > 0$. Existence and uniqueness to the above FBSDE is not straightforward has it is fully coupled (the backward process Y appears in the coefficient of the forward process E), it is degenerate in the forward direction E and the terminal condition is discontinuous. However, assuming some structural conditions, see Theorem 2.1 below for a precise statement, Carmona and Delarue [3] managed to prove the wellposedness of such singular FBSDEs. They also show that the following Markovian representation holds true for the Y-process, namely:

$$Y_t := \mathcal{V}(t, P_t, E_t), \quad 0 \leqslant t < T, \tag{1.3}$$

where the measurable function $\mathcal{V} : [0, T] \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$ is classically named the *decoupling field*.

The decoupling field $\mathcal{V}$ associated to (1.1) is a weak solution to

$$\partial_t u + \partial_e(\mathfrak{M}(p, u)) + \partial_p u b(p) + \frac{1}{2}\mathrm{Tr}\big[\sigma(p)\sigma^\top(p)\partial^2_{pp}u\big] = 0 \ \text{ and } \ u(T, p, e) = \phi(e), \tag{1.4}$$

with

$$\mathfrak{M}(p, y) := \int_0^y \mu(p, \upsilon)\, \mathrm{d}\upsilon, \quad 0 \leqslant y \leqslant 1, \tag{1.5}$$

where ∂_p denotes the Jacobian with respect to p, ∂_t the time derivative, ∂_e the derivative with respect to the e variable, $\top$ is the transpose and ∂^2_{pp} is the matrix of second derivatives with respect to the p variable.

The algorithm we study is based on a splitting scheme designed for this kind of equation and introduced in [6]. Though it could be defined at a general level for the PDE, we should note that the convergence proof relies on the FBSDE setting. Indeed, under appropriate conditions, the FBSDE appears as well-posed random characteristics for the PDE (1.4). The splitting scheme involves the iteration on a discrete time grid of the composition of two operators: a transport operator $\mathcal{T}$ (in this case the P variable is fixed) and a diffusion operator $\mathcal{D}$ (in this case the E variable is fixed). Precise definitions are given in the next section. From [6], we already know that the theoretical splitting scheme is convergent with a rate one half. This result is obtained under the same structural conditions which guarantee the well-posedness of the FBSDE. In [6], the splitting scheme is then implemented using various finite difference approximations of the transport operator and a non-linear regression using Deep Neural Networks for the diffusion operator. An alternative scheme ([6], Sect. 3.3.1) is also suggested to verify the convergence result of the main numerical procedure. The goal of this paper is to analyse precisely this alternative scheme by proving its convergence with a rate, under some regularity conditions and assuming a constant volatility coefficient in the P dynamics.

The rest of the paper is organized as follows. In the next section, we recall the theoretical setting for the study of singular FBSDEs and recall the theoretical splitting scheme. We also introduce the framework used to study the convergence of the numerical methods and give some properties of the solution $\mathcal{V}$ in this framework. In Section 3, we present the numerical algorithm and some of its variants that will be used for numerical tests. We also state our main convergence result in Theorem 3.1. Section 4 is dedicated to the proof of the convergence by studying precisely all sources of errors. The last section presents numerical results for the various schemes introduced, illustrating the theoretical convergence.

Notations. In the following we will use the following spaces:

Let $\mathbb{H} = (\mathcal{H}_t)_{t\geqslant 0}$ denotes a generic right-continuous filtration,

- For fixed $0 \leqslant a < b < +\infty$ and $I = [a, b]$ or $I = [a, b)$, $\mathcal{S}^{2,k}(I, \mathbb{H})$ is the set of $\mathbb{R}^k$-valued càdlàg[1] $\mathcal{H}_t$-adapted processes Y, s.t.

$$\|Y\|^2_{\mathcal{S}^2} := \mathbb{E}\left[\sup_{t\in I} |Y_t|^2\right] < \infty.$$

  Note that we may omit the dimension and the terminal date in the norm notation as this will be clear from the context. $\mathcal{S}^{2,k}_{\mathrm{c}}(I, \mathbb{H})$ is the subspace of processes with continuous sample paths.

[1]French acronym for right continuous with left limits.

– For fixed $0 \leqslant a < b < +\infty$, and $I = [a,b]$, we denote by $\mathcal{H}^{2,k}(I,\mathbb{H})$ the set of $\mathbb{R}^k$-valued progressively measurable processes Z, such that

$$\|Z\|_{\mathcal{H}^2}^2 := \mathbb{E}\left[\int_I |Z_t|^2 \,\mathrm{d}t\right] < \infty.$$

We denote by $\mathscr{P}(\mathbb{R})$ the space of probability measures on $\mathbb{R}$ and $\mathscr{P}_q(\mathbb{R})$, $q \geqslant 1$ the subset of probability measure which have a q-moment finite.

We denote by $\mathscr{I}$ the space of cumulative distribution function (CDF) namely functions θ given by

$$\mathbb{R} \ni x \mapsto \theta(x) = \mu((-\infty, x]) \in [0,1]$$

for some $\mu \in \mathscr{P}(\mathbb{R})$.

The set $\mathscr{C}^{1,2,1}(I \times \mathbb{R}^d \times \mathbb{R})$ (resp. $\mathscr{C}^{1,2,2}(I \times \mathbb{R}^d \times \mathbb{R})$) with $I = [0,T)$ or $I = [0,T]$, is the set of functions $f : I \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$ continuously differentiable: once in their first variable, twice in their second variable and once (resp. twice) in their third variable.

2. Review of theoretical result for singular FBSDEs

In this section, we first recall the main properties of solution to (1.1). We then define the theoretical splitting scheme studied in [6] and used to design our implemented algorithms introduced in the next section. Finally, we consider a smooth framework for the solution of (1.4) and obtain some properties useful to prove the convergence of the numerical schemes.

2.1. Well-posedness of Singular FBSDEs

We first define a class of functions to which the solution belongs.

Definition 2.1. Let $\mathcal{K}$ be the class of functions $\Phi : \mathbb{R}^d \times \mathbb{R} \to [0,1]$ such that Φ is L_ϕ-Lipschitz in the first variable for some $L_\Phi > 0$, namely

$$\left|\Phi(p,e) - \Phi(p',e)\right| \leqslant L_\Phi |p - p'| \quad \text{for all} \quad (p,p',e) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}, \tag{2.1}$$

and moreover satisfying, for all $p \in \mathbb{R}^d$,

$$\Phi(p,e) = \nu(p,(-\infty,e]) \text{ where } \nu(p,\cdot) \in \mathcal{P}(\mathbb{R}). \tag{2.2}$$

From the above definition, we have that $\Phi(p,\cdot) \in \mathscr{I}$ and is thus a non-decreasing, right continuous function and satisfies, for all $p \in \mathbb{R}^d$,

$$\lim_{e \to -\infty} \Phi(p,e) = 0 \;\text{ and }\; \lim_{e \to -\infty} \Phi(p,e) = 1. \tag{2.3}$$

We now recall the existence and uniqueness result for the singular FBSDE. It is also key to define the theoretical splitting scheme. This result is obtained for the following class of admissible coefficient functions.

Definition 2.2. Let $\mathcal{A}$ be the class of functions $B : \mathbb{R}^d \to \mathbb{R}^d$, $\Sigma : \mathbb{R}^d \to \mathcal{M}_d$, $F : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$ which are L-Lipschitz continuous functions. Moreover, F is strictly decreasing in y and satisfies, for all $p \in \mathbb{R}^d$,

$$\ell_1 |y - y'|^2 \leqslant (y - y')(F(p,y') - F(p,y)) \leqslant \ell_2 |y - y'|^2, \tag{2.4}$$

where L, ℓ_1 and ℓ_2 are positive constants.

The well-posedness result in this setting reads as follows.

Theorem 2.1 ([3], Prop. 2.10, [8], Prop. 3.2)**.** *Let $\tau > 0$, $(B, \Sigma, F) \in \mathcal{A}$ and $\Phi \in \mathcal{K}$.*

Given any initial condition $(t_0, p, e) \in [0, \tau) \times \mathbb{R}^d \times \mathbb{R}$, there exists a unique progressively measurable 4-tuple of processes $(P_t^{t_0,p,e}, E_t^{t_0,p,e}, Y_t^{t_0,p,e}, Z_t^{t_0,p,e})_{t_0 \leqslant t \leqslant \tau} \in \mathcal{S}_c^{2,d}([t_0,\tau],\mathbb{F}) \times \mathcal{S}_c^{2,1}([t_0,\tau],\mathbb{F}) \times \mathcal{S}_c^{2,1}([t_0,\tau),\mathbb{F}) \times \mathcal{H}^{2,d}([t_0,\tau],\mathbb{F})$ satisfying the dynamics

$$
\begin{aligned}
\mathrm{d}P_t^{t_0,p,e} &= B\big(P_t^{t_0,p,e}\big)\,\mathrm{d}t + \Sigma\big(P_t^{t_0,p,e}\big)\mathrm{d}W_t, && P_{t_0}^{t_0,p,e} = p \in \mathbb{R}^d,\\
\mathrm{d}E_t^{t_0,p,e} &= F\big(P_t^{t_0,p,e}, Y_t^{t_0,p,e}\big)\,\mathrm{d}t, && E_{t_0}^{t_0,p,e} = e \in \mathbb{R},\\
\mathrm{d}Y_t^{t_0,p,e} &= Z_t^{t_0,p,e}\,\mathrm{d}W_t,
\end{aligned}
\tag{2.5}
$$

and such that

$$
\mathbb{P}\left[\Phi_-\big(P_\tau^{t_0,p,e}, E_\tau^{t_0,p,e}\big) \leqslant \lim_{t\uparrow\tau} Y_t^{t_0,p,e} \leqslant \Phi_+\big(P_\tau^{t_0,p,e}, E_\tau^{t_0,p,e}\big)\right] = 1. \tag{2.6}
$$

The unique decoupling field *defined by*

$$
[0,\tau) \times \mathbb{R}^d \times \mathbb{R} \ni (t_0, p, e) \to w(t_0, p, e) = Y_{t_0}^{t_0,p,e} \in \mathbb{R}
$$

is continuous and satisfies

(1) *For any $t \in [0,\tau)$, the function $w(t,\cdot,\cdot)$ is $1/(l_1(\tau - t))$-Lipschitz continuous with respect to e.*
(2) *For any $t \in [0,\tau)$, the function $w(t,\cdot,\cdot)$ is C-Lipschitz continuous with respect to p, where C is a constant depending on L, τ and L_ϕ only.*
(3) *Given $(p,e) \in \mathbb{R}^d \times \mathbb{R}$, for any family $(p_t, e_t)_{0 \leqslant t < \tau}$ converging to (p,e) as $t \uparrow \tau$, we have*

$$
\Phi_-(p,e) \leqslant \liminf_{t\to\tau} w(t, p_t, e_t) \leqslant \limsup_{t\to\tau} w(t, p_t, e_t) \leqslant \Phi_+(p,e). \tag{2.7}
$$

(4) *For any $t \in [0,\tau)$, the function $w(t,\cdot,\cdot) \in \mathcal{K}$.*

The previous result leads us to define a non-linear operator associated to (1.1) under the following assumption, that will hold true for the rest of the paper:

Standing assumptions. The coefficients (b, σ, μ) of the singular FBSDE (1.1) belong to $\mathcal{A}$.

Definition 2.3. We define the semi-group of operator $(\Theta_h)_{h>0}$ by

$$
(0,\infty) \times \mathcal{K} \ni (h, \psi) \mapsto \Theta_h(\psi) = v(0,\cdot) \in \mathcal{K} \tag{2.8}
$$

where v is the decoupling field given in Theorem 2.1 with parameters $\tau = h$, $B = b$, $\Sigma = \sigma$, $F = \mu$ and $\Phi = \psi$.

Let us now be given a discrete time grid of $[0,T]$:

$$
\pi := \{t_0 := 0 \leqslant \cdots \leqslant t_n \leqslant \ldots t_N := T\}, \tag{2.9}
$$

for $N \geqslant 1$. For ease of presentation, we assume that the time grid π is equidistant and thus, for $0 \leqslant n \leqslant N-1$,

$$
t_{n+1} - t_n = \frac{T}{N} =: \mathfrak{h}. \tag{2.10}
$$

Using Definition 2.3, we observe that

$$
\mathcal{V}(0,\cdot) := \Theta_T(\phi) = \Theta_{t_1 - t_0} \circ \cdots \circ \Theta_{t_N - t_{N-1}}(\phi). \tag{2.11}
$$

We recall now a result that arises from the proof of the previous Theorem 2.1, see [8].

Corollary 2.1 (Approximation result)**.** *Let $\tau > 0$, $(B, \Sigma, F) \in \mathcal{A}$ and $\Phi \in \mathcal{K}$. Let $(\Phi^k)_{k\geqslant 0}$ be a sequence of smooth functions belonging to $\mathcal{K}$ and converging pointwise towards Φ as k goes to $+\infty$. For $\eta > 0$, consider then $w^{\eta,k}$ the solution to:*

$$\partial_t u + F(p, u)\partial_e u + \mathcal{L}_p u + \frac{1}{2}\eta^2\big(\partial^2_{ee} u + \Delta_{pp} u\big) = 0 \text{ and } u(\tau, \cdot) = \Phi^k \tag{2.12}$$

where Δ_{pp} is the Laplacian with respect to p, and $\mathcal{L}_p$ is the operator

$$\mathcal{L}_p(\varphi)(t, p, e) = \partial_p\varphi(t, p, e)B(p) + \frac{1}{2}\mathrm{Tr}\big[A(p)\partial^2_{pp}\big](\varphi)(t, p, e), \tag{2.13}$$

with $A = \Sigma\Sigma^\top$. For later use, we also define $\mathcal{L}^\eta := \mathcal{L}_p + \frac{1}{2}\eta^2(\partial^2_{ee} + \Delta_{pp})$. Then the functions $w^{\eta,k}$ are $\mathscr{C}^{1,2,2}([0, \tau] \times \mathbb{R}^d \times \mathbb{R})$ and $\lim_{k\to\infty} \lim_{\eta\to 0} w^{\eta,k} = w$ where the convergence is locally uniform in $[0, \tau) \times \mathbb{R}^d \times \mathbb{R}$. Moreover, for all k, η, $w^{\eta,k}(t, \cdot) \in \mathcal{K}$.

For later use, we introduce the system of FBSDEs associated to (2.12). We thus consider $(W, \tilde{B}, \tilde{\tilde{B}})$ to be independent Brownian motions with dimension respectively d, d and one, and denote $\mathbb{G} = (\mathcal{G}_t)_{t\geqslant 0}$ the filtration they generate (augmented). For $t_0 \in [0, T)$ and $(p, e) \in \mathbb{R}^d \times \mathbb{R}$, let

$$\big(P^\eta, E^{\eta,k}\big) \in \mathcal{S}^{2,d}_{\mathrm{c}}([t_0, T], \mathbb{G}) \times \mathcal{S}^{2,1}_{\mathrm{c}}([t_0, T], \mathbb{G})$$

and

$$\Big(Y^{\eta,k}, Z^{\eta,k}, \tilde{Z}^{\eta,k}, \tilde{\tilde{Z}}^{\eta,k}\Big) \in \mathcal{S}^{2,d}_{\mathrm{c}}([t_0, T], \mathbb{G}) \times \mathcal{H}^{2,d}([t_0, T], \mathbb{G}) \times \mathcal{H}^{2,d}([t_0, T], \mathbb{G}) \times \mathcal{H}^{2,1}([t_0, T], \mathbb{G})$$

be the solution to

$$\begin{cases} \mathrm{d}P^\eta_t = B(P^\eta_t)\,\mathrm{d}t + \Sigma(P^\eta_t)\,\mathrm{d}W_t + \eta\,\mathrm{d}\tilde{B}_t,\ P^\eta_{t_0} = p, \\ \mathrm{d}E^{\eta,k}_t = F\Big(P^\eta_t, Y^{\eta,k}_t\Big)\,\mathrm{d}t + \eta\,\mathrm{d}\tilde{\tilde{B}}_t,\ E^\eta_{t_0} = e, \\ \mathrm{d}Y^{\eta,k} = Z^{\eta,k}\,\mathrm{d}W_t + \tilde{Z}^{\eta,k}\,\mathrm{d}\tilde{B}_t + \tilde{\tilde{Z}}^{\eta,k}\tilde{\tilde{B}}_t. \end{cases} \tag{2.14}$$

Observe that $Y^{\eta,k}_t = w^{\eta,k}(t, P^\eta_t, E^{\eta,k}_t)$ for $t \in [t_0, T]$, in particular $Y^{\eta,k}_T = \Phi^k(P_T, E_t)$. Let us note that the above processes depend upon the initial condition (t_0, p, e) but we slightly abuse the notation in this regard by omitting to indicate the initial condition.

2.2. A theoretical splitting scheme

The numerical algorithm is based on a splitting method for the quasilinear PDE (1.4), which has been introduced in [6]. We now recall the splitting approach at a theoretical level.

We first introduce the transport step where the diffusion part is frozen.

Definition 2.4 (Transport step)**.** We set

$$(0, \infty) \times \mathcal{K} \ni (h, \psi) \mapsto \mathcal{T}_h(\psi) = \tilde{v}(0, \cdot) \in \mathcal{K},$$

where $\tilde{v}$ is the decoupling field defined in Theorem 2.1 with parameters $\tau = h$, $B = 0$, $\Sigma = 0$, $F = \mu$ and terminal condition $\Phi = \psi$.

In the definition above, $\tilde{v}(\cdot)$, defined on $[0, h] \times \mathbb{R}^d \times \mathbb{R}$, is the unique entropy solution to

$$\partial_t w + \partial_e(\mathfrak{M}(p, w)) = 0 \text{ and } \tilde{v}(h, \cdot) = \psi, \tag{2.15}$$

recall (1.4).

Remark 2.1. For later use in the convergence analysis, we introduce a small modification of the previous operator that acts on function in $\mathscr{I}$ instead of $\mathcal{K}$. Namely, we set

$$(0,\infty)\times\mathbb{R}^d\times\mathscr{I}\ni(h,p,\theta)\mapsto\tilde{\mathcal{T}}_h(p,\theta)=\hat{v}_p(0,\cdot)\in\mathscr{I}$$

where $\hat{v}_p(0,\cdot)$ is the unique entropy solution to

$$\partial_t w+\partial_e(\mathfrak{M}(p,w))=0\ \text{ and }\ \hat{v}_p(h,\cdot)=\theta. \tag{2.16}$$

One observes that $\mathcal{T}_h(\psi)(p,\cdot)=\tilde{\mathcal{T}}_h(p,\psi(p,\cdot))$.

Let us mention that, from *e.g.* Theorem 1.1 of [11], we have the following stability result:

$$\int_{\mathbb{R}}\left|\tilde{\mathcal{T}}_h(p,\psi_1(p,\cdot))(e)-\tilde{\mathcal{T}}_h(p,\psi_2(p,\cdot))(e)\right|\mathrm{d}e\leqslant\int_{\mathbb{R}}|\psi_1(p,e)-\psi_2(p,e)|\,\mathrm{d}e, \tag{2.17}$$

for $\psi_1,\psi_2\in\mathcal{K}$.

We now introduce the diffusion step, where conversely, the $E-$process is frozen to its initial value.

Definition 2.5 (Diffusion step)**.** We set

$$(0,\infty)\times\mathcal{K}\ni(h,\psi)\mapsto\mathcal{D}_h(\psi)=\bar{v}(0,\cdot)\in\mathcal{K}$$

where $\bar{v}(0,\cdot)$ is the decoupling field in Theorem 2.1 with parameters $\tau=h$, $B=b$, $\Sigma=\sigma$, $F=0$ and terminal condition $\Phi=\psi$.

Observe that, for $t\in[0,h)$,

$$\bar{v}(t,p,e)=\mathbb{E}\big[\psi\big(P_h^{t,p},e\big)\big],\ (p,e)\in\mathbb{R}^d\times\mathbb{R}\ \text{ and }\bar{v}(t,\cdot)\in\mathcal{K}. \tag{2.18}$$

We can now define the theoretical scheme on π by a backward induction. Since

$$\mathcal{V}(0,\cdot)=\Theta_T(\phi)=\prod_{0\leqslant n<N}\Theta_{t_{n+1}-t_n}(\phi), \tag{2.19}$$

the main point of the scheme is to replace one step of Θ by one step of $\mathcal{T}\circ\mathcal{D}$. This leads to the following.

Definition 2.6 (Theoretical splitting scheme)**.** We set

$$(0,\infty)\times\mathcal{K}\ni(h,\psi)\mapsto\mathcal{S}_h(\psi):=\mathcal{T}_h\circ\mathcal{D}_h(\psi)\in\mathcal{K}.$$

For $n\leqslant N$, we denote by u_n^π the solution of the following backward induction on π:
- for $n=N$, set $u_N^\pi:=\phi$,
- for $n<N$, $u_n^\pi=\mathcal{S}_{t_{n+1}-t_n}(u_{n+1}^\pi)$.

The $(u_n^\pi)_{0\leqslant n\leqslant N}$ stands for the approximation of the decoupling field $\mathcal{V}(t,\cdot)$ for $t\in\pi$.

It is proven in Proposition 2.1 of [6] a key result concerning the scheme's truncation error.

Proposition 2.1 (truncation error)**.** *Under our standing assumptions on the coefficients* (μ,b,σ)*, the following holds, for* $\psi\in\mathcal{K}$*:*

$$\int_{\mathbb{R}}|\mathcal{S}_h(\psi)(p,e)-\Theta_h(\psi)(p,e)|\,\mathrm{d}e\leqslant C_{L_\psi}\big(1+|p|^2\big)h^{\frac{3}{2}}, \tag{2.20}$$

for $p\in\mathbb{R}^d$*,* $h>0$*.*

This allows in particular to obtain the convergence with a rate $\frac{1}{2}$ of the theoretical splitting scheme to the decoupling field $\mathcal{V}$, namely

$$\int_{\mathbb{R}}|\mathcal{V}(0,p,e)-u_0^\pi(p,e)|\,\mathrm{d}e\leqslant C\big(1+|p|^2\big)\sqrt{\mathfrak{h}},\qquad p\in\mathbb{R}^d,$$

for a positive constant C, see Theorem 2.2 of [6].

2.3. Smooth setting for convergence results

We now consider a more restricted framework. In this setting, we give further properties of the value function $\mathcal{V}$ that will be used to prove our convergence results. Indeed, even if the scheme can be defined in a quite general setting, the convergence with a rate is obtained in a more regular framework.

Therefore, the main assumption we shall use is the following, see Remark 2.2 below.

Assumption 2.1. *The* decoupling field $\mathcal{V}$ *is a* $\mathscr{C}^{1,2,1}([0,T)\times\mathbb{R}^d\times\mathbb{R})$ *solution to*

$$\partial_t\mathcal{V}+\mu(p,\mathcal{V})\partial_e\mathcal{V}+b(p)\partial_p\mathcal{V}+\frac{1}{2}\bar{\sigma}^2\Delta_p\mathcal{V}=0. \tag{2.21}$$

In particular, the function $\sigma(\cdot)$ *is constant* $\sigma(\cdot):=\bar{\sigma}I_d$ *with* $\bar{\sigma}>0$ *and* I_d *the* $d\times d$ *identity matrix, recall* (1.4). *Moreover, the function* b *is twice differentiable with bounded and Lipschitz derivatives and the function* μ *is differentiable with Lipschitz and bounded derivatives.*

We first list some further properties of the function $\mathcal{V}$ in the regular setting of the previous assumption.

Proposition 2.2. *Under Assumption 2.1, the function* $\mathcal{V}$ *satisfies, for* $(t,p,e)\in[0,T)\times\mathbb{R}^d\times\mathbb{R}$,

$$0\leqslant\partial_e\mathcal{V}(t,p,e)\leqslant\frac{C}{T-t}\quad\text{and}\quad|\partial_p\mathcal{V}(t,p,e)|\leqslant C. \tag{2.22}$$

Moreover, for $(t,p)\in[0,T)\times\mathbb{R}^d$,

$$\int_{\mathbb{R}}|\partial_p\mathcal{V}(t,p,e)|\,\mathrm{d}e\leqslant C(T-t), \tag{2.23}$$

and, thus, for $p'\in\mathbb{R}^d$,

$$\int_{\mathbb{R}}\big|\mathcal{V}(t,p,e)-\mathcal{V}(t,p',e)\big|\,\mathrm{d}e\leqslant C(T-t)|p-p'|. \tag{2.24}$$

Proof. The estimates (2.22) are obtained directly from the property of $\mathcal{V}$ given in Theorem 2.1. We now study the L^1-control.

Let $\mathcal{V}^{\eta,k}$ be the smooth approximation given in Corollary 2.1 to $\mathcal{V}$. We consider also the associated FBSDE $(P^\eta,E^{\eta,k},Y^{\eta,k})$ in (2.14). We adapt the computations in the proof of Lemma 3.8 in [8], to obtain the equation corresponding to equation (3.35) there. In our simpler framework ($r=0$, ϕ does not depend on p), it reads

$$\partial_p\mathcal{V}^{\eta,k}(t_0,p,e)=\mathbb{E}\left[\int_{t_0}^T\partial_p\mu\Big(P_t^\eta,Y_t^{\eta,k}\Big)\partial_pP_t^\eta\partial_eE_t^{\eta,k}\partial_e\mathcal{V}^{\eta,k}\Big(t,P_t^\eta,E_t^{\eta,k}\Big)\,\mathrm{d}t\right] \tag{2.25}$$

where $(\partial_pP^\eta,\partial_eE^{\eta,k})$ are the tangent processes associated to $(P^\eta,E^{\eta,k})$. They are given by

$$\partial_eE_t^{\eta,k}=1+\int_{t_0}^t\partial_y\mu\big(P_s^\eta,\mathcal{V}^{\eta,k}\big(s,P_s^\eta,E_s^{\eta,k}\big)\big)\partial_e\mathcal{V}^{\eta,k}\big(s,P_s^\eta,E_s^{\eta,k}\big)\partial_eE_s^{\eta,k}\,\mathrm{d}s, \tag{2.26}$$

$$\Big(\partial_{p^{\ell'}}P_t^\eta\Big)^\ell=\mathbf{1}_{\{\ell=\ell'\}}+\int_{t_0}^t\partial_pb^\ell(P_s^\eta)\partial_{p^{\ell'}}P_s^\eta\,\mathrm{d}s\quad\text{for}\quad\ell,\ell'\in\{1,\dots,d\}. \tag{2.27}$$

Observe, in particular, that

$$\partial_eE_t^{\eta,k}=\exp\bigg(\int_{t_0}^t\partial_y\mu\big(P_s^\eta,\mathcal{V}^{\eta,k}\big(s,P_s^\eta,E_s^{\eta,k}\big)\big)\partial_e\mathcal{V}^{\eta,k}\big(s,P_s^\eta,E_s^{\eta,k}\big)\,\mathrm{d}s\bigg),\ t\in[t_0,T].$$

Moreover, it is well known in our Lipschitz setting that

$$\mathbb{E}\left[\sup_{t\in[t_0,T]}|\partial_p P_t^\eta|^\kappa\right] \leqslant C_\kappa, \tag{2.28}$$

for $\kappa \geqslant 1$.

For a positive real number R, we obtain from (2.25) that

$$\int_{-R}^{R}\left|\partial_p \mathcal{V}^{\eta,k}(t_0,p,e)\right| \mathrm{d}e \leqslant \mathbb{E}\left[\int_{t_0}^{T}\int_{-R}^{R}\left|\partial_p\mu\left(P_t^\eta,Y_t^{\eta,k}\right)\partial_p P_t^\eta\right|\partial_e\left[\mathcal{V}^{\eta,k}\left(t,P_t^\eta,E_t^{\eta,k}\right)\right]\mathrm{d}e\,\mathrm{d}t\right]$$

recalling that $\partial_e E^{\eta,k}$ and $\partial_e\mathcal{V}^{\eta,k}(\cdot)$ are non-negative. Moreover, since $\mathcal{V}^{\eta,k}$ is bounded and μ Lipschitz continuous, we compute

$$\int_{-R}^{R}\left|\partial_p \mathcal{V}^{\eta,k}(t_0,p,e)\right| \mathrm{d}e \leqslant 2\left\|\mathcal{V}^{\eta,k}\right\|_\infty\|\partial_p\mu\|_\infty(T-t_0)\mathbb{E}\left[\sup_{t\in[t_0,T]}|\partial_p P_t^\eta|\right].$$

Using (2.28), we eventually obtain

$$\int_{-R}^{R}\left|\partial_p \mathcal{V}^{\eta,k}(t_0,p,e)\right| \mathrm{d}e \leqslant C(T-t_0). \tag{2.29}$$

Since $\mathcal{V}^{\eta,k}(t_0,p+\delta,e)-\mathcal{V}^{\eta,k}(t_0,p,e)=\delta\int_0^1\partial_p\mathcal{V}^{\eta,k}(t_0,p+\lambda\delta,e)\,\mathrm{d}\lambda$, we have

$$\int_{-R}^{R}\left|\mathcal{V}^{\eta,k}(t_0,p+\delta,e)-\mathcal{V}^{\eta,k}(t_0,p,e)\right|\mathrm{d}e \leqslant \delta\int_0^1\int_{-R}^{R}\left|\partial_p\mathcal{V}^{\eta,k}(t_0,p+\lambda\delta,e)\right|\mathrm{d}e\lambda, \tag{2.30}$$

$$\leqslant C(T-t_0)\delta \tag{2.31}$$

where we used (2.29) for the last inequality. The dominated convergence theorem leads to

$$\int_{-R}^{R}\frac{1}{\delta}|\mathcal{V}(t_0,p+\delta,e)-\mathcal{V}(t_0,p,e)|\,\mathrm{d}e \leqslant C(T-t_0), \tag{2.32}$$

recalling Corollary 2.1. Taking limit as $\delta\to 0$, we obtain

$$\int_{-R}^{R}|\partial_p\mathcal{V}(t_0,p,e)|\,\mathrm{d}e \leqslant C(T-t_0). \tag{2.33}$$

The estimate (2.23) is then obtained by letting $R\to+\infty$ and invoking monotone convergence. □

The study of the regularity of the decoupling field, in particular how to prove Assumption 2.1, is outside the scope of this paper and is left for further research. We, however, give now an example of model for which Assumption 2.1 is known to hold.

Remark 2.2. Let us consider the following *toy model*, which has been studied in [4]

$$\begin{cases}\mathrm{d}P_t = \bar{\sigma}\,\mathrm{d}\mathcal{W}_t,\ P_0=p,\\ \mathrm{d}E_t=(P_t-Y_t)\,\mathrm{d}t,\ E_0=e,\\ \mathrm{d}Y_t=Z_t\,\mathrm{d}\mathcal{W}_t.\end{cases} \tag{2.34}$$

Here, $\bar{\sigma}$ is a positive constant, $\mathcal{W}$ is a one-dimensional Brownian Motion and the terminal condition (in the weak sense of (2.6)) for Y is given by $\phi(E_T)$, recall (1.2). By Theorem 2.1 above, it has a unique solution and we denote by v its decoupling field. In particular, we have $v(0,p,e) = Y_0$. Now, we introduce, as in [4], the following process

$$\bar{E}_t = E_t + (T-t)P_t, \tag{2.35}$$

and observe then that $(\bar{E}, Y)$ satisfies

$$\begin{cases} \mathrm{d}\bar{E}_t = -Y_t\,\mathrm{d}t + (T-t)\,\mathrm{d}\mathcal{W}_t \\ \mathrm{d}Y_t = Z_t\,\mathrm{d}\mathcal{W}_t \end{cases} \tag{2.36}$$

with the same terminal condition since $\bar{E}_T = E_T$. Independently from (2.34), the system (2.36) is thoroughly studied in Section 4 of [4], this system has also a unique solution with a decoupling field $Y_0 = \bar{v}(0,\bar{E}_0)$, and from Proposition 6 in [4] again, $\bar{v}$ is $\mathscr{C}^{1,2}([0,T)\times\mathbb{R})$. We deduce that $v(0,p,e) = \bar{v}(0,e+Tp)$ and more generally $v(t,p,e) = \bar{v}(t,e+(T-t)p)$ for $t \in [0,T)$. This yields that v is $\mathscr{C}^{1,2,2}([0,T)\times\mathbb{R}^d\times\mathbb{R})$.

Let us observe that Proposition 2.2 states estimates for the first order derivatives of $\mathcal{V}$. And, though $\mathcal{V}$ is assumed to be smooth in Assumption 2.1, nothing is said on the control of higher order derivatives. As we will see later on, the control of the second order derivative and third order derivative with respect to p, will be needed. This should not come as a surprise as we will approximate the P process by a discrete in time and space process, see Section 3.1.1. We shall not make further assumption on the behavior of $\mathcal{V}$ and its derivatives. Instead, we introduce a proxy to $\mathcal{V}$ that will be used in the proof of convergence of our numerical algorithm.

Let then $\mathcal{V}^\epsilon$ be a smoothing of the decoupling field $\mathcal{V}$, namely: For $\epsilon \in (0,1)$,

$$\mathcal{V}^\epsilon(t,p,e) := \int_{\mathbb{R}^d} \mathcal{V}(t,p+q,e)\varphi_\epsilon(q)\,\mathrm{d}q \quad \text{where } \ \varphi_\epsilon(q) := \frac{1}{\epsilon^d}\varphi\Big(\frac{q}{\epsilon}\Big),\ q \in \mathbb{R}^d, \tag{2.37}$$

and with $\varphi(\cdot)$ a smooth compactly supported probability density function.

We now show that $\mathcal{V}^\epsilon$ satisfies the same PDE as $\mathcal{V}$ up to an error term that we can control. From the previous Assumption 2.1, we first observe that $\mathcal{V}^\epsilon$, defined in (2.37) belongs also to $\mathscr{C}^{1,2,1}([0,T)\times\mathbb{R}^d\times\mathbb{R})$ and that it satisfies the same bounds as $\mathcal{V}$, namely

$$0 \leqslant \partial_e\mathcal{V}^\epsilon(t,p,e) \leqslant \frac{C}{T-t}, \quad |\partial_p\mathcal{V}^\epsilon(t,p,e)| \leqslant C \tag{2.38}$$

$$\text{and } \int_{\mathbb{R}} |\partial_p\mathcal{V}^\epsilon(t,p,z)|\,\mathrm{d}z \leqslant C(T-t), \quad \forall (t,p,e) \in [0,T)\times\mathbb{R}^d\times\mathbb{R}. \tag{2.39}$$

Proposition 2.3. *Under Assumption 2.1, the function $\mathcal{V}^\epsilon$ satisfies on $[0,T)\times\mathbb{R}^d\times\mathbb{R}$,*

$$\partial_t\mathcal{V}^\epsilon + b(p)\partial_p\mathcal{V}^\epsilon + \frac{1}{2}\bar{\sigma}^2\Delta_p\mathcal{V}^\epsilon + \mu(p,\mathcal{V}(t,p,e))\partial_e\mathcal{V}^\epsilon = \theta^\epsilon(t,p,e) \tag{2.40}$$

with

$$|\theta^\epsilon(t,p,e)| \leqslant C\epsilon\bigg(\int_{\mathbb{R}^d} |\partial_p\mathcal{V}(t,p+q,e)|\varphi_\epsilon(q)\,\mathrm{d}q + \partial_e\mathcal{V}^\epsilon(t,p,e)\bigg). \tag{2.41}$$

Proof. Recall that $\mathcal{V}$ satisfies (2.21), and that, for $(t,p,e) \in [0,T)\times\mathbb{R}^d\times\mathbb{R}$,

$$\partial_t\mathcal{V}^\epsilon(t,p,e) = \int_{\mathbb{R}^d} \partial_t\mathcal{V}(t,p+q,e)\varphi_\epsilon(q)\,\mathrm{d}q \text{ and } \partial^2_{pp}\mathcal{V}^\epsilon(t,p,e) = \int_{\mathbb{R}^d} \partial^2_{pp}\mathcal{V}(t,p+q,e)\varphi_\epsilon(q)\,\mathrm{d}q. \tag{2.42}$$

We thus compute, by linearity of ∂_t, Δ_p,

$$\partial_t\mathcal{V}^\epsilon(t,p,e) + \frac{1}{2}\bar{\sigma}^2\Delta_p\mathcal{V}^\epsilon(t,p,e) = -\int_{\mathbb{R}^d} b(p+q)\partial_p\mathcal{V}(t,p+q,e)\varphi_\epsilon(q)\,\mathrm{d}q \tag{2.43}$$

$$- \int_{\mathbb{R}^d} \mu(p+q, \mathcal{V}(t,p+q,e))\partial_e \mathcal{V}(t,p+q,e)\varphi_\epsilon(q)\,\mathrm{d}q \tag{2.44}$$

which leads to

$$\partial_t \mathcal{V}^\epsilon(t,p,e) + b(p)\partial_p \mathcal{V}^\epsilon(t,p,e) + \frac{1}{2}\bar{\sigma}^2 \Delta_p \mathcal{V}^\epsilon(t,p,e) \tag{2.45}$$

$$+ \mu(p, \mathcal{V}(t,p,e))\partial_e \mathcal{V}^\epsilon(t,p,e) = \theta^\epsilon(t,p,e) \tag{2.46}$$

with

$$\theta^\epsilon(t,p,e) := \int_{\mathbb{R}^d} \{b(p) - b(p+q)\}\partial_p \mathcal{V}(t,p+q,e)\varphi_\epsilon(q)\,\mathrm{d}q \tag{2.47}$$

$$+ \int_{\mathbb{R}^d} \{\mu(p, \mathcal{V}(t,p,e)) - \mu(p+q, \mathcal{V}(t,p+q,e))\}\partial_e \mathcal{V}(t,p+q,e)\varphi_\epsilon(q)\,\mathrm{d}q. \tag{2.48}$$

We compute, using the uniform Lipschitz property of b, μ and $\mathcal{V}$ in the p variable and the non-negativity of $\partial_e \mathcal{V}$, that

$$|\theta^\epsilon(t,p,e)| \leqslant C \int_{\mathbb{R}^d} |q||\partial_p \mathcal{V}(t,p+q,e)|\varphi_\epsilon(q)\,\mathrm{d}q + C\int_{\mathbb{R}^d} |q|\partial_e \mathcal{V}(t,p+q,e)\varphi_\epsilon(q)\,\mathrm{d}q \tag{2.49}$$

$$\leqslant C\epsilon\left(\int_{\mathbb{R}^d} |\partial_p \mathcal{V}(t,p+q,e)|\varphi_\epsilon(q)\,\mathrm{d}q + \partial_e \mathcal{V}^\epsilon(t,p,e)\right) \tag{2.50}$$

where we use the fact that φ has compact support to get the last inequality. □

To conclude this section, we collect some useful control in the L^1-sense of the function $\mathcal{V}^\epsilon$ and its derivatives. Their proof is very classical and we skip it.

Lemma 2.1. *Under Assumption 2.1, the followings hold, for $(t,p) \in [0,T) \times \mathbb{R}^d$,*

$$\epsilon \int_{\mathbb{R}} \left|\partial^2_{p_i p_j} \mathcal{V}^\epsilon(t,p,e)\right| \mathrm{d}e + \epsilon^2 \int_{\mathbb{R}} \left|\partial^3_{p_i p_j p_k} \mathcal{V}^\epsilon(t,p,e)\right| \mathrm{d}e \leqslant C, \tag{2.51}$$

for $i,j,k \in \{1,\dots,d\}$. And thus, for $(t,p,p') \in [0,T) \times \mathbb{R}^d \times \mathbb{R}^d$,

$$\int_{\mathbb{R}} \left|\partial_p \mathcal{V}^\epsilon(t,p,e) - \partial_p \mathcal{V}^\epsilon(t,p',e)\right| \mathrm{d}e + \epsilon\int_{\mathbb{R}} \left|\partial^2_{pp} \mathcal{V}^\epsilon(t,p,e) - \partial^2_{pp} \mathcal{V}^\epsilon(t,p',e)\right| \mathrm{d}e \leqslant \frac{C}{\epsilon}|p-p'|, \tag{2.52}$$

where C is a positive constant.

3. Numerical algorithm

In this section, we first consider various versions of the splitting scheme that will be implemented in practice. We then introduce the numerical errors that have to be taken into account in order to obtain the convergence of the scheme. The main error decomposition is given at the end of this section, see Proposition 3.1, together with our main convergence result, see Theorem 3.1. The precise study of each error is postponed to the next section.

3.1. Fully implementable schemes

The schemes we consider are all based on the approximation of the Wiener measure by a discrete and finite set of path. Indeed, on the grid π, the Brownian motion is approximated by a discrete time process denoted $\widehat{W}$. At each date t_n, $\mathfrak{S}^d_n \subset \mathbb{R}^d$ denotes the support of the law of the random variable $\widehat{W}_{t_n}$. To define $\widehat{W}$, we first define $(\Delta\widehat{W}_n := \widehat{W}_{t_{n+1}} - \widehat{W}_{t_n})_{0\leqslant n\leqslant N-1}$ which stands for discrete approximation of the Brownian increments $(W_{t_{n+1}} - W_{t_n})_{0\leqslant n\leqslant N-1}$.

Definition 3.1 (Brownian increments approximation)**.** We use the cubature formula introduced in Section A.2 of [9] which requires only $2d$ points. Denoting $(\mathfrak{e}^\ell)_{1\leqslant \ell\leqslant d}$ the canonical basis of $\mathbb{R}^d$, we set, for $1 \leqslant i \leqslant I = 2d$, $\mathbb{P}(\Delta\widehat{W}_n = \omega^i_{\mathfrak{h}}) = \frac{1}{2d}$ and $\omega^i_{\mathfrak{h}} = -\sqrt{d\mathfrak{h}}\mathfrak{e}^\ell$, if $i = 2\ell$ or $\omega^i_{\mathfrak{h}} = \sqrt{d\mathfrak{h}}\mathfrak{e}^\ell$ if $i = 2\ell - 1$.

By construction, we observe that

$$\text{for } x \in \mathfrak{S}^d_n,\ x + \Delta\widehat{W}_n \in \mathfrak{S}^d_{n+1}. \tag{3.1}$$

Moreover, the set $\mathfrak{S}^d_n$ are discrete and finite.

For later use, we make the following observation.

Lemma 3.1. *The number of points in $\mathfrak{S}^d_n$, denoted $\sharp\mathfrak{S}^d_n$, is $O(n^d)$.*

Proof. By definition, $\mathfrak{S}^d_n$ is a subgrid of $\sqrt{d\mathfrak{h}}\mathbb{Z}^d$. Without loss of generality and for ease of presentation, we assume for this proof that $\sqrt{d\mathfrak{h}} = 1$. For $z \in \mathbb{Z}^d$, we denote $|z|_1 = \sum_{\ell=1}^d |z^\ell|$. We now present two ways of describing the set $\mathfrak{S}^d_n$, for $n \geqslant 0$ and $d \geqslant 1$.

(1) Assume that $n \geqslant 2$. We first observe that the points in $\mathfrak{S}^d_n$ are the points that can be attained after n steps following at each step any direction in $\{\omega^i_{\mathfrak{h}} \mid 1 \leqslant i \leqslant 2d\}$ according to Definition 3.1. To go from step $n-1$ to step n, there are two possibilities: the direction that is followed ends up in $\mathfrak{S}^d_{n-2}$ or it allows to reach $\{z \in \mathbb{Z}^d \mid |z|_1 = n\}$ (the points that are exactly at distance n from the origin). We thus have, for $n \geqslant 2$,

$$\mathfrak{S}^d_n = \mathfrak{S}^d_{n-2} \cup \{z \in \mathbb{Z}^d \mid |z|_1 = n\}, \tag{3.2}$$

and $\mathfrak{S}^d_1 = \{\omega^i_{\mathfrak{h}} | 1 \leqslant i \leqslant 2d\}$, $\mathfrak{S}^d_0 = \{0\}$. Thus,

$$\mathfrak{S}^d_{2m} = \bigcup_{k=0}^{m} \{z \in \mathbb{Z}^d \mid |z|_1 = 2k\} \text{ and } \mathfrak{S}^d_{2m+1} = \bigcup_{k=0}^{m} \{z \in \mathbb{Z}^d \mid |z|_1 = 2k+1\}.$$

In particular, for $d = 1$, we have $\sharp\{z \in \mathbb{Z} \mid |z|_1 = n\} = 2$ for $n > 0$, which yields then that

$$\sharp\mathfrak{S}^1_n = n + 1,\ n \geqslant 0. \tag{3.3}$$

(2) For $d \geqslant 2$,$n \geqslant 1$, the set $\mathfrak{S}^d_n$ can be described by induction on the dimension. Indeed, we observe that

$$\mathfrak{S}^d_n = \bigcup_{k\in\{-n,\dots,n\}} \{(z,k) \in \mathbb{Z}^d | z \in \mathfrak{S}^{d-1}_{n-k}\}.$$

We obtain then

$$\sharp\mathfrak{S}^d_n = \sharp\mathfrak{S}^{d-1}_n + 2\sum_{k=1}^{n} \sharp\mathfrak{S}^{d-1}_{n-k}. \tag{3.4}$$

An easy induction, starting with (3.3), yields then that $\sharp\mathfrak{S}^d_n = O(n^d)$. Note that for $d = 2$, one computes $\sharp\mathfrak{S}^2_n = (n+1)^2$, but one has $\sharp\mathfrak{S}^3_1 = 6$.

$\square$

3.1.1. Schemes for generic diffusion

In this section, we treat the general case where P is solution to a SDE with Lipschitz coefficients as in (1.1). The discrete approximation on the grid π is given by the classical Euler Scheme

$$\widehat{P}_0 = P_0 \;\text{ and }\; \widehat{P}_{t_{n+1}} = \widehat{P}_{t_n} + b\left(\widehat{P}_{t_n}\right)\mathfrak{h} + \sigma\left(\widehat{P}_{t_n}\right)\Delta\widehat{W}_n, \tag{3.5}$$

where $(\Delta\widehat{W}_n)$ is introduced in Definition 3.1. For later use, we also define, for $p \in \mathbb{R}^d$ and any $0 \leqslant n \leqslant N-1$,

$$\widehat{P}^{t_n,p}_{t_{n+1}} := p + b(p)\mathfrak{h} + \sigma(p)\Delta\widehat{W}_n, \tag{3.6}$$

and, for $1 \leqslant i \leqslant 2d$,

$$\left(\widehat{P}^{t_n,p}_{t_{n+1}}\right)^i := p + b(p)\mathfrak{h} + \sigma(p)\omega^i_{\mathfrak{h}}. \tag{3.7}$$

Let us denote by $\mathcal{P}_n$ the discrete and finite support of $\widehat{P}_{t_n}$ for $0 \leqslant n \leqslant N-1$. We observe, due to the very definition of $(\widehat{P}_{t_n})$, that

$$\text{for } p \in \mathcal{P}_n,\; \widehat{P}^{t_n,p}_{t_{n+1}} \in \mathcal{P}_{n+1}. \tag{3.8}$$

In this context, we first introduce a discrete version of the operator $\mathcal{T}$, recall Definition 2.4, that will compute an approximation to (2.15) written in *forward form*: We shall use the celebrated Sticky Particle Dynamics (SPD) [2] see also Section 1.1 of [11]. The SPD is particularly simple to implement in our case, since, due to the monotonicity assumption on (μ, ψ), there is no particle colliding!

For $M \geqslant 1$, let

$$D_M := \left\{\mathrm{e} = (e_1, \dots, e_m, \dots, e_M) \in \mathbb{R}^M \mid e_1 \leqslant \dots \leqslant e_m \leqslant \dots \leqslant e_M\right\} \tag{3.9}$$

$$\text{and } \mathcal{I}_M := \left\{\theta \in \mathscr{I} | \theta(\cdot) := H * \left(\frac{1}{M}\sum_{m=1}^{M} \delta_{e_m}\right), \mathrm{e} \in D_M\right\} \tag{3.10}$$

where H is the Heaviside function $x \mapsto \mathbf{1}_{\{x \geqslant 0\}}$, $*$ the convolution operator and δ_e is the Dirac mass at $e \in \mathbb{R}$.

For later use, we introduce, for $M \geqslant 0$ and $n \in \{0, \dots, N\}$, the set of functions

$$\mathcal{K}^M_n = \{\psi : \mathcal{P}_n \times \mathbb{R} \to [0,1] \mid \text{ for } p \in \mathcal{P}_n,\, \psi(p, \cdot) \in \mathcal{I}_M\}. \tag{3.11}$$

The discrete version of $\mathcal{T}$, denoted $\mathfrak{T}$, acts on empirical CDF belonging to $\mathscr{I}_M$ and is given by

$$(0, \infty) \times \mathbb{R}^d \times \mathcal{I}_M \ni (h, p, \theta) \mapsto \mathfrak{T}^M_h(p, \theta) := H * \left(\frac{1}{M}\sum_{m=1}^{M} \delta_{e^p_m(h)}\right) \in \mathcal{I}_M. \tag{3.12}$$

Above $(e^p_m(h))_{1 \leqslant m \leqslant M}$ is a set (of positions) of particles computed as follows. Given the initial position $\mathrm{e}(0) \in D_M$ (representing θ) and velocities $(\bar{F}_m(p))_{1 \leqslant m \leqslant M}$ set to $\bar{F}_m(p) = -\int_{(m-1)/M}^{m/M} \mu(p, y)\,\mathrm{d}y$, we consider M particles $(e^p_m)_{1 \leqslant m \leqslant M}$, whose positions at time $t \in [0, h]$ are simply given by

$$e^p_m(t) = e_m(0) + \bar{F}_m(p)t. \tag{3.13}$$

We observe that $(e^p_m(t))_{1 \leqslant m \leqslant M} \in D_M$, for all $t \in [0, h]$, as $-\mu$ is non-decreasing.

We now present the approximation used for the approximation of $\mathcal{D}_{\mathfrak{h}}(\psi)$, recall Definition 2.5 and (2.18).

Let $\psi \in \mathcal{K}^M_{n+1}$ and fix $\mathrm{p} \in \mathcal{P}_n$. We observe

$$\bar{v}_n(\mathrm{p}, e) := \mathbb{E}\left[\psi\left(\widehat{P}^{t_n,\mathrm{p}}_{t_{n+1}}, e\right)\right] = \frac{1}{2d}\sum_{i=1}^{2d} \psi\left(\left(\widehat{P}^{t_n,p}_{t_{n+1}}\right)^i, e\right) \tag{3.14}$$

recall (3.7).

Since $\psi \in \mathcal{K}^M_{n+1}$, recall (3.11), we know that, for $1 \leqslant i \leqslant 2d$, there is some $\mathrm{e}^i \in D_M$, depending implicitely on $\psi((\widehat{P}^{t_n,\mathrm{p}}_{t_{n+1}})^i, \cdot)$, such that

$$\psi\left(\left(\widehat{P}^{t_n,\mathrm{p}}_{t_{n+1}}\right)^i, e\right) = H * \left(\frac{1}{M}\sum_{m=1}^{M} \delta_{e^i_m}\right)(e).$$

By linearity, (3.14) reads

$$\bar{v}_n(\mathrm{p}, e) = H * \left(\frac{1}{2dM}\sum_{i=1}^{I}\sum_{m=1}^{M} \delta_{e^i_m}\right)(e). \tag{3.15}$$

The approximation of the diffusion operator is thus given by

$$\mathcal{K}^M_{n+1} \ni \psi \mapsto \mathfrak{D}^M_n(\psi) := \bar{v}_n \in \mathcal{K}^{2Md}_n. \tag{3.16}$$

Finally, the scheme will have essentially two versions:

CASE 1: We keep all the particles at each step. The overall procedure will then be the iteration of the two operators $\mathfrak{T}$ and $\mathfrak{D}$, see Definition 3.2 below.

CASE 2: We keep the number of particles constant between two steps. To reduce the number of particles after the use of $\mathfrak{T}$, we apply another operator $\mathfrak{R}$, namely, for $M \geqslant m \geqslant 1$,

$$\mathcal{I}^M \ni \psi \mapsto \mathfrak{R}^{M,m}(\psi) \in \mathcal{I}^m. \tag{3.17}$$

Various implementations are possible, we refer to the numerical section for a precise description. The overall procedure for this case is given in Definition 3.3 below.

Let us now finally introduce the scheme formally.

Definition 3.2 (Scheme in CASE 1)**.** Fix $M \geqslant 1$ and set for $0 \leqslant n \leqslant N$, $M_n := (2d)^{N-n}M$.

(1) At $n = N$: Set $\mathrm{e}_N := (\Lambda, \ldots, \Lambda) \in D_{M_N}$ whose empirical CDF is the terminal condition $\phi = \mathbf{1}_{\{e \geqslant \Lambda\}}$, recall (1.2). Setting simply $u^{N,M}_N(p, \cdot) = \phi$, for all $p \in \mathcal{P}_N$, we observe that $u^{N,M}_N \in \mathcal{K}^{M_N}_N$.

(2) For $n < N$: Given $u^{N,M}_{n+1} \in \mathcal{K}^{M_{n+1}}_{n+1}$, define

$$\bar{u}^{N,M}_n = \mathfrak{D}^{M_{n+1}}_n\left(u^{N,M}_{n+1}\right) \in \mathcal{K}^{M_n}_n, \tag{3.18}$$

recall (3.16), and then $u^{N,M}_n$ by, for each $p \in \mathcal{P}_n$,

$$u^{N,M}_n(p, \cdot) = \mathfrak{T}^{M_n}_{\mathfrak{h}}\left(p, \bar{u}^{N,M}_n(p, \cdot)\right), \tag{3.19}$$

recall (3.12).

The approximation of $\mathcal{V}(0, P_0, \cdot)$ is then given by $u^{N,M}_0(P_0, \cdot)$.

Definition 3.3 (Scheme in CASE 2)**.** Fix $M \geqslant 1$.

(1) At $n = N$: Set $\mathrm{e}_N := (\Lambda, \cdots, \Lambda) \in D_M$ whose empirical CDF is the terminal condition $\phi = \mathbf{1}_{\{e \geqslant \Lambda\}}$, recall (1.2). Setting simply $v^{N,M}_N(p, \cdot) := \phi$, for all $p \in \mathcal{P}_N$, we observe that $v^{N,M}_N \in \mathcal{K}^M_N$.

(2) For $n < N$: Given $v_{n+1}^{N,M} \in \mathcal{K}_{n+1}^M$, define $\bar{v}_n^{N,M} = \mathfrak{D}_n^M(v_{n+1}^{N,M}) \in \mathcal{K}_n^{2dM}$, recall (3.16), and then for each $p \in \mathcal{P}_n$,

$$\check{v}_n^{N,M}(p,\cdot) = \mathfrak{R}^{2dM,M}\big(\bar{v}_n^{N,M}(p,\cdot)\big) \tag{3.20}$$

recall (3.17), and finally $v_n^{N,M}(p,\cdot)$ by,

$$v_n^{N,M}(p,\cdot) = \mathfrak{T}_{\mathfrak{h}}^M\big(p, \check{v}_n^{N,M}(p,\cdot)\big), \tag{3.21}$$

recall (3.12).

The approximation of $\mathcal{V}(0, P_0, \cdot)$ is then given by $v_0^{N,M}(P_0,\cdot)$.

Remark 3.1. Let us insist on the fact that, unless some recombination occurs at the tree level for P, the above schemes become rapidly untractable. The worst case complexity is easily obtained for the tree $\widehat{P}$: at each step n, there is at most $(2d)^n$ branches, $0 \leqslant n \leqslant N-1$. Implementing these schemes, one faces two issues: the number of operations to perform at each step when applying the discrete transport operator, see (3.12), and the number of particles to keep to be able to apply the approximation of the diffusion operator (3.15). Indeed, if one is only interested in the approximation of $\mathcal{V}(0, P_0, \cdot)$ then the sum appearing in (3.15) is only computed at time 0 but again, in order to do that, one needs to keep track of all the particles until time 0.

In **CASE 1**, see Definition 3.2, we obtain by backward induction that the number of particles at each step for a given branch is $M_n = (2d)^{N-n}M$. At each step n, the discrete transport operator must be called $(2d)^n$ times (the number of branches at most), each call requiring to recompute the position of M_{n+1} particles, see (3.13): this will cost $O((2d)^{N-1}M)$ for step $n < N$. And at step 0, one will evaluate (3.15) that will cost $O((2d)^N M)$. Overall, the complexity is $O((2d)^N MN)$.

In **CASE 2**, see Definition 3.3, the only improvement comes from the fact that the number of particles at each step for a given branch is always M. The cost of applying the discrete transport operator at step n is now $(2d)^n M$, summing on the total number of steps we obtain a complexity of $O((2d)^N M)$. And at step 0, one will evaluate (3.15) that will cost $O(2dM)$. Overall, the complexity is reduced to $O((2d)^N M)$.

3.1.2. A simplified functional Brownian setting

We now consider a special case for the process P for which the numerical implementation is less computationally demanding, see Remark 3.3 below.

Assumption 3.1. *The process P is given as a function of the Brownian motion, namely*

$$P_t = \mathfrak{P}(t, W_t) \ \text{ for } \ \mathfrak{P} : [0,T] \times \mathbb{R}^d \mapsto \mathbb{R}^d.$$

Remark 3.2. The main application of this case is when P is the Brownian motion itself (also known as the Bachelier model in financial application) or a Geometric Brownian Motion (also known as Black–Scholes model in financial application).

Compared to the previous section, the only difference is the approximation of P. In this context, it is naturally given by

$$\widehat{P}_t := \mathfrak{P}\Big(t, \widehat{W}_t\Big),\ t \in \pi. \tag{3.22}$$

Note that we slightly abuse the notation here by keeping the same ones as the ones introduced in the previous section.

So, for $0 \leqslant n \leqslant N$,

$$\mathcal{P}_n = \{\mathrm{p} = \mathfrak{P}(t_n, \mathrm{w}), \mathrm{w} \in \mathfrak{S}_n\}, \tag{3.23}$$

which is the (discrete) support of $\widehat{P}_{t_n}$. We also define, for $p = \mathfrak{P}(t_n, \mathrm{w})$,

$$\widehat{P}^{t_n,p}_{t_{n+1}} = \mathfrak{P}\Big(t_{n+1}, \mathrm{w} + \Delta\widehat{W}_n\Big) \text{ and } \Big(\widehat{P}^{t_n,p}_{t_{n+1}}\Big)^i = \mathfrak{P}\big(t_{n+1}, \mathrm{w} + \omega^i_{\mathfrak{h}}\big), \tag{3.24}$$

for $1 \leqslant i \leqslant 2d$, recall Definition 3.1.

Now that the approximation of P has been clarified in this context, we can use the various schemes given in Definitions 3.2 and 3.3.

Remark 3.3. (i) The main difference between the generic diffusion case and the functional Brownian setting comes from the numerical complexity associated to their implementation. Indeed, it is the case that the growth size of $\mathfrak{S}^d_n$ with respect to n is tamed (for low dimensional problem) because the "tree" associated to the brownian motion is naturally recombining, which is *a priori* not the case for generic diffusion. According to Lemma 3.1, $\sharp\mathfrak{S}^d_n = O(n^d)$. Following Remark 3.1, we can estimate the complexity of the algorithms in the functional Brownian setting.

In **CASE 1**, we have $M_{n+1} = O(M(N-n+1)^d)$ and $O(n^d)$ branches at each step n, we thus compute that the complexity is $O(M\sum_{n=0}^{N-1}(N-n)^d n^d) = O(MN^{2d+1})$ for large N. In **CASE 2**, we have $M_{n+1} = M$ and still $O(n^d)$ branches at each step n, we thus compute that the complexity is $O(M\sum_{n=0}^{N-1} n^d) = O(MN^{d+1})$ for large N.

Thus, the methods we study here are impacted by the "curse of dimensionality". However, we should also note that for dimension around 5 and in the case of the brownian setting, the problems are tractable numerically, in particular when working with **CASE 2**, see Section 5 for illustration.

(ii) For implementation purposes, we just use the Brownian setting. A way to control the growth for generic diffusion is to introduce some interpolation on a grid see *e.g.* [5] or to use recombination techniques developped for cubature methods, see *e.g.* [12].

3.2. Error decomposition and main result

We first present the various sources of errors introduced by the schemes above. In the sequel, we shall mainly conduct our analysis for the scheme given in CASE 1.

The error we seek to control is, for $\gamma_n = u^{N,M}_n$ or $\gamma_n = v^{N,M}_n$, recall Definitions 3.2 and 3.3,

$$\mathrm{err}(N, M) := \int_{\mathbb{R}} |\gamma_0(P_0, e) - \mathcal{V}(0, P_0, e)| \,\mathrm{d}e. \tag{3.25}$$

And we first observe

$$\mathrm{err} \leqslant \mathcal{E}_0(P_0) + \mathcal{E}^r_0, \tag{3.26}$$

with for $0 \leqslant n \leqslant N$ and $p \in \mathcal{P}_n$,

$$\mathcal{E}^r_n(p) := \int_{\mathbb{R}} |\mathcal{V}^\epsilon(t_n, p, e) - \mathcal{V}(t_n, p, e)| \,\mathrm{d}e, \tag{3.27}$$

$$\mathcal{E}_n(p) := \int_{\mathbb{R}} |\gamma_n(p, e) - \mathcal{V}^\epsilon(t_n, p, e)| \,\mathrm{d}e. \tag{3.28}$$

In CASE 1, we define, for $p \in \mathcal{P}_n$,

$$\mathcal{E}^{\mathrm{T}}_n(p) := \int_{\mathbb{R}} \Big|\mathfrak{T}^{M_n}_{\mathfrak{h}}\big(p, \bar{u}^{N,M}_n(p,\cdot)\big)(e) - \tilde{\mathcal{T}}_h\big(p, \bar{u}^{N,M}_n(p,\cdot)\big)(e)\Big| \,\mathrm{d}e, \tag{3.29}$$

$$\bar{\mathcal{E}}_n(p) := \int_{\mathbb{R}} \Big|\tilde{\mathcal{T}}_h\big(p, \bar{u}^{N,M}_n(p,\cdot)\big)(e) - \tilde{\mathcal{T}}_h\Big(p, \mathbb{E}\Big[\mathcal{V}^\epsilon\Big(t_{n+1}, \widehat{P}^{t_n,p}_{t_{n+1}}, \cdot\Big)\Big]\Big)(e)\Big| \,\mathrm{d}e, \tag{3.30}$$

recall Definition 3.2 and Remark 2.1.

Moreover, for $p \in \mathbb{R}^d$, we set

$$\mathcal{E}_n^D(p) := \int_{\mathbb{R}} \Big| \tilde{\mathcal{T}}_h\Big(p, \mathbb{E}\Big[\mathcal{V}^\epsilon\Big(t_{n+1}, \widehat{P}_{t_{n+1}}^{t_n,p}, \cdot\Big)\Big]\Big)(e) - \tilde{\mathcal{T}}_h\Big(p, \mathbb{E}\Big[\mathcal{V}^\epsilon\Big(t_{n+1}, P_{t_{n+1}}^{t_n,p}, \cdot\Big)\Big]\Big)(e)\Big| \,\mathrm{d}e, \tag{3.31}$$

$$\mathcal{E}_n^S(p) := \int_{\mathbb{R}} \Big| \tilde{\mathcal{T}}_h\Big(p, \mathbb{E}\Big[\mathcal{V}^\epsilon\Big(t_{n+1}, P_{t_{n+1}}^{t_n,p}, \cdot\Big)\Big]\Big)(e) - \mathcal{V}^\epsilon(t_n, p, e)\Big| \,\mathrm{d}e. \tag{3.32}$$

We have the following.

Lemma 3.2. *Under* CASE 1, *for* $0 \leqslant n \leqslant N-2$,

$$\mathcal{E}_n(p) \leqslant \mathcal{E}_n^{\mathrm{T}}(p) + \bar{\mathcal{E}}_n(p) + \mathcal{E}_n^D(p) + \mathcal{E}_n^S(p) \tag{3.33}$$

and

$$\mathcal{E}_{N-1}(p) \leqslant \mathcal{E}_{N-1}^{\mathrm{T}}(p) + \mathcal{E}_{N-1}^{\mathrm{r}}(p) + \int_{\mathbb{R}} |\mathcal{S}_{\mathfrak{h}}(\phi(\cdot))(p,e) - \Theta_{\mathfrak{h}}(\phi(\cdot))(p,e)| \,\mathrm{d}e \tag{3.34}$$

observing that

$$\mathcal{E}_{N-1}^{\mathrm{T}}(p) = \int_{\mathbb{R}} \big|\mathfrak{T}_{\mathfrak{h}}^M(p, \phi(\cdot))(e) - \mathcal{T}_{\mathfrak{h}}(p, \phi(\cdot)(e))\big| \,\mathrm{d}e. \tag{3.35}$$

Proof. The first inequality is straightforward. For the second, we observe that $u_N^{N,M} = \phi$ and does not depend on p variable so that $\bar{u}_{N-1}^{N,M} = \phi$.

$$\begin{aligned}
\mathcal{E}_{N-1}(p) &:= \int_{\mathbb{R}} \Big| u_{N-1}^{N,M}(p,e) - \mathcal{V}^\epsilon(t_{N-1}, p, e)\Big| \,\mathrm{d}e \\
&= \int_{\mathbb{R}} \big|\mathfrak{T}_{\mathfrak{h}}^M(p, \phi(\cdot))(e) - \mathcal{V}^\epsilon(t_{N-1}, p, e)\big| \,\mathrm{d}e \\
&\leqslant \int_{\mathbb{R}} \big|\mathfrak{T}_{\mathfrak{h}}^M(p, \phi(\cdot))(e) - \mathcal{T}_{\mathfrak{h}}(p, \phi(\cdot))(e)\big| \,\mathrm{d}e + \int_{\mathbb{R}} |\mathcal{T}_{\mathfrak{h}}(p, \phi(\cdot))(e) - \mathcal{V}(t_{N-1}, p, e)| \,\mathrm{d}e \\
&\quad + \int_{\mathbb{R}} |\mathcal{V}(t_{N-1}, p, e) - \mathcal{V}^\epsilon(t_{N-1}, p, e)| \,\mathrm{d}e.
\end{aligned}$$

We then observe that $\mathcal{T}_{\mathfrak{h}}(p, \phi(\cdot))(e) = \mathcal{S}_{\mathfrak{h}}(\phi(\cdot))(p,e)$ as again ϕ does not depend on p. □

We now comment quickly the various errors appearing above: The term $\mathcal{E}_n^{\mathrm{T}}(p)$ is the local error due to the approximation of the transport operator by the SPD. The term $\bar{\mathcal{E}}_n(p)$ will allow to propagate the local errors thanks to the L^1 stability of $\mathcal{T}_h(\cdot)$. The term $\mathcal{E}^D$ is linked to the approximation of the diffusion by the tree and $\mathcal{E}^S$ is the splitting error applied to the proxy $\mathcal{V}^\epsilon$.

Proposition 3.1 (Stability)**.** *Under* CASE 1, *the following hold*

$$\mathcal{E}_0(P_0) \leqslant \sum_{n=0}^{N-2} \mathbb{E}\Big[\mathcal{E}_n^{\mathrm{T}}\Big(\widehat{P}_{t_n}\Big) + \mathcal{E}_n^D\Big(\widehat{P}_{t_n}\Big) + \mathcal{E}_n^S\Big(\widehat{P}_{t_n}\Big)\Big] + \mathbb{E}\Big[\mathcal{E}_{N-1}\Big(\widehat{P}_{t_{N-1}}\Big)\Big]. \tag{3.36}$$

Proof. Since $\tilde{\mathcal{T}}_h$ is L^1-stable, (3.30) yields for $p \in \mathcal{P}_n$,

$$\bar{\mathcal{E}}_n(p) \leqslant \int_{\mathbb{R}} \Big| \bar{u}^{N,M}(t_n, p, e) - \mathbb{E}\Big[\mathcal{V}^\epsilon\Big(t_{n+1}, \widehat{P}_{t_{n+1}}^{t_n,p}, e\Big)\Big]\Big| \,\mathrm{d}e.$$

Under CASE 1, recall (3.18), we then get

$$\bar{\mathcal{E}}_n(p) \leqslant \mathbb{E}\left[\int_{\mathbb{R}}\left|u^{N,M}\left(t_{n+1},\widehat{P}^{t_n,p}_{t_{n+1}},e\right)-\mathcal{V}^{\epsilon}\left(t_{n+1},\widehat{P}^{t_n,p}_{t_{n+1}},e\right)\right|\mathrm{d}e\right]. \tag{3.37}$$

Thus,

$$\mathbb{E}\left[\bar{\mathcal{E}}_n\left(\widehat{P}_{t_n}\right)\right] \leqslant \mathbb{E}\left[\mathcal{E}_{n+1}\left(\widehat{P}_{t_{n+1}}\right)\right]. \tag{3.38}$$

We thus deduce from Lemma 3.2

$$\mathbb{E}\left[\mathcal{E}_n\left(\widehat{P}_{t_n}\right)\right] \leqslant \mathbb{E}\left[\mathcal{E}^{\mathrm{T}}_n\left(\widehat{P}_{t_n}\right)+\mathcal{E}^{D}_n\left(\widehat{P}_{t_n}\right)+\mathcal{E}^{S}_n\left(\widehat{P}_{t_n}\right)\right]+\mathbb{E}\left[\mathcal{E}_{n+1}\left(\widehat{P}_{t_{n+1}}\right)\right]. \tag{3.39}$$

The proof is concluded by induction observing $\mathbb{E}[\mathcal{E}_N(\widehat{P}_{t_N})] = 0$. □

The next section is dedicated to the study of the various error appearing above. We prove in Section 4.5 below the main theoretical result of this work:

Theorem 3.1. *Let Assumption 2.1 hold. Then, under* CASE 1*,*

$$\mathrm{err}(N,M) \leqslant C\left(\frac{1}{M}+\frac{\sqrt{\mathfrak{h}}}{\epsilon^2}+\epsilon\right), \tag{3.40}$$

recall (3.25). Moreover, setting $\epsilon = \mathfrak{h}^{\frac{1}{6}}$ and $M = \frac{1}{\epsilon}$, we have

$$\int_{\mathbb{R}}\left|\mathcal{V}(0,P_0,e)-u^{N,M}_0(P_0,e)\right|\mathrm{d}e \leqslant C\mathfrak{h}^{\frac{1}{6}}. \tag{3.41}$$

4. Study of the errors

4.1. Approximation of transport operator

We discuss here the error introduced by the use of the SPD approximation in our framework. The numerical analysis of this method is now well known see *e.g.* [11].

Lemma 4.1. *Under Assumption 2.1 and in* CASE 1*, the following holds*

$$\mathcal{E}^{T}_n(p) \leqslant C\frac{\mathfrak{h}}{M_n},\ p \in \mathcal{P}_n. \tag{4.1}$$

Proof. By Definition 3.2, for $p \in \mathcal{P}_n$ and $n \leqslant N-2$, we have

$$\mathcal{E}^{T}_n(p) = \int_{\mathbb{R}}\left|\mathfrak{T}^{M_n}_{\mathfrak{h}}\left(\bar{u}^{N,M}_n(p,\cdot),p\right)-\tilde{\mathcal{T}}_{\mathfrak{h}}\left(\bar{u}^{N,M}_n(p,\cdot),p\right)\right|\mathrm{d}e \tag{4.2}$$

with $\bar{u}^{N,M}_n(p,\cdot) \in \mathcal{I}_{M_n}$. For fixed p, the property of $\mathfrak{T}^{M}_{\mathfrak{h}}$ are discussed in Section 1.1 of [11]. We use here the estimate given in Theorem 3.1 of [11], observing that, in our context, there is no error due to the discretization of the initial condition. Indeed, $\bar{u}^{N,M}_n(p,\cdot)$ appears already as the CDF of an empirical distribution. Thus, straightforwardly, we get

$$\mathcal{E}^{T}_n(p) \leqslant C\frac{\mathfrak{h}}{M_n}, \tag{4.3}$$

where C does not depend on p but depends on the Lipschitz constant of $\partial_y\mu$, recall Assumption 2.1.

For the step $n = N-1$, we have,

$$\mathcal{E}^{\mathrm{T}}_{N-1}(p) = \int_{\mathbb{R}}\left|\mathfrak{T}^{M}_{\mathfrak{h}}(p,\phi(\cdot))(e)-\mathcal{T}_{\mathfrak{h}}(p,\phi(\cdot)(e))\right|\mathrm{d}e \tag{4.4}$$

recall (3.35). It turns out that for our application ϕ is trivially the CDF of an empirical distribution, recall (1.2). Thus, we indeed have, $\mathcal{E}^{T}_{N-1}(p) \leqslant C\frac{\mathfrak{h}}{M_{N-1}}$. □

4.2. Regularization error

Proposition 4.1. *Under Assumption 2.1, the following holds, for $n \leqslant N$,*

$$\mathcal{E}^r_n(p) \leqslant C\epsilon,\ p \in \mathbb{R}^d.$$

Proof. We first observe that $\mathcal{E}^r_N(p) = 0$ as ϕ does not depend on p. For $n < N$, we have that

$$\mathcal{E}^r_n(p) = \int_{\mathbb{R}} |\mathcal{V}^\epsilon(t_n, p, e) - \mathcal{V}(t_n, p, e)| \,\mathrm{d}e \tag{4.5}$$

$$= \int_{\mathbb{R}} \left| \int_{\mathbb{R}^d} \{\mathcal{V}(t_n, p+q, e) - \mathcal{V}(t_n, p, e)\} \varphi_\epsilon(q) \,\mathrm{d}q \right| \mathrm{d}e \tag{4.6}$$

$$\leqslant \int_{\mathbb{R}^d} \int_{\mathbb{R}} |\mathcal{V}(t_n, p+q, e) - \mathcal{V}(t_n, p, e)| \,\mathrm{d}e \varphi_\epsilon(q) \,\mathrm{d}q. \tag{4.7}$$

Thus, using (2.24), we compute

$$\mathcal{E}^r_0(p) \leqslant C \int_{\mathbb{R}^d} |q| \varphi_\epsilon(q) \,\mathrm{d}q \leqslant C\epsilon, \tag{4.8}$$

which concludes the proof. □

4.3. Splitting error

Recall that the splitting error is given by

$$\mathcal{E}^S_n(p) = \int_{\mathbb{R}} \left| \tilde{\mathcal{T}}_h \Big(p, \mathbb{E}\Big[\mathcal{V}^\epsilon \Big(t_{n+1}, P^{t_n,p}_{t_{n+1}}, \cdot \Big) \Big] \Big)(e) - \mathcal{V}^\epsilon(t_n, p, e) \right| \mathrm{d}e. \tag{4.9}$$

for $0 \leqslant n \leqslant N-1$, $p \in \mathbb{R}^d$.

In [6], the error due the theoretical splitting has already been studied and the results obtained there can be used in our setting. However, we should point out that here $\mathcal{V}^\epsilon$ appears instead of $\mathcal{V}$. We thus first observe the following.

Lemma 4.2. *Under Assumption 2.1, the following holds, for $p \in \mathbb{R}^d$,*

$$\mathcal{E}^S_n(p) \leqslant C\big(1 + |p|^2\big) h^{\frac{3}{2}} + \mathfrak{E}_n(p) \tag{4.10}$$

with

$$\mathfrak{E}_n(p) := \int_{\mathbb{R}} |\Theta_h(\mathcal{V}^\epsilon(t_{n+1}, \cdot))(p, e) - \mathcal{V}^\epsilon(t_n, p, e)| \,\mathrm{d}e. \tag{4.11}$$

Proof. We compute

$$\mathcal{E}^S_n(p) \leqslant \int_{\mathbb{R}} \left| \tilde{\mathcal{T}}_h \Big(p, \mathbb{E}\Big[\mathcal{V}^\epsilon \Big(t_{n+1}, P^{t_n,p}_{t_{n+1}}, \cdot \Big) \Big] \Big)(e) - \Theta_h(\mathcal{V}^\epsilon(t_{n+1}, \cdot))(p, e) \right| \mathrm{d}e$$
$$+ \int_{\mathbb{R}} |\Theta_h(\mathcal{V}^\epsilon(t_{n+1}, \cdot))(p, e) - \mathcal{V}^\epsilon(t_n, p, e)| \,\mathrm{d}e.$$

The first term in the RHS above is then controlled by invoking Proposition 2.1. □

It remains to study $\mathfrak{E}_n$, recall (4.11). The upper bound for this term is obtained in Proposition 4.2 below, which allows to conclude for the splitting error in Corollary 4.1. We first need the following result.

Lemma 4.3. *Let Assumption 2.1 hold. Then, for $n \leqslant N-2$,*

$$\sup_{(t,p,e)\in[t_n,t_{n+1}]\times\mathbb{R}^d\times\mathbb{R}} \left|\Theta_{t_{n+1}-t}(\mathcal{V}^\epsilon(t_{n+1},\cdot))(p,e) - \mathcal{V}^\epsilon(t,p,e)\right| \leqslant C\epsilon, \tag{4.12}$$

where C is a positive constant.

Proof. Without loss of generality, we do the proof for $n = 0$, working on $[0, t_1 = \mathfrak{h}]$. We denote $w(s,\cdot) = \Theta_{t_1-s}(\mathcal{V}^\epsilon(t_1,\cdot))$ for $s \in [0,t_1]$. Recall that the following FBSDE is well posed: for $s \in [0,t_1]$,

$$\begin{cases} P_s = \mathrm{p} + \displaystyle\int_0^s b(P_t)\,\mathrm{d}t + \bar{\sigma} W_s, \\ E_s = \mathrm{e} + \displaystyle\int_0^s \mu(P_t, Y_t)\,\mathrm{d}t, \\ Y_s = \mathcal{V}^\epsilon(t_1, P_{t_1}, E_{t_1}) - \displaystyle\int_s^{t_1} Z_t\,\mathrm{d}W_t, \end{cases} \tag{4.13}$$

with $Y_s = w(s, P_s, E_s)$ for $s \in [0,t_1]$. Without loss of generality, we will prove (4.12) at the point $(0,\mathrm{p},\mathrm{e})$ (and for $n=0$).

Let $V_t = Y_t - \mathcal{V}^\epsilon(t, P_t, E_t)$. Applying Ito's formula and using the martingale property of Y we get,

$$\mathrm{d}V_t = \mathrm{d}\mathcal{M}_t - \mu(P_t,Y_t)\partial_e \mathcal{V}^\epsilon(t,P_t,E_t)\,\mathrm{d}t \tag{4.14}$$

$$- \left\{\partial_t \mathcal{V}^\epsilon(t,P_t,E_t) + b(P_t)\partial_p \mathcal{V}^\epsilon(t,P_t,E_t) + \frac{1}{2}\Delta_p \mathcal{V}^\epsilon(t,P_t,E_t)\right\}\mathrm{d}t \tag{4.15}$$

where $\mathcal{M}$ is a square integrable martingale with $\mathcal{M}_0 = 0$.

Using the PDE (2.40) satisfied by $\mathcal{V}^\epsilon$ we get

$$\mathrm{d}V_t = \mathrm{d}\mathcal{M}_t - (\{\mu(P_t,Y_t) - \mu(P_t,\mathcal{V}^\epsilon(t,P_t,E_t))\}\partial_e \mathcal{V}^\epsilon(t,P_t,E_t) + \theta^\epsilon(t,P_t,E_t))\,\mathrm{d}t. \tag{4.16}$$

Observe that

$$\mu(P_t,Y_t) - \mu(P_t,\mathcal{V}^\epsilon(t,P_t,E_t)) = c_t V_t \;\text{ with }\; c_t := \int_0^1 \partial_y \mu(P_t, \mathcal{V}^\epsilon(t,P_t,E_t) + \lambda V_t)\,\mathrm{d}\lambda \tag{4.17}$$

and from the property of μ, recall (2.4), we have

$$c_t \leqslant -\ell_1 < 0. \tag{4.18}$$

We get, for $s \in [0,t_1]$

$$V_s - V_0 = -\int_0^s (c_t V_t \partial_e \mathcal{V}^\epsilon(t,P_t,E_t) + \theta^\epsilon(t,P_t,E_t))\,\mathrm{d}t + \mathcal{M}_s. \tag{4.19}$$

We set, for $0 \leqslant t \leqslant t_1$, $\mathcal{E}_t = e^{\int_0^t c_s \partial_e \mathcal{V}^\epsilon(s,P_s,E_s)\,\mathrm{d}s}$ and, we have

$$0 \leqslant \mathcal{E}_t \leqslant 1, \quad \text{for all } 0 \leqslant t \leqslant t_1, \tag{4.20}$$

since $\partial_e \mathcal{V}^\epsilon$ is non negative, recall also (4.18). We then compute

$$\mathcal{E}_s V_s - V_0 = -\int_0^s \theta^\epsilon(t,P_t,E_t)\mathcal{E}_t\,\mathrm{d}t + \mathcal{N}_s \tag{4.21}$$

with $\mathcal{N}$ a square integrable martingale such that $\mathcal{N}_0 = 0$. In particular, recalling (2.41), we get

$$V_0 \leqslant |\mathbb{E}[\mathcal{E}_s V_s]| + C\mathfrak{h}\epsilon + C\epsilon\mathbb{E}\left[\int_0^s \partial_e \mathcal{V}^\epsilon(t, P_t, E_t)\mathcal{E}_t \,\mathrm{d}t\right] \tag{4.22}$$

where we use the fact that $|\partial_p \mathcal{V}|$ is bounded and $\partial_e \mathcal{V}^\epsilon(t, P_t, E_t) \geqslant 0$. The above inequality reads also

$$|V_0| \leqslant |\mathbb{E}[\mathcal{E}_s V_s]| + C\mathfrak{h}\epsilon - C\epsilon\mathbb{E}\left[\int_0^s \frac{c_t}{|c_t|}\partial_e \mathcal{V}^\epsilon(t, P_t, E_t)\mathcal{E}_t \,\mathrm{d}t\right]. \tag{4.23}$$

Observing that

$$\dot{\mathcal{E}}_t = c_t \partial_e \mathcal{V}^\epsilon(t, P_t, E_t)\mathcal{E}_t,$$

since $|c_t| \geqslant \ell_1 > 0$, we obtain

$$|V_0| \leqslant |\mathbb{E}[\mathcal{E}_s V_s]| + C(1 + \mathfrak{h})\epsilon \tag{4.24}$$

and since $V_{t_1} = 0$

$$|w(0, \mathrm{p}, \mathrm{e}) - \mathcal{V}^\epsilon(0, \mathrm{p}, \mathrm{e})| = |V_0| \leqslant C\epsilon,$$

which concludes the proof. □

Proposition 4.2. *Let Assumption 2.1 hold. Then, for $n \leqslant N - 2$, $p \in \mathbb{R}^d$,*

$$\mathfrak{E}_n(p) = \int_{\mathbb{R}} |\Theta_{\mathfrak{h}}(\mathcal{V}^\epsilon(t_{n+1}, \cdot))(p, e) - \mathcal{V}^\epsilon(t_n, p, e)| \,\mathrm{d}e \leqslant C\mathfrak{h}\epsilon. \tag{4.25}$$

Proof. Without loss of generality, we prove the statement in the same setting as the one used for the proof of Lemma 4.3. We just stress here the dependence upon the initial condition: $V_t^{\mathrm{e}} = Y_t^{\mathrm{e}} - \mathcal{V}^\epsilon(t, P_t, E_t^{\mathrm{e}})$ since $E_0 = \mathrm{e}$. We consider the tangent process $\partial_e E$ given by

$$\partial_e E_t = 1 + \int_0^t \partial_y \mu(P_s, Y_s)\partial_e w(s, P_s, E_s)\partial_e E_s \,\mathrm{d}s, \tag{4.26}$$

$$= e^{\int_0^t \partial_y \mu(P_s, Y_s)\partial_e w(s, P_s, E_s)\,\mathrm{d}s}. \tag{4.27}$$

And we observe that $0 \leqslant \partial_e E_t \leqslant 1$, for all $0 \leqslant t \leqslant t_1$, since $\partial_e w \geqslant 0$ and μ is decreasing in y, recall (2.4).

In order to bound the error $\int_{\mathbb{R}} |V_0^e| \,\mathrm{d}e$, we will study the dynamics of $t \mapsto \int_{\mathbb{R}} |\mathbb{E}[V_t^e \partial_e E_t^e]| \,\mathrm{d}e$.

Recalling (4.19), we compute

$$\mathrm{d}(V_t^e \partial_e E_t) = \mathrm{d}\mathcal{N}_t + V_t^e \partial_y \mu(P_t, Y_t)\partial_e w(t, P_t, E_t)\partial_e E_t \,\mathrm{d}t \tag{4.28}$$

$$- (c_t V_t^e \partial_e \mathcal{V}^\epsilon(t, P_t, E_t)\partial_e E_t + \theta^\epsilon(t, P_t, E_t^e)\partial_e E_t) \,\mathrm{d}t, \tag{4.29}$$

where $\mathcal{N}$ is a square integrable martingale satisfying $\mathcal{N}_0 = 0$.

$$\mathbb{E}\big[V_{\mathfrak{h}}^e \partial_e E_{\mathfrak{h}} - V_0^e\big] = \mathbb{E}\left[\int_0^{\mathfrak{h}} V_t^e(\partial_y \mu(P_t, Y_t)\partial_e[w(t, P_t, E_t^e)] - c_t V_t^e \partial_e[\mathcal{V}^\epsilon(t, P_t, E_t^e)]) \,\mathrm{d}t\right] \tag{4.30}$$

$$+ \mathbb{E}\left[\int_0^{\mathfrak{h}} \theta^\epsilon(t, P_t, E_t)\partial_e E_t \,\mathrm{d}t\right]. \tag{4.31}$$

Since $V_{\mathfrak{h}} = 0$ and $\partial_e w \geqslant 0$, $\partial_e \mathcal{V}^\epsilon \geqslant 0$, $\partial_e E_t \geqslant 0$, we obtain

$$\int_{\mathbb{R}} |V_0^e| \,\mathrm{d}e \leqslant \int_0^{\mathfrak{h}} \mathbb{E}\left[\int_{\mathbb{R}} |V_t^e \partial_y \mu(P_t, Y_t)| \partial_e[w(t, P_t, E_t^e)] \,\mathrm{d}e\right] \mathrm{d}t \tag{4.32}$$

$$+\int_0^{\mathfrak{h}} \mathbb{E}\Big[\int_{\mathbb{R}} |c_t V_t^e| \partial_e[\mathcal{V}^\epsilon(t,P_t,E_t^e)]\,\mathrm{d}e\Big]\,\mathrm{d}t$$
$$+\int_0^{\mathfrak{h}} \mathbb{E}\Big[\int_{\mathbb{R}} |\theta^\epsilon(t,P_t,E_t)| \partial_e E_t\,\mathrm{d}e\Big]\,\mathrm{d}t.$$

We obtain, using Lemma 4.3 and the bound on $\|\partial_y\mu\|_\infty$,

$$\int_{\mathbb{R}} |V_t^e \partial_y\mu(P_t,Y_t)| \partial_e[w(t,P_t,E_t^e)]\,\mathrm{d}e \leqslant C\epsilon \int_{\mathbb{R}} \partial_e[w(t,P_t,E_t^e)]\,\mathrm{d}e, \\ \leqslant C\epsilon, \tag{4.33}$$

where we used the fact that w is bounded in the last inequality. Similar arguments, since c is bounded, lead to

$$\int_{\mathbb{R}} |c_t V_t^e| \partial_e[\mathcal{V}^\epsilon(t,P_t,E_t^e)]\,\mathrm{d}e \leqslant C\epsilon. \tag{4.34}$$

From (2.41), we know that

$$|\theta^\epsilon(t,P_t,E_t)| \leqslant C\epsilon\bigg(\int_{\mathbb{R}^d} |\partial_p\mathcal{V}(t,P_t+q,E_t^e)|\varphi_\epsilon(q)\,\mathrm{d}q + \partial_e\mathcal{V}^\epsilon(t,P_t,E_t)\bigg). \tag{4.35}$$

We compute, using the change of variable $e \mapsto \zeta = E_t^e$

$$\int_{\mathbb{R}}\int_{\mathbb{R}^d} |\partial_p\mathcal{V}(t,P_t+q,E_t^e)|\varphi_\epsilon(q)\,\mathrm{d}q\partial_e E_t^e\,\mathrm{d}e = \int_{\mathbb{R}}\int_{\mathbb{R}^d} |\partial_p\mathcal{V}(t,P_t+q,\zeta)|\varphi_\epsilon(q)\,\mathrm{d}q\,\mathrm{d}\zeta, \tag{4.36}$$
$$\leqslant C, \tag{4.37}$$

where we used the L^1-bound on $\partial_p V$ given in (2.23). We then compute

$$\mathbb{E}\Big[\int_{\mathbb{R}} |\theta^\epsilon(t,P_t,E_t)| \partial_e E_t\,\mathrm{d}e\Big] \leqslant C\epsilon(1+\mathbb{E}\Big[\int_{\mathbb{R}} \partial_e\mathcal{V}^\epsilon(t,P_t,E_t)\partial_e E_t\,\mathrm{d}e\Big]), \tag{4.38}$$
$$\leqslant C\epsilon. \tag{4.39}$$

The proof is concluded by combining the previous inequality, the estimates given in (4.33) and (4.34) with (4.32). □

Combining Lemma 4.2 with the result Proposition 4.2, we obtain straightforwardly the following.

Corollary 4.1. *Let Assumption 2.1 hold. Then, for $n \leqslant N-2$, $p \in \mathbb{R}^d$,*

$$\mathcal{E}_n^S(p) \leqslant C\big(1+|p|^2\big)\mathfrak{h}^{\frac{3}{2}} + C\epsilon\mathfrak{h}. \tag{4.40}$$

4.4. Diffusion error

We now study the term given in (3.31) and we straightforwardly observe

$$\mathcal{E}_n^D(p) = \int_{\mathbb{R}} \Big|\tilde{\mathcal{T}}_h\Big(p,\mathbb{E}\Big[\mathcal{V}^\epsilon\Big(t_{n+1},\widehat{P}_{t_{n+1}}^{t_n,p},\cdot\Big)\Big]\Big)(e) - \tilde{\mathcal{T}}_h(p,\mathbb{E}[\mathcal{V}^\epsilon(t_{n+1},P_{t_{n+1}}^{t_n,p},\cdot)])(e)\Big|\,\mathrm{d}e, \tag{4.41}$$
$$\leqslant \int_{\mathbb{R}} \Big|\mathbb{E}\Big[\mathcal{V}^\epsilon\Big(t_{n+1},\widehat{P}_{t_{n+1}}^{t_n,p},e\Big)\Big] - \mathbb{E}\Big[\mathcal{V}^\epsilon\Big(t_{n+1},P_{t_{n+1}}^{t_n,p},e\Big)\Big]\Big|\,\mathrm{d}e, \tag{4.42}$$

where we use the L^1-stability of $\tilde{\mathcal{T}}_h$.

This motivates the introduction of the following auxiliary result.

Proposition 4.3. *Let Assumption 2.1 hold. For $0 \leqslant n \leqslant N-1$, let w_n be the solution on $[t_n, t_{n+1}] \times \mathbb{R}^d$ to the following PDE:*

$$\partial_t w + b(p)\partial_p w + \frac{1}{2}\bar{\sigma}^2 \Delta_{pp} w = 0 \text{ and } w(t_{n+1}, \cdot) := \mathcal{V}^\epsilon(t_{n+1}, \cdot, e), \tag{4.43}$$

for a fixed $e \in \mathbb{R}$. Then, under Assumption 2.1, the followings hold, for $q, q' \in \mathbb{R}^d$ and $t \in [0, T)$,

$$\int_{\mathbb{R}} \left|\partial_{p^\ell} w_n(t,q,e) - \partial_{p^\ell} w_n(t,q',e)\right| \mathrm{d}e \leqslant \frac{C}{\epsilon}|q-q'| \tag{4.44}$$

and

$$\int_{\mathbb{R}} \left|\partial^2_{p^\ell p^{\ell'}} w_n(t,q,e) - \partial^2_{p^\ell p^{\ell'}} w_n(t,q',e)\right| \mathrm{d}e \leqslant \frac{C}{\epsilon^2}|q-q'|, \tag{4.45}$$

for all $\ell, \ell' \in \{1, \ldots, d\}$. Moreover,

$$\int |\partial_p w_n(t,p,e)| \,\mathrm{d}e \leqslant C. \tag{4.46}$$

Proof. For ease of presentation, the proof is done in the one-dimensional case.

By Theorem 2.3.5 in Zhang [13], we have the following expressions for the first and second order derivative with respect to vector p, for $(t,p,e) \in [t_n, t_{n+1}] \times \mathbb{R} \times \mathbb{R}$,

$$\partial_p w_n(t,p,e) = \mathbb{E}\Big[\partial_p \mathcal{V}^\epsilon\Big(t_{n+1}, P^{t,p}_{t_{n+1}}, e\Big)\partial_p P^{t,p}_{t_{n+1}}\Big], \tag{4.47}$$

$$\partial^2_{pp} w_n(t,p,e) = \mathbb{E}\Big[\partial^2_{pp}\mathcal{V}^\epsilon\Big(t_{n+1}, P^{t,p}_{t_{n+1}}, e\Big)\Big(\partial_p P^{t,p}_{t_{n+1}}\Big)^2 + \partial_p \mathcal{V}^\epsilon\Big(t_{n+1}, P^{t,p}_{t_{n+1}}, e\Big)\partial^2_{pp} P^{t,p}_{t_{n+1}}\Big], \tag{4.48}$$

where, for $t \leqslant s \leqslant t_{n+1}$,

$$P^{t,p}_s = p + \int_t^s b\big(P^{t,p}_r\big)\,\mathrm{d}r + \bar{\sigma}(W_s - W_t), \tag{4.49}$$

$$\partial_p P^{t,p}_s = 1 + \int_t^s b'\big(P^{t,p}_r\big)\partial_p P^{t,p}_r \,\mathrm{d}r, \tag{4.50}$$

$$\partial^2_{pp} P^{t,p}_s = \int_t^s \Big(b''(P_r)\big|\partial_p P^{t,p}_r\big|^2 + b'(P_r)\partial^2_{pp} P^{t,p}_r\Big)\,\mathrm{d}r. \tag{4.51}$$

In this setting, classical arguments lead to,

$$\mathbb{E}\left[\sup_{s\in[t,t_{n+1}]}\big|\partial_p P^{t,p}_s\big|^\kappa + \sup_{s\in[t,t_{n+1}]}\big|\partial^2_{pp} P^{t,p}_s\big|^\kappa\right] \leqslant C_\kappa, \tag{4.52}$$

and for $q, q' \in \mathbb{R}$, $s \in [t, T]$, $\kappa \geqslant 1$,

$$\mathbb{E}\Big[\Big|P^{t,q}_s - P^{t,q'}_s\Big|^\kappa + \Big|\partial_p P^{t,q}_s - \partial_p P^{t,q'}_s\Big|^\kappa + \Big|\partial^2_{pp} P^{t,q}_s - \partial^2_{pp} P^{t,q'}_s\Big|^\kappa\Big] \leqslant C|q-q'|^\kappa. \tag{4.53}$$

We now study the second order derivative given in (4.48).

We compute, for $(t,q,q') \in [t_n, t_{n+1}] \times \mathbb{R} \times \mathbb{R}$,

$$\int_{\mathbb{R}} \left|\partial^2_{pp} w_n(t,q,e) - \partial^2_{pp} w_n(t,q',e)\right| \mathrm{d}e \tag{4.54}$$

$$\leqslant \mathbb{E}\Big[\int_{\mathbb{R}}\Big|\partial^2_{pp}\mathcal{V}^\epsilon\Big(t_{n+1},P^{t,q}_{t_{n+1}},e\Big)\Big(\partial_p P^{t,q}_{t_{n+1}}\Big)^2-\partial^2_{pp}\mathcal{V}^\epsilon\Big(t_{n+1},P^{t,q'}_{t_{n+1}},e\Big)\Big(\partial_p P^{t,q'}_{t_{n+1}}\Big)^2\Big|\,\mathrm{d}e\Big]:=B_1 \tag{4.55}$$

$$+\,\mathbb{E}\Big[\int_{\mathbb{R}}\Big|\partial_p\mathcal{V}^\epsilon\Big(t_{n+1},P^{t,q}_{t_{n+1}},e\Big)\partial^2_{pp}P^{t,q}_{t_{n+1}}-\partial_p\mathcal{V}^\epsilon\Big(t_{n+1},P^{t,q'}_{t_{n+1}},e\Big)\partial^2_{pp}P^{t,q'}_{t_{n+1}}\Big|\,\mathrm{d}e\Big]:=B_2. \tag{4.56}$$

For the first term B_1,

$$B_1\leqslant \mathbb{E}\Big[\int_{\mathbb{R}}\Big|\partial^2_{pp}\mathcal{V}^\epsilon\Big(t_{n+1},P^{t,q}_{t_{n+1}},e\Big)\Big(\partial_p P^{t,q}_{t_{n+1}}\Big)^2-\partial^2_{pp}\mathcal{V}^\epsilon\Big(t_{n+1},P^{t,q}_{t_{n+1}},e\Big)\Big(\partial_p P^{t,q'}_{t_{n+1}}\Big)^2\Big|\,\mathrm{d}e\Big]$$
$$+\,\mathbb{E}\Big[\int_{\mathbb{R}}\Big|\partial^2_{pp}\mathcal{V}^\epsilon\Big(t_{n+1},P^{t,q}_{t_{n+1}},e\Big)\Big(\partial_p P^{t,q'}_{t_{n+1}}\Big)^2-\partial^2_{pp}\mathcal{V}^\epsilon\Big(t_{n+1},P^{t,q'}_{t_{n+1}},e\Big)\Big(\partial_p P^{t,q'}_{t_{n+1}}\Big)^2\Big|\,\mathrm{d}e\Big].$$

We then compute

$$\mathbb{E}\Big[\int_{\mathbb{R}}\Big|\partial^2_{pp}\mathcal{V}^\epsilon\Big(t_{n+1},P^{t,q}_{t_{n+1}},e\Big)\Big(\partial_p P^{t,q}_{t_{n+1}}\Big)^2-\partial^2_{pp}\mathcal{V}^\epsilon\Big(t_{n+1},P^{t,q}_{t_{n+1}},e\Big)\Big(\partial_p P^{t,q'}_{t_{n+1}}\Big)^2\Big|\,\mathrm{d}e\Big], \tag{4.57}$$

$$\leqslant \mathbb{E}\Big[\Big|\Big(\partial_p P^{t,q}_{t_{n+1}}\Big)^2-\Big(\partial_p P^{t,q'}_{t_{n+1}}\Big)^2\Big|\int_{\mathbb{R}}\Big|\partial^2_{pp}\mathcal{V}^\epsilon\Big(t_{n+1},P^{t,q}_{t_{n+1}},e\Big)\Big|\,\mathrm{d}e\Big], \tag{4.58}$$

$$\leqslant \frac{C}{\epsilon}\mathbb{E}\Big[\Big|\Big(\partial_p P^{t,q}_{t_{n+1}}\Big)^2-\Big(\partial_p P^{t,q'}_{t_{n+1}}\Big)^2\Big|\Big], \tag{4.59}$$

where we used for the last inequality Lemma 2.1. Observing that

$$\Big|\Big(\partial_p P^{t,q}_{t_{n+1}}\Big)^2-\Big(\partial_p P^{t,q'}_{t_{n+1}}\Big)^2\Big|\leqslant\Big|\partial_p P^{t,q}_{t_{n+1}}+\partial_p P^{t,q'}_{t_{n+1}}\Big|\cdot\Big|\partial_p P^{t,q}_{t_{n+1}}-\partial_p P^{t,q'}_{t_{n+1}}\Big|,$$

we combine Cauchy–Schwarz inequality with (4.53) to obtain

$$\mathbb{E}\Big[\Big|\Big(\partial_p P^{t,q}_{t_{n+1}}\Big)^2-\Big(\partial_p P^{t,q'}_{t_{n+1}}\Big)^2\Big|\Big]\leqslant C|q-q'|. \tag{4.60}$$

Combining the previous inequality with (4.59), we get

$$\mathbb{E}\Big[\int_{\mathbb{R}}\Big|\partial^2_{pp}\mathcal{V}^\epsilon\Big(t_{n+1},P^{t,q}_{t_{n+1}},e\Big)\Big(\partial_p P^{t,q}_{t_{n+1}}\Big)^2-\partial^2_{pp}\mathcal{V}^\epsilon\Big(t_{n+1},P^{t,q}_{t_{n+1}},e\Big)\Big(\partial_p P^{t,q'}_{t_{n+1}}\Big)^2\Big|\,\mathrm{d}e\Big]\leqslant\frac{C}{\epsilon}|q-q'|. \tag{4.61}$$

We also compute

$$\mathbb{E}\Big[\int_{\mathbb{R}}\Big|\partial^2_{pp}\mathcal{V}^\epsilon\Big(t_{n+1},P^{t,q}_{t_{n+1}},e\Big)\Big(\partial_p P^{t,q'}_{t_{n+1}}\Big)^2-\partial^2_{pp}\mathcal{V}^\epsilon\Big(t_{n+1},P^{t,q'}_{t_{n+1}},e\Big)\Big(\partial_p P^{t,q'}_{t_{n+1}}\Big)^2\Big|\,\mathrm{d}e\Big],$$
$$\leqslant \mathbb{E}\Big[\Big(\partial_p P^{t,q'}_{t_{n+1}}\Big)^2\int_{\mathbb{R}}\Big|\partial^2_{pp}\mathcal{V}^\epsilon\Big(t_{n+1},P^{t,q}_{t_{n+1}},e\Big)-\partial^2_{pp}\mathcal{V}^\epsilon\Big(t_{n+1},P^{t,q'}_{t_{n+1}},e\Big)\Big|\,\mathrm{d}e\Big],$$
$$\leqslant\frac{C}{\epsilon^2}|q-q'|,$$

where we used Lemma 2.1 and (4.52), for the last inequality. Combining (4.61) with the previous inequality, we finally obtain that

$$B_1\leqslant\frac{C}{\epsilon^2}|q-q'|+\frac{C}{\epsilon}|q-q'|. \tag{4.62}$$

For the term B_2 given in (4.56), we observe

$$B_2\leqslant \mathbb{E}\Big[\int_{\mathbb{R}}\Big|\partial_p\mathcal{V}^\epsilon\Big(t_{n+1},P^{t,q}_{t_{n+1}},e\Big)\partial^2_{pp}P^{t,q}_{t_{n+1}}-\partial_p\mathcal{V}^\epsilon\Big(t_{n+1},P^{t,q}_{t_{n+1}},e\Big)\partial^2_{pp}P^{t,q'}_{t_{n+1}}\Big|\,\mathrm{d}e\Big]$$

$$+\,\mathbb{E}\Big[\int_{\mathbb{R}}\Big|\partial_p\mathcal{V}^\epsilon\Big(t_{n+1},P^{t,q}_{t_{n+1}},e\Big)\partial^2_{pp}P^{t,q'}_{t_{n+1}}-\partial_p\mathcal{V}^\epsilon\Big(t_{n+1},P^{t,q'}_{t_{n+1}},e\Big)\partial^2_{pp}P^{t,q'}_{t_{n+1}}\Big|\,\mathrm{d}e\Big].$$

Combining Lemma 2.1, the estimates given in (4.53) and (4.52), with similar computations as done previously, we obtain

$$B_2\leqslant\frac{C}{\epsilon}|q-q'|+C|q-q'|. \tag{4.63}$$

Inserting (4.62) and (4.63) back into (4.56), we obtain the

$$\int_{\mathbb{R}}\big|\partial^2_{pp}w_n(t,q,e)-\partial^2_{pp}w_n(t,q',e)\big|\,\mathrm{d}e\leqslant\frac{C}{\epsilon}|q-q'|, \tag{4.64}$$

as $\epsilon\in(0,1)$.

Similar arguments allow to retrieve the estimates for the first order derivatives. □

Proposition 4.4. *Under Assumption 2.1, the following holds*

$$\mathcal{E}^D_n(p)\leqslant C\frac{\mathfrak{h}^{\frac{3}{2}}}{\epsilon^2},$$

for $0\leqslant n\leqslant N-1$ *and* $p\in\mathbb{R}^d$.

Proof. For ease of presentation, we do the proof in dimension $d=1$.

(1) We first recall Definition 3.1 and (3.6). In this context, we introduce two new discrete processes:

$$\widehat{P}^{t_n,p}_t=p+b(p)(t-t_n)+\bar\sigma\frac{\sqrt{t-t_n}}{\sqrt{\mathfrak{h}}}\Delta\widehat{W}_n,$$
$$\bar{P}^{t_n,p,\lambda}_t=p+b(p)(t-t_n)+\lambda\bar\sigma\frac{\sqrt{t-t_n}}{\sqrt{\mathfrak{h}}}\Delta\widehat{W}_n,$$

for $p\in\mathbb{R}$, $t\in[t_n,t_{n+1}]$, $\lambda\in[0,1]$ and $0\leqslant n\leqslant N-1$.
Now, recalling (4.42), we first observe

$$\mathbb{E}\Big[\mathcal{V}^\epsilon\Big(t_{n+1},\widehat{P}^{t_n,p}_{t_{n+1}},e\Big)\Big]-\mathbb{E}\Big[\mathcal{V}^\epsilon\Big(t_{n+1},P^{t_n,p}_{t_{n+1}},e\Big)\Big]=\mathbb{E}\Big[w_n\Big(t_{n+1},\widehat{P}^{t_n,p}_{t_{n+1}},e\Big)\Big]-w_n(t_n,p,e). \tag{4.65}$$

We apply the discrete Ito formula given of Proposition 14 in [7] to our much simpler framework, to obtain

$$\mathbb{E}\Big[w_n\Big(t_{n+1},\widehat{P}^{t_n,p}_{t_{n+1}},e\Big)\Big]-w_n(t_n,p,e)=\int_{t_n}^{t_{n+1}}\mathbb{E}\Big[\partial_t w_n\Big(t,\widehat{P}^{t_n,p}_t,e\Big)+b(p)\partial_p w_n\Big(t,\widehat{P}^{t_n,p}_t,e\Big)\Big]\,\mathrm{d}t$$
$$+\frac{\bar\sigma^2}{2}\mathbb{E}\Big[\int_{t_n}^{t_{n+1}}\int_0^1\partial^2_{pp}w_n\Big(t,\bar{P}^{t_n,p,\lambda}_t,e\Big)\,\mathrm{d}\lambda\,\mathrm{d}t\Big].$$

Using the PDE (4.43) satisfied by w_n, the previous equality leads to

$$\mathbb{E}\Big[w_n\Big(t_{n+1},\widehat{P}^{t_n,p}_{t_{n+1}},e\Big)\Big]-w_n(t_n,p,e)$$
$$=\int_{t_n}^{t_{n+1}}\mathbb{E}\Big[\Big(b(p)-b\Big(\widehat{P}^{t_n,p}_t\Big)\Big)\partial_p w_n\Big(t,\widehat{P}^{t_n,p}_t,e\Big)\Big]\,\mathrm{d}t \tag{4.66}$$
$$+\frac{\bar\sigma^2}{2}\mathbb{E}\Big[\int_{t_n}^{t_{n+1}}\int_0^1\Big[\partial^2_{pp}w_n\Big(t,\bar{P}^{t_n,p,\lambda}_t,e\Big)-\partial^2_{pp}w_n\Big(t,\widehat{P}^{t_n,p}_t,e\Big)\Big]\,\mathrm{d}\lambda\,\mathrm{d}t\Big]. \tag{4.67}$$

(2) For $e \in \mathbb{R}$, we denote by $A^1(e)$, the term in (4.66) and by $A^2(e)$, the term in (4.67).
We compute

$$\begin{aligned}\left|A^1(e)\right| &\leqslant \int_{t_n}^{t_{n+1}} \mathbb{E}\left[\left|\left(b(p) - b\left(\widehat{P}_t^{t_n,p}\right)\right)\right|\left|\partial_p w_n\left(t, \widehat{P}_t^{t_n,p}, e\right)\right|\right] \mathrm{d}t, \\ &\leqslant C\mathfrak{h}^{\frac{1}{2}} \int_{t_n}^{t_{n+1}} \mathbb{E}\left[\left|\partial_p w_n\left(t, \widehat{P}_t^{t_n,p}, e\right)\right|\right] \mathrm{d}t,\end{aligned}$$

and thus

$$\begin{aligned}\int_{\mathbb{R}} \left|A^1(e)\right| \mathrm{d}e &\leqslant C\mathfrak{h}^{\frac{1}{2}} \int_{t_n}^{t_{n+1}} \mathbb{E}\left[\int_{\mathbb{R}} \left|\partial_p w_n\left(t, \widehat{P}_t^{t_n,p}, e\right)\right| \mathrm{d}e\right] \mathrm{d}t, \\ &\leqslant C\mathfrak{h}^{\frac{3}{2}}, \end{aligned} \tag{4.68}$$

where we used (4.46) for the last inequality.
Using Proposition 4.3, we compute

$$\begin{aligned}\int_{\mathbb{R}} \left|\partial^2_{pp} w_n\left(t, \bar{P}_t^{t_n,p,\lambda}, e\right) - \partial^2_{pp} w_n\left(t, \widehat{P}_t^{t_n,p}, e\right)\right| \mathrm{d}e &\leqslant \frac{C}{\epsilon^2}\left|\bar{P}_t^{t_n,p,\lambda} - \widehat{P}_t^{t_n,p}\right|, \\ &\leqslant \frac{C}{\epsilon^2}\sqrt{\mathfrak{h}}.\end{aligned}$$

This yields

$$\int_{\mathbb{R}} |A_2(e)| \,\mathrm{d}e \leqslant \frac{C}{\epsilon^2}\mathfrak{h}^{\frac{3}{2}}. \tag{4.69}$$

Combining the previous inequality with (4.68) and (4.66), (4.67), we get

$$\int_{\mathbb{R}} \left|\mathbb{E}\left[w_n\left(t_{n+1}, \widehat{P}_{t_{n+1}}^{t_n,p}, e\right)\right] - w_n(t_n, p, e)\right| \mathrm{d}e \leqslant \frac{C}{\epsilon^2}\mathfrak{h}^{\frac{3}{2}}.$$

The proof is concluded recalling (4.65) and (4.42).

□

4.5. Proof of Theorem 3.1

The proof of our main result is now almost straightforward. From Lemma 4.1, Corollary 4.1 and Proposition 4.4, we observe that

$$\mathbb{E}\left[\mathcal{E}_n^{\mathrm{T}}\left(\widehat{P}_{t_n}\right)\right] \leqslant C\frac{\mathfrak{h}}{M},\; \mathbb{E}\left[\mathcal{E}_n^{D}\left(\widehat{P}_{t_n}\right)\right] \leqslant C\frac{\mathfrak{h}^{\frac{3}{2}}}{\epsilon^2} \;\text{ and }\; \mathbb{E}\left[\mathcal{E}_n^{S}\left(\widehat{P}_{t_n}\right)\right] \leqslant C\mathfrak{h}\left(\mathfrak{h}^{\frac{1}{2}} + \epsilon\right). \tag{4.70}$$

Summing over n, we get

$$\sum_{n=0}^{N-2} \mathbb{E}\left[\mathcal{E}_n^{\mathrm{T}}\left(\widehat{P}_{t_n}\right) + \mathcal{E}_n^{D}\left(\widehat{P}_{t_n}\right) + \mathcal{E}_n^{S}\left(\widehat{P}_{t_n}\right)\right] \leqslant C\left(\frac{1}{M} + \frac{\sqrt{\mathfrak{h}}}{\epsilon^2} + \epsilon\right). \tag{4.71}$$

For the last step, we obtain

$$\mathbb{E}\left[\mathcal{E}_{N-1}\left(\widehat{P}_{t_{N-1}}\right)\right] \leqslant C\left(\frac{h}{M} + \epsilon + \mathfrak{h}\left(\mathfrak{h}^{\frac{1}{2}} + \epsilon\right)\right) \tag{4.72}$$

combining (3.34) with Proposition 4.1 and Proposition 2.1. The previous inequality combined with (4.71) and Proposition 3.1 yields (3.40).

Balancing optimally the errors in (3.40) allows to conclude by proving (3.41). □

5. Numerics

In this section, we realize numerical experiments to test in practice the schemes introduced in Section 3.

5.1. Setting

We first introduce the models we will use. They have already been considered in [6], which facilitates the comparison with previous numerical results.

Let us first define the following toy model where the process P corresponds to a Brownian motion and is a multidimensional version of the model given in Remark 2.2:

Example 5.1 (Linear model)**.**

$$\mathrm{d}P_t = \sigma\,\mathrm{d}W_t, \mathrm{d}E_t = \left(\frac{1}{\sqrt{d}}\sum_{\ell=1}^{d} P_t^\ell - Y_t\right)\mathrm{d}t \tag{5.1}$$

with terminal function $(p,e) \mapsto \phi(p,e) = \mathbf{1}_{\{e \geqslant 0\}}$ and where W is a d-dimensional Brownian motion and $\sigma > 0$.

We will also consider a multiplicative model:

Example 5.2 (Multiplicative model)**.** Let W be a d-dimensional Brownian motion. For all $\ell \in \{1,\dots,d\}$, we set

$$\mathrm{d}P_t^\ell = \mu P_t^\ell\,\mathrm{d}t + \sigma P_t^\ell\,\mathrm{d}W_t^\ell,\ P_0^\ell = 1,\ \text{and}\ \mathrm{d}E_t = \tilde{\mu}(P_t, Y_t)\,\mathrm{d}t \tag{5.2}$$

with $(p,y) \mapsto \tilde{\mu}(p,y) = \left(\prod_{\ell=1}^{d} p^\ell\right)^{\frac{1}{\sqrt{d}}} e^{-\theta y}$, for some $\theta > 0$. The terminal condition is given by $(p,e) \mapsto \phi(p,e) = \mathbf{1}_{\{e \geqslant 0\}}$.

As mentioned in [6], one advantage of these models is that they can be reduced to models with lower dimension (at most 2) whatever the value of d. For comparison, we will also use some numerical *proxy*: for the first model in Example 5.1, we use the method considered in [1] applied to the one dimensional reduction of the model, for the second model in Example 5.2 we use the NN & Upwind scheme. Both methods are introduced and discussed in [6].

We also observe that the above models fit into the setting of Section 3.1.2, recall Assumption 3.1. As shown below and as expected, they are thus numerically tractable. We shall use in this regard the scheme introduced in Definition 3.2, linked to CASE 1. We will also consider CASE 2, where the number of particles is capped. To this end, we need to define precisely the operator $\mathfrak{R}^{M,m}$ appearing in Definition 3.3, which is the purpose of the next section.

5.2. Definition of $\mathfrak{R}$

There are various possible implementations for the operator $\mathfrak{R}$ whose goal is to tame the computational cost, recall Definition 3.3. Using this operator $\mathfrak{R}$ introduces another numerical error. In view of the convergence study in the previous sections that focuses on L^1-norm, one could aim to optimise this error measurement also in the implementation of $\mathfrak{R}$. Namely, with the notations of Definition 3.3 (step 2), this means that we would like to have, for each $0 \leqslant n \leqslant N-1$, $p \in \mathbb{R}^d$,

$$\int_{\mathbb{R}} \left|\check{v}_n^{N,M}(p,e) - \bar{v}_n^{N,M}(p,e)\right| \mathrm{d}e \tag{5.3}$$

minimal. The solution to this problem is known, see Section 2 of [11]. Indeed, $\check{v}_n^{N,M}(p,\cdot)$ and $\bar{v}_n^{N,M}(p,\cdot)$ are CDF of probability measures on $\mathbb{R}$, respectively denoted $\hat{\mu}^{N,M}(p)$ and $\bar{\mu}^{N,M}(p)$ for later use. It turns out that minimising (5.3) corresponds exactly to minimise the $\mathcal{W}_1$-distance between the two distributions. Let us recall the general approximation result stated in Section 2 of [11].

Lemma 5.1. *Let F be the CDF associated to a probability distribution $\nu \in \mathcal{P}_1(\mathbb{R})$. We note F^{-1} its generalized inverse, given by*

$$F^{-1}(v) = \inf\{y \in \mathbb{R} | F(y) \geqslant v\},\ v \in (0,1). \tag{5.4}$$

Let $\hat{\nu}_M = \frac{1}{M}\sum_{m=1}^{M} \delta_{e_m}$, then the minimum of

$$\mathcal{W}_1(\mu, \hat{\nu}_M) = \int_{\mathbb{R}} |H * \nu(x) - H * \hat{\nu}_M(x)| \,\mathrm{d}x$$

is achieved for $e_m = F^{-1}(\frac{2m-1}{2M})$, $1 \leqslant m \leqslant M$.

In our framework, the above result can be specialized since the starting distribution $\bar{\mu}^{N,M}(p)$ is already atomic (with $2dM$ atoms).

Corollary 5.1. *For $\xi \in \mathcal{D}_{2dM}$, let $\nu := \frac{1}{2dM}\sum_{m=1}^{2dM} \delta_{\xi_m}$. Then the optimal $\hat{\nu}_M$, recall Lemma 5.1 is given by*

$$\hat{\nu}_M = \frac{1}{M}\sum_{m=1}^{M} \delta_{e_m} \text{ with } \ e_m = \xi_{(2m-1)d}. \tag{5.5}$$

Proof. According to Lemma 5.1 and the definition of F^{-1} in (5.4), one should find, for $1 \leqslant m \leqslant M$,

$$i^\star(m) := \min\left\{ i \in \{1, \ldots, M\} \,|\, F(\xi_i) \geqslant \frac{2m-1}{2M} \right\}. \tag{5.6}$$

Observing that $F(\xi_i) = \frac{i}{2dM}$, we obtain that $i^\star(m) = (2m-1)d$, which concludes the proof of the corollary. $\square$

Definition 5.1 (Operator $\mathfrak{R}^{M,m}$)**.** For $M \geqslant m \geqslant 1$, we first denote S^M the operator of sorting M particles in ascending order, namely

$$\mathbb{R}^M \ni \xi \mapsto S^M(\xi) \in \mathcal{D}_M. \tag{5.7}$$

At each step n, for a set of $2dM$ particles namely $e \in \mathbb{R}^{2dM}$, $S^{2dM}(e)$ could be written as follows:

$$S^{2dM}(\xi) = \left(\widetilde{S}^{2d,1}(\xi), \cdots, \widetilde{S}^{2d,M}(\xi)\right), \tag{5.8}$$

where for $1 \leqslant i \leqslant M, \widetilde{S}^{2d,i}(\xi) := \left(S_j^{2dM}(\xi)\right)_{2d(i-1)<j\leqslant 2di} \in \mathcal{D}_{2d}$, with $S_i^{2dM}(\xi)$ denoting the i-th coordinate of vector $S^{2dM}(\xi)$. For $\psi^\xi \in \mathcal{I}^{2dM}$ associated to $\xi \in \mathcal{D}_{2dM}$, we recall that

$$\psi^\xi = H * \left(\frac{1}{2dM}\sum_{m=1}^{2dM} \delta_{S_m^{2dM}(\xi)}\right),$$

we then define the following operator

$$\psi^\xi \ni \mathcal{I}^{2dM} \mapsto \mathfrak{R}^{2dM,M}\left(\psi^\xi\right) := H * \left(\frac{1}{M}\sum_{m=1}^{M} \delta_{\tilde{e}_m}\right) \in \mathcal{I}^M, \tag{5.9}$$

where four different choices for $\tilde{e}_m$ are considered, leading to four different implementations. Namely, for each $1 \leqslant m \leqslant M$,

TABLE 1. $L1$-error and $L\infty$-error for model Example 5.1 with different numbers of particles with respect to `leftmost,mean,rightmost,optim`, note that the time steps $N = 20$.

Sigma	Number of particles	Method	$L1$-error	$L\infty$-error
1.0	1000	`leftmost`	5.71e−3	1.67e−2
		`mean`	4.96e−3	1.50e−2
		`rightmost`	5.39e−3	1.80e−2
		`optim`	3.72e−3	1.02e−2
	5000	`leftmost`	5.05e−3	1.45e−2
		`mean`	4.92e−3	1.46e−2
		`rightmost`	4.92e−3	1.53e−2
		`optim`	3.47e−3	1.01e−2
0.3	1000	`leftmost`	1.34e−3	0.85e−2
		`mean`	0.88e−3	0.61e−2
		`rightmost`	1.25e−3	0.91e−2
		`optim`	1.19e−3	0.69e−2
	5000	`leftmost`	0.94e−3	0.68e−2
		`mean`	0.88e−3	0.61e−2
		`rightmost`	0.92e−3	0.65e−2
		`optim`	0.56e−3	0.46e−2
0.01	1000	`leftmost`	0.27e−3	0.39e−2
		`mean`	0.13e−3	0.20e−2
		`rightmost`	0.48e−3	0.43e−2
		`optim`	0.11e−3	0.17e−2
	5000	`leftmost`	0.09e−3	0.13e−2
		`mean`	0.08e−3	0.17e−2
		`rightmost`	0.13e−3	0.18e−2
		`optim`	0.06e−3	0.11e−2

`mean`: The mean position of $2d$ particles of $\widetilde{S}^{2d,m}(\xi)$, namely

$$\tilde{e}_m := \frac{1}{2d} \sum_{j=2d(m-1)+1}^{2dm} S_j^{2dM}(\xi). \tag{5.10}$$

`leftmost`: The minimum position among the $2d$ particles of $\widetilde{S}^{2d,m}(\xi)$ is kept,

$$\tilde{e}_m := \min\left\{S_j^{2dM}(\xi), 2d(m-1)+1 \leqslant j \leqslant 2dm\right\} = S_{2d(m-1)+1}^{2dM}(\xi). \tag{5.11}$$

`rightmost`: The maximum position among the $2d$ particles of $\widetilde{S}^{2d,m}(\xi)$ is kept,

$$\tilde{e}_m := \max\left\{S_j^{2dM}(\xi), 2d(m-1)+1 \leqslant j \leqslant 2dm\right\} = S_{2dm}^{2dM}(\xi). \tag{5.12}$$

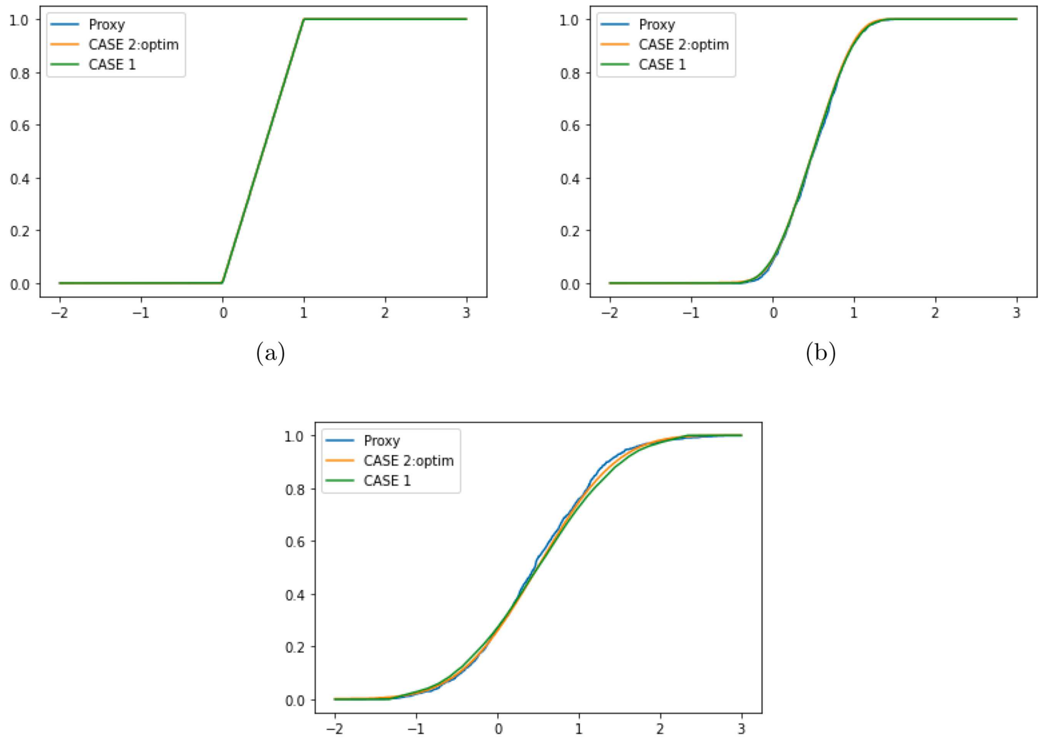
`optim`: The optimal position according to Corollary 5.1 is kept, $\tilde{e}_m := S_{d(2m-1)}^{2dM}(\xi)$.

5.3. Numerical results

First of all, to validate empirically our different implementations of operator $\mathfrak{R}$, we report the L^1-error and L^∞-error against the “proxy” solution for different numbers of particles, based on the different strategies to reduce the number of particles, see Definition 5.1 for $\sigma = 0.01, 0.3, 1.0$, see Table 1. We observe that the four implementations of $\mathfrak{R}$: `leftmost`, `mean`, `rightmost`, `optim` of positions of particles, yield practically the same

TABLE 2. Computational cost in Example 5.1 for different dimension d (for the P-variable) for CASE 1 and CASE 2 schemes.

Dimension	$d = 1$	$d = 4$
Time for CASE 1	46.96 s	784 s
Time for CASE 2	0.3 s	1.3 s



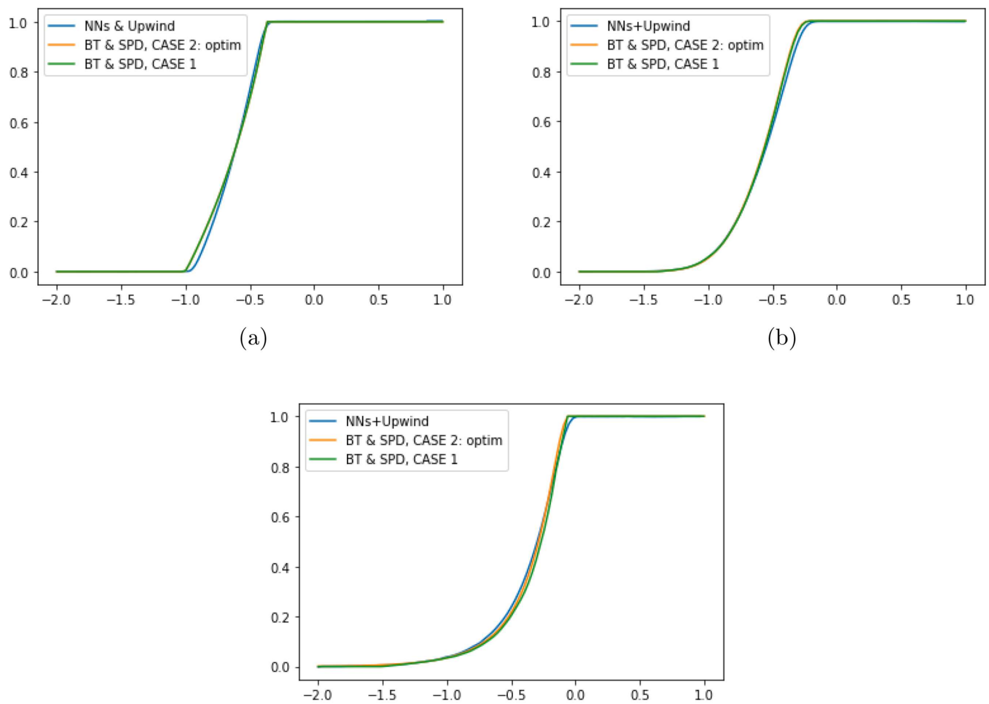
(a) (b) (c)

FIGURE 1. Model of Example 5.1: Comparison of the two methods CASE 1 & CASE2 with $d = 4$. The Proxy solution is given by the same particle method on the one-dimensional PDE. For $BT\&SPD$: both CASE 1 and CASE 2, the number of particles is $M = 3500$ and the number of time steps $N = 20$. (a) $\sigma = 0.01$. (b) $\sigma = 0.3$. (c) $\sigma = 1.0$.

numerical results for different levels of volatilities. The `optim` strategy appears to be better, as expected. Hence in the following, we only consider the CASE 2 implemented with operator $\mathfrak{R}$ `optim`. The clear advantage of considering the CASE 2 scheme is that the number of particles does not grow exponentially during the iterations, hence it takes less computational time compared to CASE 1, see Table 2.

To validate empirically our numerical schemes, we plot the value function $\mathcal{V}$ of model Example 5.1 using scheme CASE 1 and CASE 2 against the *proxy* solution presented in [1], see Figure 1. As expected, both CASE 1 and CASE 2 scheme could reproduce correctly the entropy solution of the PDE (2.1). We also observe that the solutions obtained by CASE 1 and CASE 2 scheme are quite close.

Apart from the linear model, we have also tested numerically our CASE 1 and CASE 2 schemes in model Example 5.2 against a NN & Upwind method based on splitting schemes presented in [6], see Figure 2. We



(a) (b)

(c)

FIGURE 2. Model of Example 5.2: Comparison of the two methods CASE 1 & CASE2 with $d = 4$. The Proxy solution is given by the Neural nets & Upwind in [6]. For $BT\&SPD$: both CASE 1 and CASE 2, the number of particles is $M = 3500$ and the number of time steps $N = 20$. (a) $\sigma = 0.01$. (b) $\sigma = 0.3$. (c) $\sigma = 1.0$.

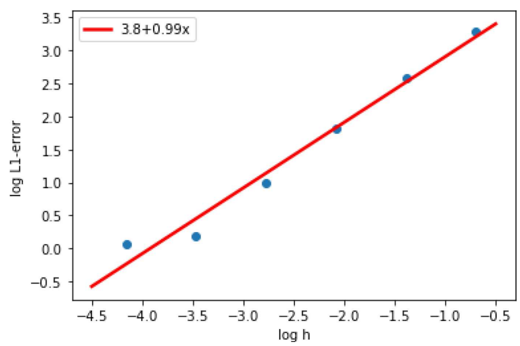


FIGURE 3. Convergence rate on $\mathfrak{h}$ for the model in Example 5.1 with parameters $d = 4, \sigma = 1.0$.

note that the function μ in Example 5.2 is always non negative, thus we use the Upwind scheme which is less "diffusive" than the Lax-Friedrichs scheme as argued in [6]. As expected, both CASE 1 and CASE 2 schemes give satisfying numerical results for different levels of volatility.

At last, we want to empirically estimate the convergence rate of the error introduced by our numerical scheme. We consider the model Example 5.1 where $\sigma = 1.0$. We consider a set of number of time steps $N :=$

$\{2, 4, 8, 16, 32, 64\}$ and time step $\mathfrak{h} := \frac{T}{N}$, and compute the $L1$-error for BT & SPD method. Note that the proxy solution is always given by method in [1] applied to one dimensional equivalent model. The empirical convergence rate with respect to the time step is close to one see Figure 3, which is much better than the order $\frac{1}{6}$ obtained in Theorem 3.1.

Acknowledgments

The authors would like to thank Benjamin Jourdain for fruitful discussions.

Data availability statement

No new data/codes were created or analyzed in this study.

References

[1] M. Bossy and D. Talay, Convergence rate for the approximation of the limit law of weakly interacting particles: application to the burgers equation. *Ann. Appl. Probab.* **6** (1996) 818–861.

[2] Y. Brenier and E. Grenier, Sticky particles and scalar conservation laws. *SIAM J. Numer. Anal.* **35** (1998) 2317–2328.

[3] R. Carmona and F. Delarue, Singular FBSDEs and scalar conservation laws driven by diffusion processes. *Probab. Theory Relat. Fields* **157** (2013) 333–388.

[4] R. Carmona, F. Delarue, G.-E. Espinosa and N. Touzi, Singular forward–backward stochastic differential equations and emissions derivatives. *Ann. Appl. Probab.* **23** (2013) 1086–1128.

[5] J.-F. Chassagneux and C. Garcia Trillos, Cubature method to solve BSDEs: error expansion and complexity control. *Math. Comput.* **89** (2020) 1895–1932.

[6] J.-F. Chassagneux and M. Yang, Numerical approximation of singular forward-backward SDEs. *J. Comput. Phys.* **468** (2022) 111459.

[7] J.-F. Chassagneux, D. Crisan and F. Delarue, Numerical method for FBSDEs of Mckean–Vlasov type. *Ann. Appl. Probab.* **29** (2019) 1640–1684.

[8] J.-F. Chassagneux, H. Chotai and D. Crisan, Modelling multiperiod carbon markets using singular forward-backward SDEs. *Math. Oper. Res.* **48** (2023) 463–497.

[9] L.G. Gyurkó and T.J. Lyons, Efficient and practical implementations of cubature on wiener space, in Stochastic Analysis 2010. Springer (2011) 73–111.

[10] S. Howison and D. Schwarz, Risk-neutral pricing of financial instruments in emission markets: a structural approach. *SIAM J. Finan. Math.* **3** (2012) 709–739.

[11] B. Jourdain and J. Reygner, Optimal convergence rate of the multitype sticky particle approximation of one-dimensional diagonal hyperbolic systems with monotonic initial data. *Discrete Contin. Dyn. Syst. – A* **36** (2016).

[12] C. Litterer and T. Lyons, High order recombination and an application to cubature on Wiener space. *Ann. Appl. Probab.* **22** (2012) 1301–1327.

[13] J. Zhang, *Some fine properties of backward stochastic differential equations*. Ph.D. thesis, Purdue University (2001).