Multilevel Picard iterations for solving smooth semilinear parabolic heat equations

Weinan E[1], Martin Hutzenthaler[2], Arnulf Jentzen[3] & Thomas Kruse[4]

[1] Department of Mathematics and Program in Applied and Computational Mathematics, Princeton University, Princeton, NJ 08544-1000, USA, e-mail: weinan@math.princeton.edu

[2] Faculty of Mathematics, University of Duisburg-Essen, 45117 Essen, Germany, e-mail: martin.hutzenthaler@uni-due.de

[3] Seminar für Angewandte Mathematik, ETH Zurich, 8092 Zürich, Switzerland, e-mail: arnulf.jentzen@sam.math.ethz.ch

[4] Faculty of Mathematics, University of Duisburg-Essen, 45117 Essen, Germany, e-mail: thomas.kruse@uni-due.de

February 26, 2019

arXiv:1607.03295v4 [math.NA] 22 Feb 2019

Abstract

We introduce a new family of numerical algorithms for approximating solutions of general high-dimensional semilinear parabolic partial differential equations at single space-time points. The algorithm is obtained through a delicate combination of the Feynman-Kac and the Bismut-Elworthy-Li formulas, and an approximate decomposition of the Picard fixed-point iteration with multilevel accuracy. The algorithm has been tested on a variety of semilinear partial differential equations that arise in physics and finance, with very satisfactory results. Analytical tools needed for the analysis of such algorithms, including a semilinear Feynman-Kac formula, a new class of semi-norms and their recursive inequalities, are also introduced. They allow us to prove for semilinear heat equations with gradient-independent nonlinearity that the computational complexity of the proposed algorithm is bounded by $O(d\,\varepsilon^{-(4+\delta)})$ for any $\delta \in (0,\infty)$ under suitable assumptions, where $d \in \mathbb{N}$ is the dimensionality of the problem and $\varepsilon \in (0,\infty)$ is the prescribed accuracy.

1 Introduction and main results

High-dimensional partial differential equations (PDEs) arise naturally in many important areas including quantum mechanics, statistical physics, financial engineering, economics, etc. Yet developing efficient and practical algorithms for these high-dimensional PDEs has been a long-standing problem and indeed one of the most challenging tasks in mathematics. The difficulty lies in the "curse of dimensionality" [4], i.e., the complexity of the problem goes up exponentially as a function of dimension, which is a well-known obstacle that is also at the heart of many other important subjects such as high-dimensional statistics and the modeling of many-body systems.

For *linear* parabolic PDEs, the Feynman-Kac formula establishes an explicit representation of the solution of the PDE as the expectation of the solution of an appropriate stochastic differential equation (SDE). Monte Carlo methods together with suitable discretizations of the SDE (see, e.g., [34, 33, 30, 29]) then allow to approximate the solution at any single point in space-time with a computational complexity that grows as $O(d\varepsilon^{-(2+\delta)})$ for any $\delta > 0$ where d is the dimensionality of the problem and ε is the accuracy required (cf., e.g., [22, 19, 24, 25]).

In the seminal papers [36, 37, 35], Pardoux & Peng established a generalized *nonlinear Feynman-Kac formula* that gives an explicit representation of the solutions of a semilinear parabolic PDE through the solution of an appropriate backward stochastic differential equation (BSDE). Solving the BSDEs numerically, however, requires in general suitable discretizations of nested conditional expectations (see, e.g., [7, 42]) and the straightforward Monte Carlo method applied to these nested conditional expectations results in an algorithm with a computational complexity that grows polynomially in d but exponentially in ε^{-1}. Other discretization methods for the nested conditional expectations proposed in the literature include the quantization tree method (see [3]), the regression method based on Malliavin calculus or based on kernel estimation (see [7]), the projection on function spaces method (see [21]), the cubature on Wiener space method (see [12]), and the Wiener chaos

AMS 2010 subject classification: 65M75

Key words and phrases: curse of dimensionality, high-dimensional PDEs, high-dimensional semilinear BSDEs, multilevel Picard iteration, multilevel Monte Carlo method

decomposition method (see [8]). None of these algorithms meets the requirement that the computational complexity grows at most polynomially both in d and ε^{-1} (see [16, Subsections 6.1–6.6] for a detailed discussion of these approximation methods).

Another probabilistic representation for the solutions of some semilinear parabolic PDEs with polynomial nonlinearity has been established in Skorohod [40] by means of *branching diffusion processes*. Recently this classical representation has been extended to more general analytic nonlinearities [26, 28, 27]. This probabilistic representation has been successfully used to obtain a Monte Carlo approximation method for semilinear parabolic PDEs with a computational complexity that grows polynomially both in d and ε^{-1}. However, not only is this method only applicable to a special class of PDEs, it also requires the terminal/initial condition to be quite small (see [16, Subsection 6.7] for a detailed discussion).

In this paper we propose a new family of numerical algorithms for approximating solutions of general high-dimensional semilinear parabolic PDEs (and BSDEs) at single space-time points; see (12) below for the definition of our approximations. For semilinear heat equations with gradient-independent nonlinearities we prove that the computational complexity (see Corollary 3.18 below for the precise meaning hereof) of our proposed algorithm is $O(d\,\varepsilon^{-(4+\delta)})$ for any $\delta > 0$ under suitable assumptions including the strong smoothness assumption that the constant in (86) below is finite; see Corollary 3.18 below for details. Under the assumptions of Corollary 3.18, to the best of our knowledge, no implementable approximation method was known in the literature to overcome the curse of dimensionality. The analysis of more general coefficient functions and nonlinearities is deferred to future publications. The algorithm, which we will call "multilevel Picard iteration", is a delicate combination of the Feynman-Kac and Bismut-Elworthy-Li formulas, and a decomposition of the Picard iteration with multilevels of accuracy. The efficiency and accuracy of the proposed algorithm has been tested on a variety of semilinear parabolic PDEs that arise in physics and finance. These details are presented in [16]. To get a feeling about the performance of the algorithm: To evaluate $u(1,0)$ for the solution of

$$\partial_t u = \tfrac{1}{2}\Delta u + u - u^3 \tag{1}$$

with $d = 100, \varepsilon = 0.01, u(0,x) = (1+\max\{|x_1|^2, ..., |x_{100}|^2)^{-1}$ requires 10 seconds of runtime on a 2.8 GHz Intel i7 processor with 16 GB RAM.

We also introduce the tools needed to analyze these high-dimensional algorithms. Some of these tools are quite non-standard (e.g. the semi-norms (18) and the recursive inequality (54) involving different semi-norms). Using these tools, we are able to establish rigorously the bounds for the computational complexity mentioned above.

1.1 Notation

Since the proposed algorithm relies heavily on the Feynman-Kac formula, we will adopt the notations and conventions in stochastic analysis. In addition, we frequently use the following notation. We denote by $\|\cdot\| : (\cup_{n\in\mathbb{N}}\mathbb{R}^n) \to [0,\infty)$ and $\langle\cdot,\cdot\rangle\colon (\cup_{n\in\mathbb{N}}(\mathbb{R}^n\times\mathbb{R}^n)) \to [0,\infty)$ the functions that satisfy for all $n\in\mathbb{N}$, $v=(v_1,\ldots,v_n)$, $w=(w_1,\ldots,w_n)\in\mathbb{R}^n$ that $\|v\| = \big[\sum_{i=1}^n |v_i|^2\big]^{1/2}$ and $\langle v,w\rangle = \sum_{i=1}^n v_i w_i$. For every topological space $(E,\mathcal{E})$ we denote by $\mathcal{B}(E)$ the Borel-sigma-algebra on $(E,\mathcal{E})$. For all measurable spaces $(A,\mathcal{A})$ and $(B,\mathcal{B})$ we denote by $\mathcal{M}(\mathcal{A},\mathcal{B})$ the set of $\mathcal{A}/\mathcal{B}$-measurable functions from A to B. For all metric spaces (E,d_E) and (F,d_F) we denote by $\operatorname{Lip}(E,F)$ the set of all globally Lipschitz continuous functions from E to F. For every $d\in\mathbb{N}$ we denote by $\mathbb{R}^{d\times d}_{\mathrm{Inv}}$ the set of invertible matrices in $\mathbb{R}^{d\times d}$. For every $d\in\mathbb{N}$ and every $A\in\mathbb{R}^{d\times d}$ we denote by $A^*\in\mathbb{R}^{d\times d}$ the transpose of A. For every $d\in\mathbb{N}$ and every $x=(x_1,\ldots,x_d)\in\mathbb{R}^d$ we denote by $\operatorname{diag}(x)\in\mathbb{R}^{d\times d}$ the diagonal matrix with diagonal entries $x_1,\ldots,x_d$. For every $T\in(0,\infty)$ we denote by $\mathcal{Q}_T$ the set given by $\mathcal{Q}_T = \{w\colon [0,T]\to\mathbb{R}\colon w^{-1}(\mathbb{R}\backslash\{0\}) \text{ is a finite set}\}$. We denote by $\lfloor\cdot\rfloor\colon\mathbb{R}\to\mathbb{Z}$ and $[\cdot]^+\colon\mathbb{R}\to[0,\infty)$ the functions that satisfy for all $x\in\mathbb{R}$ that $\lfloor x\rfloor = \max(\mathbb{Z}\cap(-\infty,x])$. and $[x]^+ = \max\{x,0\}$. We denote by $\frac{0}{0}$, $0\cdot\infty$, and 0^0 the real numbers given by $\frac{0}{0}=0$, $0\cdot\infty=0$, and $0^0=1$.

2 Multilevel Picard iteration for semilinear parabolic PDEs

2.1 A fixed-point equation for semilinear PDEs

Let $T>0$, $d\in\mathbb{N}$, let $g\colon\mathbb{R}^d\to\mathbb{R}$, $f\colon\mathbb{R}\times\mathbb{R}^d\to\mathbb{R}$, $u\colon[0,T]\times\mathbb{R}^d\to\mathbb{R}$, $\mu\colon[0,T]\times\mathbb{R}^d\to\mathbb{R}^d$, and $\sigma=(\sigma_1,\ldots,\sigma_d)\colon[0,T]\times\mathbb{R}^d\to\mathbb{R}^{d\times d}_{\mathrm{Inv}}$ be sufficiently regular functions, assume that $u(T,x)=g(x)$ and

$$\partial_t u + f(u,\sigma^*\nabla u) + \langle\mu,\nabla u\rangle + \tfrac{1}{2}\operatorname{Trace}(\sigma\sigma^*\operatorname{Hess}u) = 0, \tag{2}$$

for $t\in[0,T)$, $x\in\mathbb{R}^d$, let $(\Omega,\mathcal{F},\mathbb{P},(\mathbb{F}_t)_{t\in[0,T]})$ be a stochastic basis (cf., e.g., [38, Appendix E]), let $W=(W^1,\ldots,W^d)\colon[0,T]\times\Omega\to\mathbb{R}^d$ be a standard $(\mathbb{F}_t)_{t\in[0,T]}$-Brownian motion, and for every $s\in[0,T]$, $x\in\mathbb{R}^d$ let $X^{s,x}\colon[s,T]\times\Omega\to\mathbb{R}^d$ and $D^{s,x}\colon[s,T]\times\Omega\to\mathbb{R}^{d\times d}$ be $(\mathbb{F}_t)_{t\in[s,T]}$-adapted stochastic processes with continuous sample paths which satisfy that for all $t \in [s,T]$ it holds $\mathbb{P}$-a.s. that

$$\begin{aligned} X^{s,x}_t &= x + \int_s^t \mu(r, X^{s,x}_r)\, dr + \sum_{j=1}^d \int_s^t \sigma_j(r, X^{s,x}_r)\, dW^j_r, \\ D^{s,x}_t &= \mathrm{I}_{\mathbb{R}^{d\times d}} + \int_s^t (\tfrac{\partial}{\partial x}\mu)(r, X^{s,x}_r)\, D^{s,x}_r\, dr + \sum_{j=1}^d \int_s^t (\tfrac{\partial}{\partial x}\sigma_j)(r, X^{s,x}_r)\, D^{s,x}_r\, dW^j_r \end{aligned} \tag{3}$$

(cf., e.g., [31, Chapter 5], [23], or [2] for existence and uniqueness results for stochastic differential equations of the form (3)). For every $s \in [0,T]$ the processes $D^{s,x}$, $x \in \mathbb{R}^d$, are in a suitable sense the *derivative processes* of $X^{s,x}$, $x \in \mathbb{R}^d$, with respect to $x \in \mathbb{R}^d$. Using the *Feynman-Kac formula*, we have from (2)

$$u(s,x) = \mathbb{E}[g(X^{s,x}_T)] + \int_s^T \mathbb{E}[f(u(t, X^{s,x}_t), [\sigma(t, X^{s,x}_t)]^*(\nabla u)(t, X^{s,x}_t))]\, dt \tag{4}$$

for all $(s,x) \in [0,T) \times \mathbb{R}^d$. In (4) the derivative of u appears on the right-hand side and, therefore, (4) does not provide a closed fixed point equation. To obtain such a closed fixed point equation we now bring the *Bismut-Elworthy-Li formula* into play (see, e.g., Elworthy & Li [17, Theorem 2.1] or Da Prato & Zabczyk [13, Theorem 2.1]). This gives us

$$\begin{aligned} [\sigma(s,x)]^*(\nabla u)(s,x) &= \mathbb{E}\Big[g(X^{s,x}_T) \tfrac{[\sigma(s,x)]^*}{T-s} \textstyle\int_s^T \big[\sigma(r, X^{s,x}_r)^{-1} D^{s,x}_r\big]^* dW_r\Big] \\ &\quad + \int_s^T \mathbb{E}\Big[f(u(t, X^{s,x}_t), [\sigma(t, X^{s,x}_t)]^*(\nabla u)(t, X^{s,x}_t))\, \tfrac{[\sigma(s,x)]^*}{t-s} \textstyle\int_s^t \big[\sigma(r, X^{s,x}_r)^{-1} D^{s,x}_r\big]^*\, dW_r\Big]\, dt, \end{aligned} \tag{5}$$

for all $(s,x) \in [0,T) \times \mathbb{R}^d$. Now let $\mathbf{u}^\infty \in \mathrm{Lip}([0,T] \times \mathbb{R}^d, \mathbb{R}^{1+d})$ be defined by $\mathbf{u}^\infty(s,x) = \big(u(s,x), [\sigma(s,x)]^*(\nabla u)(s,x)\big)$ for all $(s,x) \in [0,T) \times \mathbb{R}^d$. Let $\Phi\colon \mathrm{Lip}([0,T] \times \mathbb{R}^d, \mathbb{R}^{1+d}) \to \mathrm{Lip}([0,T] \times \mathbb{R}^d, \mathbb{R}^{1+d})$ be defined by

$$\begin{aligned} \big(\Phi(\mathbf{v})\big)(s,x) &= \mathbb{E}\Big[g(X^{s,x}_T) \Big(1, \tfrac{[\sigma(s,x)]^*}{T-s} \textstyle\int_s^T \big[\sigma(r, X^{s,x}_r)^{-1} D^{s,x}_r\big]^* dW_r\Big)\Big] \\ &\quad + \int_s^T \mathbb{E}\Big[f(\mathbf{v}(t, X^{s,x}_t))\, \big(1, \tfrac{[\sigma(s,x)]^*}{t-s} \textstyle\int_s^t \big[\sigma(r, X^{s,x}_r)^{-1} D^{s,x}_r\big]^*\, dW_r\big)\Big]\, dt. \end{aligned} \tag{6}$$

for all $\mathbf{v} \in \mathrm{Lip}([0,T] \times \mathbb{R}^d, \mathbb{R}^{1+d})$, $(s,x) \in [0,T) \times \mathbb{R}^d$. Combining (6) with (4) and (5) gives

$$\mathbf{u}^\infty = \Phi(\mathbf{u}^\infty). \tag{7}$$

Next we define a sequence of Picard iterations associated to (6),

$$\mathbf{u}_k(s,x) = (\mathbf{\Phi}(\mathbf{u}_{k-1}))(s,x) \tag{8}$$

for all $k \in \mathbb{N}$, $s \in [0,T)$, $x \in \mathbb{R}^d$. This sequence of Picard iterations has already been studied in the literature; see, e.g. Thereom 7.3.4 in [41] or [5]. Under suitable assumptions, e.g., Thereom 7.3.4 in [41] ensures that for all $s \in [0,T)$, $x \in \mathbb{R}^d$ it holds that $\lim_{k\to\infty} \mathbf{u}_k(s,x) = \mathbf{u}^\infty(s,x)$. Observe that for all $k \in \mathbb{N}$, $s \in [0,T)$, $x \in \mathbb{R}^d$ it holds that

$$\begin{aligned} \mathbf{u}_k(s,x) &= \mathbf{u}_1(s,x) + \sum_{l=1}^{k-1} [\mathbf{u}_{l+1}(s,x) - \mathbf{u}_l(s,x)] = (\mathbf{\Phi}(\mathbf{u}_0))(s,x) + \sum_{l=1}^{k-1} \Big[(\mathbf{\Phi}(\mathbf{u}_l))(s,x) - (\mathbf{\Phi}(\mathbf{u}_{l-1}))(s,x)\Big] \\ &= \mathbb{E}\Big[g(X^{s,x}_T) \Big(1, \tfrac{[\sigma(s,x)]^*}{T-s} \textstyle\int_s^T \big[\sigma(r, X^{s,x}_r)^{-1} D^{s,x}_r\big]^* dW_r\Big)\Big] \\ &\quad + \sum_{l=0}^{k-1} \int_s^T \mathbb{E}\Big[\big(f(\mathbf{u}_l(t, X^{s,x}_t)) - \mathbb{1}_{\mathbb{N}}(l) f(\mathbf{u}_{l-1}(t, X^{s,x}_t))\big)\, \big(1, \tfrac{[\sigma(s,x)]^*}{t-s} \textstyle\int_s^t \big[\sigma(r, X^{s,x}_r)^{-1} D^{s,x}_r\big]^*\, dW_r\big)\Big]\, dt. \end{aligned} \tag{9}$$

Next we incorporate a zero expectation term to slightly reduce the variance when approximating the expectation involving g by Monte Carlo approximations. More precisely, for all $k \in \mathbb{N}$, $s \in [0,T)$, $x \in \mathbb{R}^d$ it holds that

$$\begin{aligned} \mathbf{u}_k(s,x) &= (g(x), 0) + \mathbb{E}\Big[(g(X^{s,x}_T) - g(x)) \Big(1, \tfrac{[\sigma(s,x)]^*}{T-s} \textstyle\int_s^T \big[\sigma(r, X^{s,x}_r)^{-1} D^{s,x}_r\big]^* dW_r\Big)\Big] \\ &\quad + \sum_{l=0}^{k-1} \int_s^T \mathbb{E}\Big[\big(f(\mathbf{u}_l(t, X^{s,x}_t)) - \mathbb{1}_{\mathbb{N}}(l) f(\mathbf{u}_{l-1}(t, X^{s,x}_t))\big)\, \big(1, \tfrac{[\sigma(s,x)]^*}{t-s} \textstyle\int_s^t \big[\sigma(r, X^{s,x}_r)^{-1} D^{s,x}_r\big]^*\, dW_r\big)\Big]\, dt. \end{aligned} \tag{10}$$

In this telescope expansion, we will apply a fundamental idea of Heinrich [24, 25] and Giles [18] (control variates were also used, e.g., in [32, 20]) and approximate the continuous quantities (expectation and time integral) by discrete ones (Monte Carlo averages and quadrature formulas respectively) with different degrees of accuracy at different levels of the Picard iteration. Since for large $l \in \mathbb{N}$ the difference between $\mathbf{u}_l$ and $\mathbf{u}_{l-1}$ is small, say ρ^{-l}, it suffices to approximate the expectation and the time integral with lower accuracy, say $\rho^{-(k-l)}$, at level $l \in \{0, \ldots, k-1\}$ for the k-th approximation. More precisely, we denote by $(q_s^{k,\rho})_{k\in\mathbb{N}_0,\rho\in(0,\infty),s\in[0,T)} \subseteq \mathcal{Q}_T$ a family of quadrature formulas on $C([0,T],\mathbb{R})$ that we employ to approximate the time integrals $\int_s^T \ldots dt$, $s \in [0,T]$, appearing on the right-hand side of (10). We denote by $\Theta = \cup_{n\in\mathbb{N}}\mathbb{R}^n$ a set that allows to index families of independent random variables which we need for the Monte Carlo approximations. We denote by $(\mathfrak{m}_{k,\rho})_{k\in\mathbb{N}_0,\rho\in(0,\infty)}$ and $(m_{k,\rho})_{k\in\mathbb{N}_0,\rho\in(0,\infty)} \subseteq \mathbb{N}$ families of natural numbers that specify the number of Monte Carlo samples for approximating the expectations involving g and f on the right-hand side of (10). In Section 3.1 we will take $\mathfrak{m}_{k,\rho} = m_{k,\rho} = \rho^k$ for every $k \in \mathbb{N}_0$, $\rho \in (0,\infty)$ and we take $q^{k,\rho}$ as the Gauß-Legendre quadrature rule with $\lfloor\rho\rfloor$ nodes. Furthermore, for every $k \in \mathbb{N}_0$, $\rho \in (0,\infty)$, $\theta \in \Theta$, $(s,x) \in [0,T] \times \mathbb{R}^d$ we denote by $(\mathcal{X}^\theta_{k,\rho}(s,x,t))_{t\in[s,T]}$ and $(\mathcal{I}^\theta_{k,\rho}(s,x,t))_{t\in(s,T]}$ the stochastic processes that we employ to approximate the processes $(X_t^{s,x})_{t\in[s,T]}$ and $\big(1, \frac{[\sigma(s,x)]^*}{t-s}\int_s^t \big[\sigma(r,X_r^{s,x})^{-1}D_r^{s,x}\big]^* dW_r\big)_{t\in(s,T]}$. More specifically, we choose for every $k \in \mathbb{N}_0$, $\rho \in (0,\infty)$, $\theta \in \Theta$, $(s,x) \in [0,T] \times \mathbb{R}^d$ the processes $(\mathcal{X}^\theta_{k,\rho}(s,x,t))_{t\in[s,T]}$ and $(\mathcal{I}^\theta_{k,\rho}(s,x,t))_{t\in(s,T]}$ such that for all $t \in (s,T]$,

$$\begin{aligned}
&\mathcal{X}^\theta_{k,\rho}(s,x,t) \approx X_t^{s,x}, \\
&\mathcal{I}^\theta_{k,\rho}(s,x,t) \approx \big(1, \tfrac{[\sigma(s,x)]^*}{t-s}\textstyle\int_s^t \big[\sigma(r,X_r^{s,x})^{-1}D_r^{s,x}\big]^* dW_r\big).
\end{aligned} \tag{11}$$

2.2 The approximation scheme

Let $T \in (0,\infty)$, $d \in \mathbb{N}$, $\Theta = \cup_{n\in\mathbb{N}}\mathbb{R}^n$, let $g\colon \mathbb{R}^d \to \mathbb{R}$, $f\colon \mathbb{R}^{d+1} \to \mathbb{R}$, $\mu\colon [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$, $\sigma\colon [0,T]\times\mathbb{R}^d \to \mathbb{R}^{d\times d}_{\mathrm{Inv}}$ be measurable functions, let $(q_s^{k,\rho})_{k\in\mathbb{N}_0,\rho\in(0,\infty),s\in[0,T)} \subseteq \mathcal{Q}_T$, $(\mathfrak{m}_{k,\rho})_{k\in\mathbb{N}_0,\rho\in(0,\infty)}$, $(m_{k,\rho})_{k\in\mathbb{N}_0,\rho\in(0,\infty)} \subseteq \mathbb{N}$, let $(\Omega,\mathcal{F},\mathbb{P},(\mathbb{F}_t)_{t\in[0,T]})$ be a stochastic basis, let $W^\theta\colon [0,T]\times\Omega \to \mathbb{R}^d$, $\theta \in \Theta$, be independent standard $(\mathbb{F}_t)_{t\in[0,T]}$-Brownian motions with continuous sample paths, for every $l \in \mathbb{Z}$, $\rho \in (0,\infty)$, $\theta \in \Theta$, $x \in \mathbb{R}^d$, $s \in [0,T)$, $t \in [s,T]$ let $\mathcal{X}^\theta_{l,\rho}(s,x,t)\colon \Omega \to \mathbb{R}^d$ and $\mathcal{I}^\theta_{l,\rho}(s,x,t)\colon \Omega \to \mathbb{R}^{1+d}$ be functions, and for every $\theta \in \Theta$, $\rho \in (0,\infty)$ let $\mathbf{U}^\theta_{k,\rho}\colon [0,T]\times\mathbb{R}^d\times\Omega \to \mathbb{R}^{d+1}$, $k \in \mathbb{N}_0$, be functions that satisfy for all $k \in \mathbb{N}$, $(s,x) \in [0,T)\times\mathbb{R}^d$ that

$$\begin{aligned}
\mathbf{U}^\theta_{k,\rho}(s,x) &= (g(x),0) + \sum_{i=1}^{\mathfrak{m}_{k,\rho}} \tfrac{1}{\mathfrak{m}_{k,\rho}} \big[g(\mathcal{X}^{(\theta,0,-i)}_{k,\rho}(s,x,T)) - g(x)\big]\,\mathcal{I}^{(\theta,0,-i)}_{k,\rho}(s,x,T) \\
&\quad + \sum_{l=0}^{k-1}\sum_{i=1}^{m_{k-l,\rho}}\sum_{t\in[s,T]} \tfrac{q_s^{k-l,\rho}(t)}{m_{k-l,\rho}} \Big[f\Big(\mathbf{U}^{(\theta,l,i,t)}_{l,\rho}\big(t,\mathcal{X}^{(\theta,l,i)}_{k-l,\rho}(s,x,t)\big)\Big) \\
&\quad - \mathbb{1}_{\mathbb{N}}(l)\, f\Big(\mathbf{U}^{(\theta,-l,i,t)}_{[l-1]^+,\rho}\big(t,\mathcal{X}^{(\theta,l,i)}_{k-l,\rho}(s,x,t)\big)\Big)\Big]\,\mathcal{I}^{(\theta,l,i)}_{k-l,\rho}(s,x,t).
\end{aligned} \tag{12}$$

Observe that the approximation scheme (12) employs Picard fixed-point iteration (cf., e.g., [5]), multilevel/-multigrid techniques (see, e.g., [24, 25, 19, 11]), discretizations of the SDE system (3), as well as quadrature approximations for the time integrals. The numerical approximations (12) are full history recursive in the sense that for every $(k,\rho) \in \mathbb{N}\times(0,\infty)$ the full history $\mathbf{U}^{(\cdot)}_{0,\rho}$, $\mathbf{U}^{(\cdot)}_{1,\rho}$, $\ldots$, $\mathbf{U}^{(\cdot)}_{k-1,\rho}$ needs to be computed recursively in order to compute $\mathbf{U}^{(\cdot)}_{k,\rho}$. In this sense the numerical approximations (12) are full history recursive multilevel Picard approximations. Finally we remark that all multilevel Picard approximations on the right-hand side of (12) are independent since all Brownian motions W^θ, $\theta \in \Theta$, are independent. This independence is useful for the mathematical analysis and allows an implementation with a simple recursive structure (cf. Subsection 3.2).

2.3 Numerical simulations of high-dimensional semilinear PDEs

We applied the algorithm (12) to approximate the solutions at single space-time points of several semilinear PDEs from physics and financial mathematics such as

(i) a PDE arising from the recursive pricing model with default risk due to Duffie, Schroder, & Skiadas [15],

(ii) a PDE arising from the valuation of derivative contracts with counterparty credit risk (see, e.g., Burgard & Kjaer [9] and Henry-Labordère [26] for derivations of the PDE),

(iii) a PDE arising from pricing models for financial markets with different interest rates for borrowing and lending due to Bergman [6],

(iv) a version of the Allen-Cahn equation with a double well potential, and

(v) a PDE with an explicit solution whose three-dimensional version has been considered in Chassagneux [10].

We took $d = 100$. All simulations are performed on a computer with a 2.8 GHz Intel i7 processor and 16 GB RAM. We refer to [16] for the simulation results, MATLAB codes and further details concerning the numerical simulations. These results suggest that the proposed algorithm is highly efficient and quite practical for dealing with these high-dimensional PDEs.

3 Convergence rate for the multilevel Picard iteration

In this section we establish the convergence rate for semilinear heat equations in the case where the nonlinearity is independent of the gradient of the solution and satisfies the Lipschitz-type condition (13) below and when the Gauß-Legendre formula (15) (see, e.g., [14] for more details) is used as the quadrature rule.

3.1 Setting

Let $T, L \in (0,\infty)$, $d \in \mathbb{N}$, $g \in C^2(\mathbb{R}^d, \mathbb{R})$, $\Theta = \cup_{n\in\mathbb{N}}\mathbb{R}^n$, let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathbb{F}_t)_{t\in[0,T]})$ be a stochastic basis, let $W^\theta\colon [0,T]\times\Omega \to \mathbb{R}^d$, $\theta\in\Theta$, be independent standard $(\mathbb{F}_t)_{t\in[0,T]}$-Brownian motions with continuous sample paths, let $f\colon [0,T]\times\mathbb{R}^d\times\mathbb{R}\to\mathbb{R}$ be a Borel measurable function which satisfies for all $t\in[0,T]$, $x\in\mathbb{R}^d$, $u_1,u_2\in\mathbb{R}$ that

$$|f(t,x,u_1) - f(t,x,u_2)| \le L|u_1-u_2|, \tag{13}$$

let $F\colon \mathcal{M}(\mathcal{B}([0,T]\times\mathbb{R}^d),\mathcal{B}(\mathbb{R})) \to \mathcal{M}(\mathcal{B}([0,T]\times\mathbb{R}^d),\mathcal{B}(\mathbb{R}))$ be the function which satisfies for all $t\in[0,T]$, $x\in\mathbb{R}^d$, $u\in\mathcal{M}(\mathcal{B}([0,T]\times\mathbb{R}^d),\mathcal{B}(\mathbb{R}))$ that $(F(u))(t,x) = f(t,x,u(t,x))$, let $u^\infty = (u^\infty(r,y))_{(r,y)\in[0,T]\times\mathbb{R}^d} \in C^{1,2}([0,T]\times\mathbb{R}^d,\mathbb{R})$ satisfy for all $r\in[0,T]$, $y\in\mathbb{R}^d$ that $u^\infty(T,y) = g(y)$ and

$$\partial_r u^\infty(r,y) + \frac{1}{2}(\Delta_y u^\infty)(r,y) + (F(u^\infty))(r,y) = 0, \tag{14}$$

for every $n\in\mathbb{N}$ let $(c_i^n)_{i\in\{1,\dots,n\}} \subseteq [-1,1]$ be the n distinct roots of the Legendre polynomial $[-1,1]\ni x\mapsto \frac{1}{2^n n!}\frac{d^n}{dx^n}[(x^2-1)^n]\in\mathbb{R}$, for every $n\in\mathbb{N}$, $a\in\mathbb{R}$, $b\in[a,\infty)$ let $q^{n,[a,b]}\colon [a,b]\to\mathbb{R}$ be the function which satisfies for all $t\in[a,b]$ that

$$q^{n,[a,b]}(t) = \begin{cases} \int_a^b \left[\prod_{\substack{i\in\{1,\dots,n\},\\ c_i^n\neq\frac{2t-(a+b)}{b-a}}} \frac{2x-(b-a)c_i^n-(a+b)}{2t-(b-a)c_i^n-(a+b)}\right] dx & : (a<b) \text{ and } \left(\frac{2t-(a+b)}{b-a}\in\{c_1^n,\dots,c_n^n\}\right) \\ 0 & : \text{else}, \end{cases} \tag{15}$$

let $(\bar{q}^{n,Q})_{n,Q\in\mathbb{N}_0}\subseteq\mathcal{Q}_T$ satisfy for all $n,Q\in\mathbb{N}$, $t\in[0,T]$ that $\bar{q}^{0,Q}(t) = \mathbb{1}_{\{0\}}(t)$ and

$$\bar{q}^{n,Q}(t) = \sum_{s\in[0,t]} \bar{q}^{n-1,Q}(s)\, q^{Q,[s,T]}(t), \tag{16}$$

let $(U^\theta_{n,M,Q})_{n,M,Q\in\mathbb{Z},\theta\in\Theta}\subseteq\mathcal{M}(\mathcal{B}([0,T]\times\mathbb{R}^d)\otimes\mathcal{F},\mathcal{B}(\mathbb{R}))$ satisfy for all $n,M,Q\in\mathbb{N}$, $\theta\in\Theta$, $(t,x)\in[0,T]\times\mathbb{R}^d$ that $U^\theta_{0,M,Q}(t,x) = 0$ and

$$\begin{aligned} U^\theta_{n,M,Q}(t,x) &= \frac{1}{M^n}\sum_{i=1}^{M^n} g(x + W_T^{(\theta,0,-i)} - W_t^{(\theta,0,-i)}) \\ &+ \sum_{l=0}^{n-1}\sum_{i=1}^{M^{n-l}}\sum_{s\in[t,T]} \frac{q^{Q,[t,T]}(s)}{M^{n-l}}\big(F(U^{(\theta,l,i,s)}_{l,M,Q}) - \mathbb{1}_{\mathbb{N}}(l)F(U^{(\theta,-l,i,s)}_{l-1,M,Q})\big)(s, x + W_s^{(\theta,l,i)} - W_t^{(\theta,l,i)}), \end{aligned} \tag{17}$$

for every $n,Q\in\mathbb{N}_0$ let $\|\cdot\|_{n,Q} : \mathcal{M}(\mathcal{B}([0,T]\times\mathbb{R}^d)\otimes\mathcal{F},\mathcal{B}(\mathbb{R}))\to[0,\infty]$ be the function which satisfies

$$\|V\|_{n,Q} = \sum_{t\in[0,T]} \bar{q}^{n,Q}(t)\left[\sup_{s\in[t,T]}\sup_{u\in[0,s]}\sup_{z\in\mathbb{R}^d}\sqrt{\mathbb{E}\big[|V(s,z+W_u^0)|^2\big]}\right] \tag{18}$$

for all $V\in\mathcal{M}(\mathcal{B}([0,T]\times\mathbb{R}^d)\otimes\mathcal{F},\mathcal{B}(\mathbb{R}))$.

3.2 Pseudocode

In this subsection a mathematical style pseudocode illustrates that the multilevel Picard approximations (17) can be easily implemented. We assume that the time horizon $T\in(0,\infty)$, the dimension $d\in\mathbb{N}$, the terminal condition $g\colon\mathbb{R}^d\to\mathbb{R}$, the (gradient-independent) nonlinearity $f\colon[0,T]\times\mathbb{R}^d\times\mathbb{R}\to\mathbb{R}$, the basis for the number of Monte-Carlo samples $M \in \mathbb{N}$, the number of quadrature nodes $Q \in \mathbb{N}$, increasingly ordered roots $c \in [-1,1]^Q$ of the Q-th Legendre polynomial, and the corresponding Legendre quadrature weights $w \in [0,\infty)^Q$ are global variables. For an implementation in MATLAB see [16].

Algorithm 1 Multilevel Picard approximation

```
function MLP(n, t, x)
    c_loc ← (T − t)c/2 + (T + t)/2;                ▷ Quadrature nodes on [t, T]
    d ← c_loc − [t; c_loc(1 : (Q − 1))];           ▷ Increments between consecutive quadrature nodes
    w_loc ← (T − t)w/2;                            ▷ Quadrature weights on [t, T]
    Generate M^n realizations W(i) ∈ R^d, i ∈ {1, ..., M^n}, of independent standard normally distributed
    random vectors;
    u ← (1/M^n) Σ_{i=1}^{M^n} g(x + sqrt(T − t) W(i));
    for l ← 0 to (n − 1) do
        X(i) ← x for all i ∈ {1, ..., M^{n−l}};
        for k ← 1 to Q do
            Generate M^{n−l} realizations W(i) ∈ R^d, i ∈ {1, ..., M^{n−l}}, of independent standard normally
            distributed random vectors;
            X(i) ← X(i) + sqrt(d(k)) W(i) for all i ∈ {1, ..., M^{n−l}};
            u ← u + (w_loc(k)/M^{n−l}) Σ_{i=1}^{M^{n−l}} f(c_loc(k), X(i), MLP(l, c_loc(k), X(i)));
            if l > 0 then
                u ← u − (w_loc(k)/M^{n−l}) Σ_{i=1}^{M^{n−l}} f(c_loc(k), X(i), MLP(l − 1, c_loc(k), X(i)));
            end if
        end for
    end for
    return u;
end function
```

3.3 Sketch of the proof

Throughout this subsection assume the setting in Subsection 3.1 and let $N, M, Q \in \mathbb{N}$. Theorem 3.11 provides an upper bound for the distance between the approximation $U^0_{N,M,Q}$ and the PDE solution u^∞ measured in the semi-norms $\|\cdot\|_{n,Q}$, $n \in \mathbb{N}_0$, given in (18). We establish this bound by splitting the global error $\|U^0_{N,M,Q} - u^\infty\|_{n,Q}$ into the Monte Carlo error $\|U^0_{N,M,Q} - \mathbb{E}[U^0_{N,M,Q}]\|_{n,Q}$ and the time discretization error $\|\mathbb{E}[U^0_{N,M,Q}] - u^\infty\|_{n,Q}$. To analyze the time discretization error, we employ the Feynman-Kac formula to obtain

$$u^\infty(s,x) = \mathbb{E}\left[g(x + W^0_{T-s}) + \int_s^T (F(u^\infty))(t, x + W^0_{t-s})\,dt\right] \tag{19}$$

for all $s \in [0,T], x \in \mathbb{R}^d$ (see Lemma 3.10 below). Moreover, the approximations admit the following Feynman-Kac-type representation

$$\mathbb{E}\big[U^0_{N,M,Q}(s,x)\big] = \mathbb{E}\left[g(x + W^0_{T-s}) + \sum_{t \in [s,T]} q^{Q,[s,T]}(t)\big(F(U^0_{N-1,M,Q})\big)(t, x + W^0_{t-s})\right] \tag{20}$$

for all $s \in [0,T], x \in \mathbb{R}^d$ (see Lemma 3.9 below). This, (19) and the Lipschitz-type assumption (13) show that the time discretization error is bounded from above by the error of the $(N-1)$-th approximation $\|U^0_{N-1,M,Q} - u^\infty\|_{n+1,Q}$ and the error of the Gauß-Legendre quadrature rule applied to the function $[s,T] \ni t \mapsto \mathbb{E}[F(u^\infty)(t, x + W_{t-s})] \in \mathbb{R}$ (see (52) below). Combining this with the established bound for the Monte Carlo error (see (50) below) results in the recursive inequality for the global error (54) that can be handled using a discrete Gronwall-type inequality. The error representation for Gauß-Legendre quadrature rules allows to further simplify the global error under suitable regularity assumptions (see Corollary 3.14 below). In Section 3.7 we provide upper bounds for the number of realizations of scalar standard normal random variables and for the number of function evaluations of f and g required to compute one realization of $U^0_{N,M,Q}(t,x)$ for a single point $(t,x) \in [0,T] \times \mathbb{R}^d$ in space-time. This and Corollary 3.14 prove in the case of the semilinear heat equation (14) that the computational complexity (see Corollary 3.14 for the precise definition hereof) of our proposed scheme grows linearly in the space dimension d and polynomially in the inverse accuracy ε^{-1} under suitable assumptions (see Corollary 3.17 below).

3.4 Preliminary results for the Gauß-Legendre quadrature rules

Lemma 3.1 (Gauß-Legendre over different intervals)**.** *Assume the setting in Subsection 3.1, let $n \in \mathbb{N}$, $s \in [0,T)$, $t \in [0,s]$, and let $\psi\colon [0,T] \to [0,\infty]$ be a non-increasing function. Then we have*

$$\sum_{r\in[s,T]} q^{n,[s,T]}(r)\,\psi(r) \le \sum_{r\in[t,T]} q^{n,[t,T]}(r)\,\psi(r). \tag{21}$$

Proof. Note that (15) and the integral transformation theorem with the substitution $[s,T] \ni x \mapsto (x-s)\frac{T-t}{T-s}+t \in [t,T]$ show that

$$\begin{aligned}
&\sum_{r\in[s,T]} q^{n,[s,T]}(r)\,\psi(r) = \sum_{i=1}^{n} q^{n,[s,T]}(\tfrac{T-s}{2}c_i^n + \tfrac{T+s}{2})\,\psi(\tfrac{T-s}{2}c_i^n + \tfrac{T+s}{2})\\
&= \sum_{i=1}^{n}\left[\int_s^T \left(\prod_{j\in\{1,\dots,n\}\setminus\{i\}} \tfrac{2(x-s)-(T-s)c_j^n-(T-s)}{(T-s)c_i^n+(T+s)-(T-s)c_j^n-(T+s)}\right)dx\right]\psi(\tfrac{T-s}{2}c_i^n + \tfrac{T+s}{2})\\
&= \sum_{i=1}^{n}\tfrac{T-s}{T-t}\left[\int_t^T \left(\prod_{j\in\{1,\dots,n\}\setminus\{i\}} \tfrac{2\frac{y-t}{T-t}-c_j^n-1}{c_i^n-c_j^n}\right)dy\right]\psi(\tfrac{T-s}{2}c_i^n + \tfrac{T+s}{2})\\
&= \sum_{i=1}^{n}\tfrac{T-s}{T-t}\left[\int_t^T \left(\prod_{j\in\{1,\dots,n\}\setminus\{i\}} \tfrac{2(y-t)-(T-t)c_j^n-(T-t)}{(T-t)c_i^n+(T+t)-(T-t)c_j^n-(T+t)}\right)dy\right]\psi(\tfrac{T-s}{2}c_i^n + \tfrac{T+s}{2})\\
&= \sum_{i=1}^{n}\tfrac{T-s}{T-t}\,q^{n,[t,T]}(\tfrac{T-t}{2}c_i^n + \tfrac{T+t}{2})\,\psi(\tfrac{T-s}{2}c_i^n + \tfrac{T+s}{2}).
\end{aligned} \tag{22}$$

Observe that the fact that $t \le s$ and the fact that $\forall\, i \in \{1,\dots,n\}\colon c_i^n \in [-1,1]$ ensure that for all $i \in \{1,\dots,n\}$ it holds that $\frac{T-s}{2}c_i^n + \frac{T+s}{2} \ge \frac{T-t}{2}c_i^n + \frac{T+t}{2}$. This and the fact that ψ is non-increasing imply for all $i \in \{1,\dots,n\}$ that $\psi(\frac{T-s}{2}c_i^n + \frac{T+s}{2}) \le \psi(\frac{T-t}{2}c_i^n + \frac{T+t}{2})$. Combining this with (22), (15), and the fact that $\frac{T-s}{T-t} \le 1$ proves that

$$\sum_{r\in[s,T]} q^{n,[s,T]}(r)\,\psi(r) \le \sum_{i=1}^{n} q^{n,[t,T]}(\tfrac{T-t}{2}c_i^n + \tfrac{T+t}{2})\,\psi(\tfrac{T-t}{2}c_i^n + \tfrac{T+t}{2}) = \sum_{r\in[t,T]} q^{n,[t,T]}(r)\,\psi(r). \tag{23}$$

□

Lemma 3.2. *Assume the setting in Subsection 3.1 and let $Q \in \mathbb{N}$. Then, for all $n \in \mathbb{N}_0$, $k \in \mathbb{N}_0 \cap [0, 2Q-n)$ we have*

$$\sum_{t\in[0,T]} \bar{q}^{n,Q}(t)\,\tfrac{(T-t)^k}{k!} = \tfrac{T^{n+k}}{(n+k)!}. \tag{24}$$

Proof. First, note that the fact that the Gauß-Legendre quadrature rule $C([0,T],\mathbb{R}) \ni \varphi \mapsto \sum_{t\in[0,T]} q^{Q,[0,T]}(t)\varphi(t) \in \mathbb{R}$ integrates polynomials of order less than $2Q$ exactly implies that for all $s \in [0,T]$, $k \in \mathbb{N}_0 \cap [0,2Q)$ it holds that

$$\sum_{t\in[s,T]} q^{Q,[s,T]}(t)\,\tfrac{(T-t)^k}{k!} = \int_s^T \tfrac{(T-t)^k}{k!}\,dt = \tfrac{(T-s)^{k+1}}{(k+1)!}. \tag{25}$$

We now prove (24) by induction on $n \in \mathbb{N}_0$. For the base case $n = 0$ we note that for all $k \in \mathbb{N}_0$ it holds that

$$\sum_{t\in[0,T]} \bar{q}^{0,Q}(t)\,\tfrac{(T-t)^k}{k!} = \sum_{t\in[0,T]} \mathbb{1}_{\{0\}}(t)\,\tfrac{(T-t)^k}{k!} = \tfrac{T^k}{k!}. \tag{26}$$

This establishes (24) in the base case $n = 0$. For the induction step $\mathbb{N}_0 \ni n \to n+1 \in \mathbb{N}$ we observe that (25) and the induction hypothesis imply that for all $k \in \mathbb{N}_0 \cap [0, 2Q-n-1)$ it holds that

$$\begin{aligned}
\sum_{t\in[0,T]} \bar{q}^{n+1,Q}(t)\,\tfrac{(T-t)^k}{k!} &= \sum_{t\in[0,T]}\left[\sum_{s\in[0,t]} \bar{q}^{n,Q}(s)\,q^{Q,[s,T]}(t)\right]\tfrac{(T-t)^k}{k!} = \sum_{s\in[0,T]} \bar{q}^{n,Q}(s)\left[\sum_{t\in[s,T]} q^{Q,[s,T]}(t)\,\tfrac{(T-t)^k}{k!}\right]\\
&= \sum_{s\in[0,T]} \bar{q}^{n,Q}(s)\,\tfrac{(T-s)^{k+1}}{(k+1)!} = \tfrac{T^{n+1+k}}{(n+1+k)!}.
\end{aligned} \tag{27}$$

This finishes the induction step $\mathbb{N}_0 \ni n \to n+1 \in \mathbb{N}$. Induction hence establishes (24). The proof of Lemma 3.2 is thus completed. □

3.5 Preliminary results for the semi-norms

We refer to a $[0,\infty]$-valued function as semi-norm if it is subadditive and absolutely homogeneous. In particular, we do not require semi-norms to have finite values. The proof of the following lemma is clear and therefore omitted.

Lemma 3.3 (Seminorm property)**.** *Assume the setting in Subsection 3.1 and let $k \in \mathbb{N}_0$. Then the function $\mathcal{M}(\mathcal{B}([0,T]\times\mathbb{R}^d)\otimes\mathcal{F},\mathcal{B}(\mathbb{R})) \ni U \mapsto \|U\|_{k,Q} \in [0,\infty]$ is a semi-norm in the sense that it is subadditive, nonnegative, and absolutely homogeneous.*

The following lemma implies that Monte Carlo averages converge in our semi-norms with rate 1/2.

Lemma 3.4 (Linear combinations of iid random variables)**.** *Assume the setting in Subsection 3.1, let $k \in \mathbb{N}_0$, $n, Q \in \mathbb{N}$, $r_1,\ldots,r_n \in \mathbb{R}$, and let $V_1,\ldots,V_n \in \mathcal{M}(\mathcal{B}([0,T]\times\mathbb{R}^d)\otimes\mathcal{F},\mathcal{B}(\mathbb{R}))$ satisfy for all $(s,x)\in[0,T]\times\mathbb{R}^d$ that $V_1(s,x),\ldots,V_n(s,x)$ are integrable random variables which are independent and identically distributed and which are independent of W^0. Then*

$$\left\|\sum_{i=1}^n r_i(V_i-\mathbb{E}[V_i])\right\|_{k,Q} = \|(V_1-\mathbb{E}[V_1])\|_{k,Q}\sqrt{\sum_{i=1}^n |r_i|^2} \le \|V_1\|_{k,Q}\sqrt{\sum_{i=1}^n |r_i|^2}. \tag{28}$$

Proof. The definition (18) of the semi-norm and the fact that for all $(s,x)\in[0,T]\times\mathbb{R}^d$ it holds that $(V_i(s,x))_{i\in\{1,\ldots,n\}}$ are independent of W^0 and are independent and identically distributed imply that

$$\begin{aligned}
&\left\|\sum_{i=1}^n r_i(V_i-\mathbb{E}[V_i])\right\|_{k,Q}\\
&= \sum_{t\in[0,T]} \bar{q}^{k,Q}(t)\left[\sup_{s\in[t,T]}\sup_{u\in[0,s]}\sup_{z\in\mathbb{R}^d}\mathbb{E}\left[\mathbb{E}\left[\left|\sum_{i=1}^n r_i\Big(V_i(s,z+W_u^0)-\mathbb{E}\left[V_i(s,z+W_u^0)|W^0\right]\Big)\right|^2\middle|W^0\right]\right]\right]^{\frac12}\\
&= \sum_{t\in[0,T]} \bar{q}^{k,Q}(t)\left[\sup_{s\in[t,T]}\sup_{u\in[0,s]}\sup_{z\in\mathbb{R}^d}\mathbb{E}\left[\operatorname{Var}\left(\sum_{i=1}^n r_iV_i(s,z+W_u^0)|W^0\right)\right]\right]^{\frac12}\\
&= \sum_{t\in[0,T]} \bar{q}^{k,Q}(t)\left[\sup_{s\in[t,T]}\sup_{u\in[0,s]}\sup_{z\in\mathbb{R}^d}\mathbb{E}\left[\sum_{i=1}^n |r_i|^2\operatorname{Var}\Big(V_1(s,z+W_u^0)\Big|W^0\Big)\right]\right]^{\frac12}\\
&= \|V_1-\mathbb{E}[V_1]\|_{k,Q}\sqrt{\sum_{i=1}^n|r_i|^2}\\
&\le \sum_{t\in[0,T]} \bar{q}^{k,Q}(t)\left[\sup_{s\in[t,T]}\sup_{u\in[0,s]}\sup_{z\in\mathbb{R}^d}\mathbb{E}\left[\mathbb{E}\left[\left|V_1(s,z+W_u^0)\right|^2\middle|W^0\right]\right]\right]^{\frac12}\sqrt{\sum_{i=1}^n|r_i|^2}\\
&= \|V_1\|_{k,Q}\sqrt{\sum_{i=1}^n|r_i|^2}.
\end{aligned} \tag{29}$$

□

Lemma 3.5 (Lipschitz property)**.** *Assume the setting in Subsection 3.1, let $k\in\mathbb{N}_0$, $Q\in\mathbb{N}$, and let $U,V\in\mathcal{M}(\mathcal{B}([0,T]\times\mathbb{R}^d)\otimes\mathcal{F},\mathcal{B}(\mathbb{R}))$. Then*

$$\|F(U)-F(V)\|_{k,Q} \le L\|U-V\|_{k,Q}. \tag{30}$$

Proof. The definition (18) of the semi-norm and the global Lipschitz property (13) of F imply that

$$\begin{aligned}
\|F(U)-F(V)\|_{k,Q} &= \sum_{t\in[0,T]} \bar{q}^{k,Q}(t)\left[\sup_{s\in[t,T]}\sup_{u\in[0,s]}\sup_{z\in\mathbb{R}^d}\mathbb{E}\big[|(F(U))(s,z+W_u^0)-(F(V))(s,z+W_u^0)|^2\big]\right]^{\frac12}\\
&\le \sum_{t\in[0,T]} \bar{q}^{k,Q}(t)\left[\sup_{s\in[t,T]}\sup_{u\in[0,s]}\sup_{z\in\mathbb{R}^d}L^2\mathbb{E}\big[|U(s,z+W_u^0)-V(s,z+W_u^0)|^2\big]\right]^{\frac12}\\
&= L\|U-V\|_{k,Q}.
\end{aligned} \tag{31}$$

□

Lemma 3.6. *Assume the setting in Subsection 3.1, let* $k \in \mathbb{N}_0$*,* $Q \in \mathbb{N}$*, and let* $U \in \mathcal{M}(\mathcal{B}([0,T]\times\mathbb{R}^d)\otimes\mathcal{F}, \mathcal{B}(\mathbb{R}))$ *satisfy for all* $(s,x) \in [0,T]\times\mathbb{R}^d$ *that* $U(s,x)$ *and* W^0 *are independent. Then*

$$\left\| [0,T]\times\mathbb{R}^d \ni (s,z) \mapsto \sum_{t\in[s,T]} q^{Q,[s,T]}(t) U(t, z+W^0_t - W^0_s) \in \mathbb{R} \right\|_{k,Q} \leq \|U\|_{k+1,Q} \tag{32}$$

Proof. The definition (18) of the semi-norm, the triangle inequality, independence, Lemma 3.1, and the definition (16) of $\bar{q}^{k+1,Q}$ yield that

$$\begin{aligned}
&\left\| [0,T]\times\mathbb{R}^d \ni (s,z) \mapsto \sum_{r\in[s,T]} q^{Q,[s,T]}(r) U(r, z+W^0_r - W^0_s) \in \mathbb{R} \right\|_{k,Q} \\
&= \sum_{t\in[0,T]} \bar{q}^{k,Q}(t) \left[\sup_{s\in[t,T]} \sup_{u\in[0,s]} \sup_{z\in\mathbb{R}^d} \left(\mathbb{E}\left[\left| \sum_{r\in[s,T]} q^{Q,[s,T]}(r) U(r, z+W^0_u + W^0_r - W^0_s) \right|^2 \right] \right)^{\frac{1}{2}} \right] \\
&\leq \sum_{t\in[0,T]} \bar{q}^{k,Q}(t) \left[\sup_{s\in[t,T]} \sup_{u\in[0,s]} \sup_{z\in\mathbb{R}^d} \sum_{r\in[s,T]} q^{Q,[s,T]}(r) \left(\mathbb{E}\left[\left| U(r, z+W^0_u + W^0_r - W^0_s) \right|^2 \right] \right)^{\frac{1}{2}} \right] \\
&\leq \sum_{t\in[0,T]} \bar{q}^{k,Q}(t) \sup_{s\in[t,T]} \sum_{r\in[s,T]} q^{Q,[s,T]}(r) \left[\sup_{v\in[r,T]} \sup_{u\in[0,v]} \sup_{z\in\mathbb{R}^d} \mathbb{E}\left[\left| U(v, z+W^0_u) \right|^2 \right] \right]^{\frac{1}{2}} \\
&= \sum_{t\in[0,T]} \bar{q}^{k,Q}(t) \sum_{r\in[t,T]} q^{Q,[t,T]}(r) \left[\sup_{v\in[r,T]} \sup_{u\in[0,v]} \sup_{z\in\mathbb{R}^d} \mathbb{E}\left[\left| U(v, z+W^0_u) \right|^2 \right] \right]^{\frac{1}{2}} \\
&= \sum_{r\in[0,T]} \bar{q}^{k+1,Q}(r) \left[\sup_{v\in[r,T]} \sup_{u\in[0,v]} \sup_{z\in\mathbb{R}^d} \mathbb{E}\left[\left| U(v, z+W^0_u) \right|^2 \right] \right]^{\frac{1}{2}} \\
&= \|U\|_{k+1,Q}.
\end{aligned} \tag{33}$$

□

Lemma 3.7 (Monotonicity)**.** *Assume the setting in Subsection 3.1, let* $k \in \mathbb{N}_0$*,* $Q \in \mathbb{N}$*, let* $\mathcal{G} \subseteq \mathcal{F}$ *be a* σ*-algebra, and let* $U, V \in \mathcal{M}(\mathcal{B}([0,T]\times\mathbb{R}^d)\otimes\mathcal{F}, \mathcal{B}(\mathbb{R}))$ *satisfy* $|U| \leq |V|$*. Then*

$$\left\| \mathbb{E}\big[|U|\big|\mathcal{G}\big] \right\|_{k,Q} \leq \|U\|_{k,Q} \leq \|V\|_{k,Q}. \tag{34}$$

Proof. The definition (18) of the semi-norm, Jensen's inequality, and the hypothesis that $|U| \leq |V|$ imply that

$$\begin{aligned}
\left\| \mathbb{E}\big[|U|\big|\mathcal{G}\big] \right\|_{k,Q} &= \sum_{t\in[0,T]} \bar{q}^{k,Q}(t) \left[\sup_{s\in[t,T]} \sup_{u\in[0,s]} \sup_{z\in\mathbb{R}^d} \left(\mathbb{E}\left[\left| \mathbb{E}\big[|U(s,z+W^0_u)|\big|\mathcal{G}\big] \right|^2 \right] \right)^{\frac{1}{2}} \right] \\
&\leq \sum_{t\in[0,T]} \bar{q}^{k,Q}(t) \left[\sup_{s\in[t,T]} \sup_{u\in[0,s]} \sup_{z\in\mathbb{R}^d} \left(\mathbb{E}\left[\left| U(s,z+W^0_u) \right|^2 \right] \right)^{\frac{1}{2}} \right] = \|U\|_{k,Q} \\
&\leq \sum_{t\in[0,T]} \bar{q}^{k,Q}(t) \left[\sup_{s\in[t,T]} \sup_{u\in[0,s]} \sup_{z\in\mathbb{R}^d} \left(\mathbb{E}\left[\left| V(s,z+W^0_u) \right|^2 \right] \right)^{\frac{1}{2}} \right] = \|V\|_{k,Q}.
\end{aligned} \tag{35}$$

□

The following lemma specifies the values of our semi-norms of constant functions. It follows directly from the definition (18) of the semi-norms and from Lemma 3.2. Its proof is therefore omitted.

Lemma 3.8 (Seminorm of constants)**.** *Assume the setting in Subsection 3.1 and let* $Q \in \mathbb{N}$*,* $k \in \mathbb{N}_0 \cap [0, 2Q-1]$*. Then* $\|1\|_{k,Q} = \frac{T^k}{k!}$*.*

3.6 Error analysis for multilevel Picard iteration

Lemma 3.9 (Approximations are integrable)**.** *Assume the setting in Subsection 3.1, let* $z \in \mathbb{R}^d$*,* $M, Q \in \mathbb{N}$*, and assume for all* $s \in [0,T]$*,* $t \in [s,T]$ *that* $\mathbb{E}\big[|g(z+W^0_t)| + |(F(0))(t, z+W^0_s)|\big] < \infty$*. Then* *(i) for all $n \in \mathbb{N}_0$, $\theta \in \Theta$, $s \in [0,T]$, $t \in [s,T]$ it holds that*

$$\mathbb{E}\Big[\big|U^\theta_{n,M,Q}(t, z + W^0_s)\big| + \big|\big(F(U^\theta_{n,M,Q})\big)(t, z + W^0_s)\big|\Big] < \infty \tag{36}$$

and

(ii) for all $n \in \mathbb{N}_0$, $\theta \in \Theta$, $s \in [0,T]$ it holds that

$$\mathbb{E}\big[U^\theta_{n+1,M,Q}(s,z)\big] = \mathbb{E}\left[g(z + W^0_{T-s}) + \sum_{t\in[s,T]} q^{Q,[s,T]}(t)\big(F(U^\theta_{n,M,Q})\big)(t, z + W^0_{t-s})\right]. \tag{37}$$

Proof. We prove (i) by induction on $n \in \mathbb{N}_0$. For the base case $n = 0$ we note that for all $\theta \in \Theta$, $s \in [0,T]$, $t \in [s,T]$ it holds that

$$\mathbb{E}\Big[\big|U^\theta_{0,M,Q}(t, z + W^0_s)\big| + \big|\big(F(U^\theta_{0,M,Q})\big)(t, z + W^0_s)\big|\Big] = \mathbb{E}\Big[\big|((F(0))(t, z + W^0_s)\big|\Big] < \infty. \tag{38}$$

This establishes (i) in the base case $n = 0$. For the induction step $\mathbb{N}_0 \ni n \to n+1 \in \mathbb{N}$ let $n \in \mathbb{N}_0$ and assume that (i) holds for $n = 0$, $n = 1$, …, $n = n$. The induction hypothesis and (17) imply that for all $\theta \in \Theta$, $s \in [0,T]$, $t \in [s,T]$ it holds that

$$\begin{aligned}
&\mathbb{E}\Big[\big|U^\theta_{n+1,M,Q}(t, z + W^0_s)\big|\Big] \le \mathbb{E}\Big[\big|g(z + W^0_{T-t+s})\big|\Big] \\
&\quad + \sum_{l=0}^{n} \sum_{r\in[t,T]} \frac{q^{Q,[t,T]}(r)}{M^{n+1-l}} \sum_{i=1}^{M^{n+1-l}} \sum_{k\in\{l-1,l\}\cap\mathbb{N}_0} \max_{j\in\{-l,l\}} \mathbb{E}\Big[\big|\big(F(U^{(\theta,j,i,r)}_{k,M,Q})\big)(r, z + W^0_{s+r-t})\big|\Big] < \infty.
\end{aligned} \tag{39}$$

Combining this with (13) proves for all $\theta \in \Theta$, $s \in [0,T]$, $t \in [s,T]$ that

$$\begin{aligned}
&\mathbb{E}\Big[\big|\big(F(U^\theta_{n+1,M,Q})\big)(t, z + W^0_s)\big|\Big] \\
&\le \mathbb{E}\Big[\big|\big(F(U^\theta_{n+1,M,Q})\big)(t, z + W^0_s) - \big(F(0)\big)(t, z + W^0_s)\big|\Big] + \mathbb{E}\Big[\big|\big(F(0)\big)(t, z + W^0_s)\big|\Big] \\
&\le L\mathbb{E}\Big[\big|U^\theta_{n+1,M,Q}(t, z + W^0_s)\big|\Big] + \mathbb{E}\Big[\big|\big(F(0)\big)(t, z + W^0_s)\big|\Big] < \infty.
\end{aligned} \tag{40}$$

This finishes the induction step $\mathbb{N}_0 \ni n \to n+1 \in \mathbb{N}$. Induction hence establishes (i). Next we note that (17), the fact that $(U^\theta_{n,M,Q})_{n\in\mathbb{N}_0}$, $\theta \in \Theta$, are identically distributed, and a telescope argument yield that for all $n \in \mathbb{N}_0$, $\theta \in \Theta$, $s \in [0,T]$ it holds that

$$\begin{aligned}
&\mathbb{E}\big[U^\theta_{n+1,M,Q}(s,z))\big] - \mathbb{E}\big[g(z + W^0_{T-s})\big] \\
&= \sum_{l=0}^{n} \sum_{v\in[s,T]} q^{Q,[s,T]}(v)\, \mathbb{E}\Big[\big(F(U^0_{l,M,Q}) - \mathbb{1}_{\mathbb{N}}(l)F(U^0_{l-1,M,Q})\big)(v, z + W^0_v - W^0_s)\Big] \\
&= \sum_{v\in[s,T]} q^{Q,[s,T]}(v)\, \mathbb{E}\Big[\big(F(U^0_{n,M,Q})\big)(v, z + W^0_v - W^0_s)\Big] \\
&= \mathbb{E}\left[\sum_{v\in[s,T]} q^{Q,[s,T]}(v)\, \big(F(U^\theta_{n,M,Q})\big)(v, z + W^0_{v-s})\right].
\end{aligned} \tag{41}$$

This establishes (ii). The proof of Lemma 3.9 is thus completed. ◻

Lemma 3.10 (Nonlinear Feynman-Kac formula)**.** *Assume the setting in Subsection 3.1, let $z \in \mathbb{R}^d$, and assume for all $s \in [0,T]$ that*

$$\mathbb{E}\left[\sup_{t\in[s,T]} \big|u^\infty(t, z + W^0_{t-s})\big| + \int_s^T \big|\big(F(0)\big)(t, z + W^0_{t-s})\big|\, dt\right] < \infty. \tag{42}$$

Then

(i) for all $s \in [0,T]$ it holds that

$$\mathbb{E}\left[\sup_{t\in[s,T]} |u^\infty(t, z + W^0_t - W^0_s)| + \int_s^T |(F(u^\infty))(t, z + W^0_t - W^0_s)|\, dt\right] < \infty \tag{43}$$

and

(ii) for all $s \in [0,T]$ it holds that

$$u^\infty(s,z) - \mathbb{E}\big[g(z+W^0_{T-s})\big] = \mathbb{E}\left[\int_s^T (F(u^\infty))(t,z+W^0_{t-s})\,dt\right]. \tag{44}$$

Proof. Note that (13) and (42) imply (i). Next Itô's formula and the PDE (14) ensure that for all $s \in [0,T]$, $t \in [s,T]$ it holds $\mathbb{P}$-a.s. that

$$\begin{aligned} & u^\infty(t,z+W^0_t-W^0_s) - u^\infty(s,z) \\ &= \int_s^t \big(\tfrac{\partial}{\partial r}u^\infty + \tfrac{1}{2}\Delta_y u^\infty\big)(r,z+W^0_r-W^0_s)\,dr + \int_s^t \langle (\nabla_y u^\infty)(r,z+W^0_r-W^0_s),\, dW^0_r\rangle \\ &= -\int_s^t (F(u^\infty))(r,z+W^0_r-W^0_s)\,dr + \int_s^t \langle (\nabla_y u^\infty)(r,z+W^0_r-W^0_s),\, dW^0_r\rangle. \end{aligned} \tag{45}$$

This and (43) show that for all $s \in [0,T]$ it holds that $\mathbb{E}\big[\sup_{t\in[s,T]} \big| \int_s^t \langle (\nabla_y u^\infty)(r,z+W^0_r-W^0_s),\, dW^0_r\rangle \big|\big] < \infty$. This ensures that $\mathbb{E}\big[\int_s^T \langle (\nabla_y u^\infty)(t,z+W^0_t-W^0_s),\, dW^0_t\rangle\big] = 0$. This and (45) prove for all $s \in [0,T]$ that

$$u^\infty(s,z) - \mathbb{E}[g(z+W^0_{T-s})] = u^\infty(s,z) - \mathbb{E}[u^\infty(T,z+W^0_T-W^0_s)] = \mathbb{E}\left[\int_s^T (F(u^\infty))(t,z+W^0_{t-s})\,dt\right]. \tag{46}$$

This finishes the proof of Lemma 3.10. □

Theorem 3.11. *Assume the setting in Subsection 3.1, let $M, Q \in \mathbb{N}$, $N \in \mathbb{N} \cap [1, 2Q-1]$, $\theta \in \Theta$, and assume for all $z \in \mathbb{R}^d$, $s \in [0,T]$ that*

$$\mathbb{E}\left[\sup_{t\in[s,T]} \big|u^\infty(t,z+W^0_{t-s})\big| + \int_s^T \big|\big(F(0)\big)(t,z+W^0_{t-s})\big|\,dt\right] < \infty. \tag{47}$$

Then we have

$$\begin{aligned} &\big\|U^\theta_{N,M,Q} - u^\infty\big\|_{0,Q} \leq (1+2L)^{N-1} \Bigg\{ L \sup_{i\in\{1,2,\ldots,N\}} \tfrac{\|u^\infty\|_{i,Q}}{\sqrt{M^{N-i}}} \\ &+ \Big[\sup_{i\in\{0,1,\ldots,N-1\}} \tfrac{T^i}{i!\sqrt{M^{N-i}}}\Big] \left[\sup_{z\in\mathbb{R}^d}\sup_{s\in[0,T]} \big\|g(z+W^0_s)\big\|_{L^2(\mathbb{P};\mathbb{R})} + T \sup_{z\in\mathbb{R}^d}\sup_{r,s\in[0,T]} \big\|(F(0))(r,z+W^0_s)\big\|_{L^2(\mathbb{P};\mathbb{R})}\right] \\ &+ e^T \sup_{\substack{t\in[0,T],\\ r\in[0,t],\\ z\in\mathbb{R}^d}} \left\|\mathbb{E}\Big[\sum_{s\in[t,T]} q^{Q,[t,T]}(s)\,(F(u^\infty))(s,z+W^0_{r+s-t}) - \int_t^T (F(u^\infty))(s,z+W^0_{r+s-t})\,ds \,\Big|\, W^0_r\Big]\right\|_{L^2(\mathbb{P};\mathbb{R})} \Bigg\}. \end{aligned} \tag{48}$$

Proof. Throughout this proof assume w.l.o.g. that the right-hand side of (48) is finite, assume w.l.o.g. that $\theta = 0$ (the case $\theta \neq 0$ follows from the case $\theta = 0$), let $\varepsilon \in [0,\infty)$ be the real number given by

$$\varepsilon = \sup_{\substack{t\in[0,T],\, u\in[0,t]\\ z\in\mathbb{R}^d}} \left\|\mathbb{E}\Big[\sum_{s\in[t,T]} q^{Q,[t,T]}(s)\,(F(u^\infty))(s,z+W^0_{u+s-t}) - \int_t^T (F(u^\infty))(s,z+W^0_{u+s-t})\,ds \,\Big|\, W^0_u\Big]\right\|_{L^2(\mathbb{P};\mathbb{R})},$$

and let $(e_n)_{n\in\{0,1,\ldots,N\}} \subseteq [0,\infty]$ be the extended real numbers which satisfy for all $n \in \{0,1,\ldots,N\}$ that

$$e_n = \sup\Big\{\sqrt{M^{-j}}\,\big\|U^0_{n,M,Q} - u^\infty\big\|_{k,Q} : k,j \in \mathbb{N}_0, k+j+n = N\Big\}. \tag{49}$$

First, we analyze the *Monte Carlo error.* Item (i) of Lemma 3.9 shows for all $n \in \mathbb{N}_0$, $(t,z) \in [0,T]\times\mathbb{R}^d$, $s \in [0,t]$ that $\mathbb{E}[|U^0_{n,M,Q}(t,z+W^0_s)|] < \infty$. The triangle inequality, independence, Lemma 3.4, Lemma 3.7, Lemma 3.6, Lemma 3.8, and Lemma 3.5 imply that for all $n \in \mathbb{N}$, $k \in \mathbb{N}_0$ it holds that

$$
\begin{aligned}
&\left\| U^0_{n,M,Q} - \mathbb{E}\big[U^0_{n,M,Q}\big] \right\|_{k,Q} \\
&\leq \left\| [0,T]\times\mathbb{R}^d \ni (t,z) \mapsto M^{-n} \sum_{i=1}^{M^n} \Big(g(z+W^{(\theta,0,-i)}_T - W^{(\theta,0,-i)}_t) - \mathbb{E}\big[g(z+W^{(\theta,0,-i)}_T - W^{(\theta,0,-i)}_t)\big] \Big) \in \mathbb{R} \right\|_{k,Q} \\
&\quad + \sum_{l=0}^{n-1} \Bigg\| [0,T]\times\mathbb{R}^d \ni (t,z) \mapsto \\
&\qquad M^{l-n} \sum_{i=1}^{M^{n-l}} \sum_{r\in[t,T]} q^{Q,[t,T]}(r) \big(F(U^{(\theta,l,i,r)}_{l,M,Q}) - \mathbb{1}_{\mathbb{N}}(l) F(U^{(\theta,-l,i,r)}_{l-1,M,Q}) \big)(r, z+W^{(\theta,l,i)}_r - W^{(\theta,l,i)}_t) \\
&\qquad - M^{l-n} \sum_{i=1}^{M^{n-l}} \sum_{r\in[t,T]} q^{Q,[t,T]}(r) \mathbb{E}\left[\big(F(U^{(\theta,l,i,r)}_{l,M,Q}) - \mathbb{1}_{\mathbb{N}}(l) F(U^{(\theta,-l,i,r)}_{l-1,M,Q}) \big)(r, z+W^{(\theta,l,i)}_r - W^{(\theta,l,i)}_t) \right] \in \mathbb{R} \Bigg\|_{k,Q} \\
&\leq \tfrac{1}{\sqrt{M^n}} \left\| [0,T]\times\mathbb{R}^d \ni (t,z) \mapsto g(z+W^0_T - W^0_t) \in \mathbb{R} \right\|_{k,Q} \\
&\quad + \sum_{l=0}^{n-1} \tfrac{1}{\sqrt{M^{n-l}}} \Bigg\| [0,T]\times\mathbb{R}^d \ni (t,z) \mapsto \sum_{r\in[t,T]} q^{Q,[t,T]}(r) \\
&\qquad \cdot \big(F(U^{(\theta,1,1,r)}_{l,M,Q}) - \mathbb{1}_{\mathbb{N}}(l) F(U^{(\theta,-1,1,r)}_{l-1,M,Q}) \big)(r, z+W^0_r - W^0_t) \in \mathbb{R} \Bigg\|_{k,Q} \\
&\leq \tfrac{1}{\sqrt{M^n}} \sup_{z\in\mathbb{R}^d} \sup_{s\in[0,T]} \left\| g(z+W^0_s) \right\|_{L^2(\mathbb{P};\mathbb{R})} \|1\|_{k,Q} + \sum_{l=0}^{n-1} \tfrac{1}{\sqrt{M^{n-l}}} \left\| F(U^{(\theta,1,1,0)}_{l,M,Q}) - \mathbb{1}_{\mathbb{N}}(l) F(U^{(\theta,-1,1,0)}_{l-1,M,Q}) \right\|_{k+1,Q} \\
&\leq \tfrac{1}{\sqrt{M^n}} \sup_{z\in\mathbb{R}^d} \sup_{s\in[0,T]} \left\| g(z+W^0_s) \right\|_{L^2(\mathbb{P};\mathbb{R})} \tfrac{T^k}{k!} + \tfrac{1}{\sqrt{M^n}} \|F(0)\|_{k+1,Q} + L \sum_{l=1}^{n-1} \tfrac{1}{\sqrt{M^{n-l}}} \left\| U^{(\theta,1,1,0)}_{l,M,Q} - U^{(\theta,-1,1,0)}_{l-1,M,Q} \right\|_{k+1,Q} \\
&\leq \tfrac{1}{\sqrt{M^n}} \sup_{z\in\mathbb{R}^d} \sup_{s\in[0,T]} \left\| g(z+W^0_s) \right\|_{L^2(\mathbb{P};\mathbb{R})} \tfrac{T^k}{k!} + \tfrac{1}{\sqrt{M^n}} \sup_{z\in\mathbb{R}^d} \sup_{r,s\in[0,T]} \left\| (F(0))(r, z+W^0_s) \right\|_{L^2(\mathbb{P};\mathbb{R})} \tfrac{T^{k+1}}{(k+1)!} \\
&\quad + L \sum_{l=0}^{n-1} \left(\tfrac{\mathbb{1}_{(0,n)}(l)}{\sqrt{M^{n-l}}} + \tfrac{\mathbb{1}_{(-\infty,n-1)}(l)}{\sqrt{M^{n-l-1}}} \right) \left\| U^0_{l,M,Q} - u^\infty \right\|_{k+1,Q}.
\end{aligned} \tag{50}
$$

Next we analyze the *time discretization error*. Item (ii) of Lemma 3.9 and Item (ii) of Lemma 3.10 ensure that for all $n \in \mathbb{N}$, $s \in [0,T]$, $z \in \mathbb{R}^d$ it holds $\mathbb{P}$-a.s. that

$$
\mathbb{E}\big[U^0_{n,M,Q}(s,z))\big] - u^\infty(s,z) = \mathbb{E}\left[\sum_{t\in[s,T]} q^{Q,[s,T]}(t) \big(F(U^0_{n-1,M,Q}) \big)(t, z+W^0_{t-s}) - \int_s^T (F(u^\infty))(t, z+W^0_{t-s})\, dt \right]. \tag{51}
$$

This, the triangle inequality, Lemma 3.7, Lemma 3.6, Lemma 3.5, and Lemma 3.8 demonstrate for all $n \in \mathbb{N}$, $k \in \mathbb{N}_0 \cap [0, 2Q-1]$ that

$$
\left\| \mathbb{E}\big[U^0_{n,M,Q}\big] - u^\infty \right\|_{k,Q} \tag{52}
$$

$$
\begin{aligned}
&\leq \left\| [0,T]\times\mathbb{R}^d \ni (s,z) \mapsto \mathbb{E}\left[\sum_{t\in[s,T]} q^{Q,[s,T]}(t) \big(F(U^0_{n-1,M,Q}) - F(u^\infty) \big)(t, z+W^0_{t-s}) \right] \right\|_{k,Q} \\
&\quad + \left\| [0,T]\times\mathbb{R}^d \ni (s,z) \mapsto \mathbb{E}\left[\sum_{t\in[s,T]} q^{Q,[s,T]}(t) \big(F(u^\infty) \big)(t, z+W^0_{t-s}) - \int_s^T (F(u^\infty))(t, z+W^0_{t-s})\, dt \right] \right\|_{k,Q}
\end{aligned}
$$

$$
\leq \left\| [0,T]\times\mathbb{R}^d \ni (s,z) \mapsto \sum_{t\in[s,T]} q^{Q,[s,T]}(t) \big(F(U^0_{n-1,M,Q}) - F(u^\infty) \big)(t, z+W^0_{t-s}) \right\|_{k,Q} + \varepsilon \|1\|_{k,Q} \tag{53}
$$

$$
\begin{aligned}
&\leq \left\| F(U^0_{n-1,M,Q}) - F(u^\infty) \right\|_{k+1,Q} + \varepsilon \|1\|_{k,Q} \\
&\leq L \left\| U^0_{n-1,M,Q} - u^\infty \right\|_{k+1,Q} + \varepsilon \tfrac{T^k}{k!}.
\end{aligned}
$$

In the next step we combine the established bounds for the Monte Carlo error and the time discretization error to obtain a bound for the *global error*. More formally, observe that (50) and (52) ensure that for all $n \in \mathbb{N}$,

$k \in \mathbb{N}_0 \cap [0, 2Q-1]$ it holds that

$$\begin{aligned}
&\big\|U^0_{n,M,Q} - u^\infty\big\|_{k,Q} \le \big\|U^0_{n,M,Q} - \mathbb{E}\big[U^0_{n,M,Q}\big]\big\|_{k,Q} + \big\|\mathbb{E}\big[U^0_{n,M,Q}\big] - u^\infty\big\|_{k,Q} \\
&\le \tfrac{1}{\sqrt{M^n}} \tfrac{T^k}{k!} \Big[\sup_{z\in\mathbb{R}^d} \sup_{s\in[0,T]} \big\|g(z+W^0_s)\big\|_{L^2(\mathbb{P};\mathbb{R})} + T \sup_{z\in\mathbb{R}^d} \sup_{r,u\in[0,T]} \big\|(F(0))(r, z+W^0_u)\big\|_{L^2(\mathbb{P};\mathbb{R})} \Big] \\
&\quad + L \sum_{l=0}^{n-1} \Big(\tfrac{\mathbb{1}_{(0,n)}(l)}{\sqrt{M^{n-l-1}}} + \tfrac{\mathbb{1}_{(-\infty,n-1)}(l)}{\sqrt{M^{n-l-1}}} \Big) \big\|U^0_{l,M,Q} - u^\infty\big\|_{k+1,Q} + L \big\|U^0_{n-1,M,Q} - u^\infty\big\|_{k+1,Q} + \varepsilon \tfrac{T^k}{k!} \\
&= \tfrac{1}{\sqrt{M^n}} \tfrac{T^k}{k!} \Big[\sup_{z\in\mathbb{R}^d} \sup_{s\in[0,T]} \big\|g(z+W^0_s)\big\|_{L^2(\mathbb{P};\mathbb{R})} + T \sup_{z\in\mathbb{R}^d} \sup_{r,u\in[0,T]} \big\|(F(0))(r, z+W^0_u)\big\|_{L^2(\mathbb{P};\mathbb{R})} \Big] \\
&\quad + L \tfrac{\|u^\infty\|_{k+1,Q}}{\sqrt{M^{n-1}}} + 2L \sum_{l=1}^{n-1} \tfrac{1}{\sqrt{M^{n-l-1}}} \big\|U^0_{l,M,Q} - u^\infty\big\|_{k+1,Q} + \varepsilon \tfrac{T^k}{k!}.
\end{aligned} \tag{54}$$

Hence, we obtain that for all $j \in \mathbb{N}_0$, $n \in \mathbb{N}$, $k \in \mathbb{N}_0 \cap [0, 2Q-1]$ it holds that

$$\begin{aligned}
&\sqrt{M^{-j}} \big\|U^0_{n,M,Q} - u^\infty\big\|_{k,Q} \le \tfrac{L\|u^\infty\|_{k+1,Q}}{\sqrt{M^{n+j-1}}} + \varepsilon \tfrac{T^k}{k!\sqrt{M^j}} + 2L \sum_{l=1}^{n-1} \sqrt{M^{-j-n+l+1}} \big\|U^0_{l,M,Q} - u^\infty\big\|_{k+1,Q} \\
&\quad + \tfrac{T^k}{k!\sqrt{M^{j+n}}} \Big[\sup_{z\in\mathbb{R}^d} \sup_{s\in[0,T]} \big\|g(z+W^0_s)\big\|_{L^2(\mathbb{P};\mathbb{R})} + T \sup_{z\in\mathbb{R}^d} \sup_{r,u\in[0,T]} \big\|(F(0))(r, z+W^0_u)\big\|_{L^2(\mathbb{P};\mathbb{R})} \Big].
\end{aligned} \tag{55}$$

This shows for all $n \in \{1, 2, \ldots, N\}$ that

$$\begin{aligned}
&e_n \le L \sup_{k\in\{0,1,\ldots,N-1\}} \tfrac{\|u^\infty\|_{k+1,Q}}{\sqrt{M^{N-k-1}}} + \varepsilon e^T + 2L \sum_{l=1}^{n-1} e_l \\
&\quad + \Big[\sup_{i\in\{0,1,\ldots,N-1\}} \tfrac{T^i}{i!\sqrt{M^{N-i}}} \Big] \Big[\sup_{z\in\mathbb{R}^d} \sup_{s\in[0,T]} \big\|g(z+W^0_s)\big\|_{L^2(\mathbb{P};\mathbb{R})} + T \sup_{z\in\mathbb{R}^d} \sup_{r,u\in[0,T]} \big\|(F(0))(r, z+W^0_u)\big\|_{L^2(\mathbb{P};\mathbb{R})} \Big].
\end{aligned} \tag{56}$$

Combining this with the discrete Gronwall-type inequality in Agarwal [1, Corollary 4.1.2] proves that

$$\begin{aligned}
&\big\|U^0_{N,M,Q} - u^\infty\big\|_{0,Q} = e_N \le (1+2L)^{N-1} \Big\{ L \sup_{i\in\{1,2,\ldots,N\}} \tfrac{\|u^\infty\|_{i,Q}}{\sqrt{M^{N-i}}} + \varepsilon e^T \\
&+ \Big[\sup_{i\in\{0,1,\ldots,N-1\}} \tfrac{T^i}{i!\sqrt{M^{N-i}}} \Big] \Big[\sup_{z\in\mathbb{R}^d} \sup_{s\in[0,T]} \big\|g(z+W^0_s)\big\|_{L^2(\mathbb{P};\mathbb{R})} + T \sup_{z\in\mathbb{R}^d} \sup_{r,u\in[0,T]} \big\|(F(0))(r, z+W^0_u)\big\|_{L^2(\mathbb{P};\mathbb{R})} \Big] \Big\}.
\end{aligned} \tag{57}$$

This completes the proof of Theorem 3.11. ☐

In the proof of the following result, Corollary 3.12, an upper bound for the quadrature error on the right-hand side of (48) is derived under the hypothesis that the solution of the PDE is sufficiently smooth and regular.

Corollary 3.12. *Assume the setting in Subsection 3.1, assume that $u^\infty \in C^\infty([0,T] \times \mathbb{R}^d, \mathbb{R})$, assume for all $k \in \mathbb{N}_0$, $x \in \mathbb{R}^d$, $t \in [0,T]$ that*

$$\mathbb{E}\bigg[\sup_{s\in[t,T]} \big| \big((\tfrac{\partial}{\partial r} + \tfrac{1}{2}\Delta_y)^k u^\infty \big)(s, x + W^0_{s-t}) \big| \bigg] < \infty, \tag{58}$$

and let $M, Q \in \mathbb{N}$, $N \in \mathbb{N} \cap [1, 2Q)$. Then it holds for all $\theta \in \Theta$ that

$$\begin{aligned}
&\big\|U^\theta_{N,M,Q} - u^\infty\big\|_{0,Q} \le (1+2L)^{N-1} \bigg\{ L \sup_{i\in\{1,2,\ldots,N\}} \tfrac{\|u^\infty\|_{i,Q}}{\sqrt{M^{N-i}}} \\
&+ \bigg[\sup_{i\in\{0,1,\ldots,N-1\}} \tfrac{T^i}{i!\sqrt{M^{N-i}}} \bigg] \bigg[\sup_{z\in\mathbb{R}^d} \sup_{s\in[0,T]} \big\|g(z+W^0_s)\big\|_{L^2(\mathbb{P};\mathbb{R})} + T \sup_{z\in\mathbb{R}^d} \sup_{r,u\in[0,T]} \Big\|(F(0))(r, z+W^0_u)\Big\|_{L^2(\mathbb{P};\mathbb{R})} \bigg] \\
&+ e^T \sup_{t\in[0,T]} \bigg[\sup_{u\in[0,t]} \sup_{z\in\mathbb{R}^d} \bigg\| \sup_{s\in[t,T]} \big| \mathbb{E}\big[\big((\tfrac{\partial}{\partial r} + \tfrac{1}{2}\Delta_y)^{2Q+1} u^\infty \big)(s, x + W^0_{s-t}) \big] \big| \Big|_{x=z+W^0_u} \bigg\|_{L^2(\mathbb{P};\mathbb{R})} \tfrac{[Q!]^4 (T-t)^{2Q+1}}{(2Q+1)[(2Q)!]^3} \bigg] \bigg\}.
\end{aligned} \tag{59}$$

Proof. Throughout this proof assume w.l.o.g. that $\sup_{z\in\mathbb{R}^d} \sup_{t,s\in[0,T]} \mathbb{E}\big[|g(z+W^0_t)| + |(F(0))(t, z+W^0_s)| \big] < \infty$ (otherwise the right-hand side of (59) is infinite and the proof of (59) is clear). Observe that (58) and the dominated convergence theorem ensure that for every $k \in \mathbb{N}_0$, $x \in \mathbb{R}^d$, $t \in [0,T]$ it holds that the function

$$[t,T] \ni s \mapsto \mathbb{E}\big[\big((\tfrac{\partial}{\partial r} + \tfrac{1}{2}\Delta_y)^k u^\infty \big)(s, x + W^0_{s-t}) \big] \in \mathbb{R} \tag{60}$$

is continuous. The assumption that $u^\infty \in C^\infty([0,T]\times\mathbb{R}^d,\mathbb{R})$ and Itô's formula imply that for all $x\in\mathbb{R}^d$, $t\in[0,T]$, $s\in[t,T]$, $k\in\mathbb{N}$ it holds $\mathbb{P}$-a.s. that

$$\begin{aligned}&\big((\tfrac{\partial}{\partial r}+\tfrac{1}{2}\Delta_y)^k u^\infty\big)(s,x+W^0_s-W^0_t)-\big((\tfrac{\partial}{\partial r}+\tfrac{1}{2}\Delta_y)^k u^\infty\big)(t,x)\\&=\int_t^s\big((\tfrac{\partial}{\partial r}+\tfrac{1}{2}\Delta_y)^{k+1}u^\infty\big)(v,x+W^0_v-W^0_t)\,dv+\int_t^s\big\langle\big(\nabla_y(\tfrac{\partial}{\partial r}+\tfrac{1}{2}\Delta_y)^k u^\infty\big)(v,x+W^0_v-W^0_t),\,dW^0_v\big\rangle.\end{aligned}\tag{61}$$

This and (58) show that for all $x\in\mathbb{R}^d$, $t\in[0,T]$, $k\in\mathbb{N}$ it holds that $\mathbb{E}\big[\sup_{s\in[t,T]}\big|\int_t^s\big\langle\big(\nabla_y(\frac{\partial}{\partial r}+\frac{1}{2}\Delta_y)^k u^\infty\big)(v,x+W^0_v-W^0_t),\,dW^0_v\big\rangle\big|\big]<\infty$. This implies that for all $x\in\mathbb{R}^d$, $t\in[0,T]$, $s\in[t,T]$, $k\in\mathbb{N}$ it holds that $\mathbb{E}\big[\int_t^s\big\langle\big(\nabla_y(\frac{\partial}{\partial r}+\frac{1}{2}\Delta_y)^k u^\infty\big)(v,x+W^0_v-W^0_t),\,dW^0_v\big\rangle\big]=0$. This, (61), and Fubini's theorem show that for all $x\in\mathbb{R}^d$, $t\in[0,T]$, $s\in[t,T]$, $k\in\mathbb{N}$ it holds that

$$\begin{aligned}&\mathbb{E}\big[\big((\tfrac{\partial}{\partial r}+\tfrac{1}{2}\Delta_y)^k u^\infty\big)(s,x+W^0_s-W^0_t)\big]-\big((\tfrac{\partial}{\partial r}+\tfrac{1}{2}\Delta_y)^k u^\infty\big)(t,x)\\&=\int_t^s\mathbb{E}\big[\big((\tfrac{\partial}{\partial r}+\tfrac{1}{2}\Delta_y)^{k+1}u^\infty\big)(v,x+W^0_v-W^0_t)\big]\,dv.\end{aligned}\tag{62}$$

Equation (62) (with $k=1$) together with (60) (with $k=2$) implies for every $x\in\mathbb{R}^d$, $t\in[0,T)$ that the function $[t,T]\ni s\mapsto\mathbb{E}\big[\big((\frac{\partial}{\partial r}+\frac{1}{2}\Delta_y)u^\infty\big)(s,x+W^0_s-W^0_t)\big]\in\mathbb{R}$ is continuously differentiable. Induction, (60), and (62) prove that for every $x\in\mathbb{R}^d$, $t\in[0,T]$ it holds that the function $[t,T]\ni s\mapsto\mathbb{E}\big[\big((\frac{\partial}{\partial r}+\frac{1}{2}\Delta_y)u^\infty\big)(s,x+W^0_s-W^0_t)\big]\in\mathbb{R}$ is infinitely often differentiable. This, induction, and (62) demonstrate that for all $k\in\mathbb{N}$, $x\in\mathbb{R}^d$, $t\in[0,T)$, $s\in[t,T]$ it holds that

$$\tfrac{\partial^k}{\partial s^k}\mathbb{E}\big[\big((\tfrac{\partial}{\partial r}+\tfrac{1}{2}\Delta_y)u^\infty\big)(s,x+W^0_s-W^0_t)\big]=\mathbb{E}\big[\big((\tfrac{\partial}{\partial r}+\tfrac{1}{2}\Delta_y)^{k+1}u^\infty\big)(s,x+W^0_s-W^0_t)\big].\tag{63}$$

Equation (14) and the error representation for the Gauß-Legendre quadrature rule (see, e.g., [14, Display (2.7.12)]) imply for all $x\in\mathbb{R}^d$, $t\in[0,T)$ that there exists a real number $\xi\in[t,T]$ such that

$$\begin{aligned}&\sum_{s\in[t,T]}q^{Q,[t,T]}(s)\mathbb{E}\big[(F(u^\infty))(s,x+W^0_s-W^0_t)\big]-\int_t^T\mathbb{E}\big[(F(u^\infty))(s,x+W^0_s-W^0_t)\big]\,ds\\&=\int_t^T\mathbb{E}\big[\big((\tfrac{\partial}{\partial r}+\tfrac{1}{2}\Delta_y)u^\infty\big)(s,x+W^0_s-W^0_t)\big]\,ds-\sum_{s\in[t,T]}q^{Q,[t,T]}(s)\mathbb{E}\big[\big((\tfrac{\partial}{\partial r}+\tfrac{1}{2}\Delta_y)u^\infty\big)(s,x+W^0_s-W^0_t)\big]\\&=\Big(\tfrac{\partial^{2Q}}{\partial s^{2Q}}\mathbb{E}\,\big[\big((\tfrac{\partial}{\partial r}+\tfrac{1}{2}\Delta_y)u^\infty\big)(s,x+W^0_s-W^0_t)\big]\Big)\Big|_{s=\xi}\tfrac{[Q!]^4(T-t)^{2Q+1}}{(2Q+1)[(2Q)!]^3}.\end{aligned}\tag{64}$$

This and (63) prove that

$$\begin{aligned}&\sup_{\substack{t\in[0,T],\,u\in[0,t]\\z\in\mathbb{R}^d}}\Big\|\mathbb{E}\Big[\sum_{s\in[t,T]}q^{Q,[t,T]}(s)(F(u^\infty))(s,z+W^0_{u+s-t})-\int_t^T(F(u^\infty))(s,z+W^0_{u+s-t})\,ds\,\Big|\,W^0_u\Big]\Big\|_{L^2(\mathbb{P};\mathbb{R})}\\&\le\sup_{t\in[0,T)}\sup_{u\in[0,t]}\sup_{z\in\mathbb{R}^d}\left\{\left\|\sup_{s\in[t,T]}\Big(\tfrac{\partial^{2Q}}{\partial s^{2Q}}\mathbb{E}\big[\big((\tfrac{\partial}{\partial r}+\tfrac{1}{2}\Delta_y)u^\infty\big)(s,x+W^0_s-W^0_t)\big]\Big|_{x=z+W^0_u}\Big)\right\|_{L^2(\mathbb{P};\mathbb{R})}\tfrac{[Q!]^4(T-t)^{2Q+1}}{(2Q+1)[(2Q)!]^3}\right\}\\&\le\sup_{t\in[0,T]}\sup_{u\in[0,t]}\sup_{z\in\mathbb{R}^d}\left\{\left\|\sup_{s\in[t,T]}\big|\mathbb{E}\big[\big((\tfrac{\partial}{\partial r}+\tfrac{1}{2}\Delta_y)^{2Q+1}u^\infty\big)(s,x+W^0_s-W^0_t)\big]\big|\Big|_{x=z+W^0_u}\right\|_{L^2(\mathbb{P};\mathbb{R})}\tfrac{[Q!]^4(T-t)^{2Q+1}}{(2Q+1)[(2Q)!]^3}\right\}.\end{aligned}\tag{65}$$

Theorem 3.11 together with (65) implies (59). The proof of Corollary 3.12 is thus completed. □

The following result, Corollary 3.13, establishes an upper bound for the L^2-error between the solution of the PDE and our approximations (17) if the sup-norm of the n-th derivative of the solution of the PDE grows sufficiently slowly as $\mathbb{N}\ni n\to\infty$.

Corollary 3.13. *Assume the setting in Subsection 3.1, assume that $u^\infty\in C^\infty([0,T]\times\mathbb{R}^d,\mathbb{R})$, let $\alpha\in[0,1/4]$, and let $C\in[0,\infty]$ be the extended real number given by*

$$\begin{aligned}C&=L\Big[\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}|u^\infty(t,x)|\Big]+\Big[\sup_{x\in\mathbb{R}^d}|g(x)|\Big]+T\Big[\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}|(F(0))(t,x)|\Big]\\&\quad+Te^T\Big[\sup_{k\in\mathbb{N}}\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}(k!)^{\alpha-1}\big|\big((\tfrac{\partial}{\partial r}+\tfrac{1}{2}\Delta_y)^k u^\infty\big)(t,x)\big|\Big].\end{aligned}\tag{66}$$

Then it holds for all $M,Q\in\mathbb{N}$, $N\in\mathbb{N}\cap[0,2Q)$ that

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}\big\|U^0_{N,M,Q}(t,x)-u^\infty(t,x)\big\|_{L^2(\mathbb{P};\mathbb{R})}\le C(1+2L)^N\max\Big\{\tfrac{T^{2Q}}{Q^{2\alpha Q}},\tfrac{\exp(T\sqrt{M})}{M^{N/2}}\Big\}.\tag{67}$$

Proof. To prove (67) we assume w.l.o.g. that $C \in [0,\infty)$. Observe that the Stirling-type formula in Robbins [39, Displays (1)–(2)] proves for all $n \in \mathbb{N}$ that

$$\sqrt{2\pi n}\left[\frac{n}{e}\right]^n \leq n! \leq \sqrt{2\pi n}\left[\frac{n}{e}\right]^n e^{\frac{1}{12}} \tag{68}$$

This together with the fact that $\sqrt{e} \leq 2$ and the fact that $\forall\, n \in \mathbb{N}\colon \pi e^{\frac{1}{3}} n \leq 8^n$ shows for all $n \in \mathbb{N}$ that

$$\begin{aligned} \tfrac{n^{2\alpha n}((2n+1)!)^{1-\alpha}[n!]^4}{(2n+1)[(2n)!]^3} &\leq \tfrac{n^{2\alpha n}[n!]^4}{[(2n)!]^{2+\alpha}} \leq \tfrac{n^{2\alpha n}\left[\sqrt{2\pi}n^{n+\frac{1}{2}}e^{-n+\frac{1}{12}}\right]^4}{\left[\sqrt{2\pi}(2n)^{2n+\frac{1}{2}}e^{-2n}\right]^{2+\alpha}} = (\sqrt{2\pi})^{2-\alpha} n^{1-\frac{\alpha}{2}} e^{\frac{1}{3}+2n\alpha} 2^{-(2n+\frac{1}{2})(2+\alpha)} \\ &\leq 2\pi n e^{\frac{1}{3}+\frac{n}{2}} 2^{-4n-1} = \pi e^{\frac{1}{3}} n (\sqrt{e})^n 2^{-4n} \leq \pi e^{\frac{1}{3}} n 2^{-3n} \leq 1. \end{aligned} \tag{69}$$

Next note that Lemma 3.2 and (18) imply that for all $Q \in \mathbb{N}$, $i \in \{0,1,\ldots,2Q-1\}$ it holds that

$$\|u^\infty\|_{i,Q} \leq \left[\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |u^\infty(t,x)|\right]\left[\sum_{s\in[0,T]} \bar{q}^{i,Q}(s)\right] = \left[\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |u^\infty(t,x)|\right]\frac{T^i}{i!}. \tag{70}$$

The assumption that $C \in [0,\infty)$ allows us to apply Corollary 3.12 to obtain for all $M,Q \in \mathbb{N}$, $N \in \mathbb{N}\cap[0,2Q)$ that

$$\begin{aligned} &\sup_{t\in[0,T]}\sup_{z\in\mathbb{R}^d}\left\|U^0_{N,M,Q}(t,z) - u^\infty(t,z)\right\|_{L^2(\mathbb{P};\mathbb{R})} \\ &\leq \sup_{t\in[0,T]}\sup_{z\in\mathbb{R}^d}\sup_{u\in[0,t]}\left\|U^0_{N,M,Q}(t,z+W^0_u) - u^\infty(t,z+W^0_u)\right\|_{L^2(\mathbb{P};\mathbb{R})} = \left\|U^0_{N,M,Q} - u^\infty\right\|_{0,Q} \\ &\leq (1+2L)^{N-1}\Bigg\{L\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}|u^\infty(t,x)|\sup_{i\in\{0,1,\ldots,N\}}\tfrac{T^i}{i!\sqrt{M^{N-i}}} \\ &\quad+\sup_{i\in\{0,1,\ldots,N\}}\tfrac{T^i}{i!\sqrt{M^{N-i}}}\sup_{z\in\mathbb{R}^d}\left[\sup_{s\in[0,T]}\left\|g(z+W^0_s)\right\|_{L^2(\mathbb{P};\mathbb{R})} + T\sup_{r,u\in[0,T]}\left\|(F(0))(r,z+W^0_u)\right\|_{L^2(\mathbb{P};\mathbb{R})}\right] \\ &\quad+e^T\sup_{t\in[0,T]}\sup_{u\in[0,t]}\sup_{z\in\mathbb{R}^d}\left\|\sup_{s\in[t,T]}\left|\mathbb{E}\left[\left((\tfrac{\partial}{\partial r}+\tfrac{1}{2}\Delta_y)^{2Q+1}u^\infty\right)(s,x+W^0_{s-t})\right]\right|\Big|_{x=z+W^0_u}\right\|_{L^2(\mathbb{P};\mathbb{R})}\tfrac{[Q!]^4(T-t)^{2Q+1}}{(2Q+1)[(2Q)!]^3}\Bigg\} \\ &\leq (1+2L)^N\Bigg\{e^T T^{2Q+1}\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}\left|\left((\tfrac{\partial}{\partial r}+\tfrac{1}{2}\Delta_y)^{2Q+1}u^\infty\right)(t,x)\right|\tfrac{[Q!]^4}{(2Q+1)[(2Q)!]^3} \\ &\quad+\tfrac{1}{\sqrt{M^N}}\sup_{i\in\{0,1,\ldots,N\}}\left(\tfrac{(\sqrt{M}T)^i}{i!}\right)\left[L\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}|u^\infty(t,x)| + \sup_{x\in\mathbb{R}^d}|g(x)| + T\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}|(F(0))(t,x)|\right]\Bigg\}. \end{aligned} \tag{71}$$

This, (69), and the fact that $\sup_{i\in\{0,1,\ldots,N\}}\frac{(\sqrt{M}T)^i}{i!} \leq e^{T\sqrt{M}}$ imply for all $M,Q \in \mathbb{N}$, $N \in \mathbb{N}\cap[0,2Q)$ that

$$\begin{aligned} &\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}\left\|U^0_{N,M,Q}(t,x) - u^\infty(t,x)\right\|_{L^2(\mathbb{P};\mathbb{R})} \\ &\leq \tfrac{(1+2L)^N}{Q^{2\alpha Q}}e^T T^{2Q+1}\left[\sup_{k\in\mathbb{N}}\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}(k!)^{\alpha-1}\left|\left((\tfrac{\partial}{\partial r}+\tfrac{1}{2}\Delta_y)^k u^\infty\right)(t,x)\right|\right]\left[\sup_{n\in\mathbb{N}}\tfrac{n^{2\alpha n}((2n+1)!)^{1-\alpha}[n!]^4}{(2n+1)[(2n)!]^3}\right] \\ &\quad+\left(\tfrac{1+2L}{\sqrt{M}}\right)^N e^{T\sqrt{M}}\left[L\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}|u^\infty(t,x)| + \sup_{x\in\mathbb{R}^d}|g(x)| + T\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}|(F(0))(t,x)|\right] \\ &\leq \tfrac{(1+2L)^N}{Q^{2\alpha Q}}e^T T^{2Q+1}\left[\sup_{k\in\mathbb{N}}\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}(k!)^{\alpha-1}\left|\left((\tfrac{\partial}{\partial r}+\tfrac{1}{2}\Delta_y)^k u^\infty\right)(t,x)\right|\right] \\ &\quad+\left(\tfrac{1+2L}{\sqrt{M}}\right)^N e^{T\sqrt{M}}\left[L\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}|u^\infty(t,x)| + \sup_{x\in\mathbb{R}^d}|g(x)| + T\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}|(F(0))(t,x)|\right]. \end{aligned} \tag{72}$$

This establishes (67). The proof of Corollary 3.13 is thus completed. □

The next result, Corollary 3.14, provides an upper bound for the L^2-error between the solution of the PDE and our approximations (17) if the parameters $N,M,Q \in \mathbb{N}$ satisfy $N = M = Q$. Corollary 3.14 is a direct consequence of Corollary 3.13.

Corollary 3.14. *Assume the setting in Subsection 3.1, assume that* $u^\infty \in C^\infty([0,T] \times \mathbb{R}^d, \mathbb{R})$*, let* $\alpha \in [0, 1/4]$*, and let* $C \in [0, \infty]$ *be the extended real number given by*

$$\begin{aligned} C = L \left[\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |u^\infty(t,x)| \right] + \left[\sup_{x\in\mathbb{R}^d} |g(x)| \right] + T \left[\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |(F(0))(t,x)| \right] \\ + Te^T \left[\sup_{k\in\mathbb{N}} \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} (k!)^{\alpha-1} \left| \left((\tfrac{\partial}{\partial r} + \tfrac{1}{2}\Delta_y)^k u^\infty \right)(t,x) \right| \right]. \end{aligned} \tag{73}$$

Then it holds for all $N \in \mathbb{N}$ *that*

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \left\| U^0_{N,N,N}(t,x) - u^\infty(t,x) \right\|_{L^2(\mathbb{P};\mathbb{R})} \leq C \left[\frac{(1+2L)e^T}{N^{2\alpha}} \right]^N. \tag{74}$$

3.7 Analysis of the computational complexity and overall rate of convergence

In Lemma 3.15 $\mathrm{RN}_{n,M,Q}$ is the number of realizations of a scalar standard normal random variable required to compute one realization of the random variable $U^\theta_{n,M,Q}(t,x)\colon \Omega \to \mathbb{R}$. In Lemma 3.16 $\mathrm{FE}_{n,M,Q}$ is the number of function evaluations of f and g required to compute one realization of $U^\theta_{n,M,Q}(t,x)\colon \Omega \to \mathbb{R}$.

Lemma 3.15. *Assume the setting in Subsection 3.1 and let* $(\mathrm{RN}_{n,M,Q})_{n,M,Q\in\mathbb{Z}} \subseteq \mathbb{N}_0$ *be natural numbers which satisfy for all* $n, M, Q \in \mathbb{N}$ *that* $\mathrm{RN}_{0,M,Q} = 0$ *and*

$$\mathrm{RN}_{n,M,Q} \leq dM^n + \sum_{l=0}^{n-1} \left[QM^{n-l}(d + \mathrm{RN}_{l,M,Q} + \mathbb{1}_{\mathbb{N}}(l) \cdot \mathrm{RN}_{l-1,M,Q}) \right]. \tag{75}$$

Then for all $N \in \mathbb{N}$*, we have*

$$\mathrm{RN}_{N,N,N} \leq 8dN^{2N}.$$

Proof. Inequality (75) implies for all $n, Q \in \mathbb{N}$, $M \in \mathbb{N} \cap [2, \infty)$ that

$$\begin{aligned} (M^{-n} \cdot \mathrm{RN}_{n,M,Q}) &\leq d + \sum_{l=0}^{n-1} \left[QM^{-l}(d + \mathrm{RN}_{l,M,Q} + \mathbb{1}_{\mathbb{N}}(l) \cdot \mathrm{RN}_{l-1,M,Q}) \right] \\ &\leq d\left(1 + \tfrac{MQ}{M-1}\right) + (1 + \tfrac{1}{M})Q \left[\sum_{l=0}^{n-1} (M^{-l} \cdot \mathrm{RN}_{l,M,Q}) \right]. \end{aligned} \tag{76}$$

The fact that $\forall\, M, Q \in \mathbb{N}\colon \mathrm{RN}_{0,M,Q} = 0$ and the discrete Gronwall-type inequality in Agarwal [1, Corollary 4.1.2] hence prove that for all $n, Q \in \mathbb{N}$, $M \in \mathbb{N} \cap [2, \infty)$ it holds that

$$(M^{-n} \cdot \mathrm{RN}_{n,M,Q}) \leq d\left(1 + \tfrac{MQ}{M-1}\right)(1 + (1 + \tfrac{1}{M})Q)^{n-1} \leq \tfrac{d(M+(M+1)Q)^n}{M^{n-1}(M-1)}. \tag{77}$$

Hence, we obtain that for all $N \in \mathbb{N} \cap [2, \infty)$ it holds that

$$\mathrm{RN}_{N,N,N} \leq \tfrac{Nd}{N-1}(N + (N+1)N)^N = \tfrac{N}{N-1}(1 + \tfrac{2}{N})^N dN^{2N} \leq 8dN^{2N}. \tag{78}$$

This and the fact that $\mathrm{RN}_{1,1,1} \leq 2d$ complete the proof of Lemma 3.15. ☐

Lemma 3.16. *Assume the setting in Subsection 3.1 and let* $(\mathrm{FE}_{n,M,Q})_{n,M,Q\in\mathbb{Z}} \subseteq \mathbb{N}_0$ *be natural numbers which satisfy for all* $n, M, Q \in \mathbb{N}$ *that* $\mathrm{FE}_{0,M,Q} = 0$ *and*

$$\mathrm{FE}_{n,M,Q} \leq M^n + \sum_{l=0}^{n-1} \left[QM^{n-l}(1 + \mathrm{FE}_{l,M,Q} + \mathbb{1}_{\mathbb{N}}(l) + \mathbb{1}_{\mathbb{N}}(l) \cdot \mathrm{FE}_{l-1,M,Q}) \right]. \tag{79}$$

Then for all $N \in \mathbb{N}$*, we have*

$$\mathrm{FE}_{N,N,N} \leq 8N^{2N}.$$

The proof of Lemma 3.16 is analogous to the proof of Lemma 3.15 and therefore omitted. In the proof of Corollary 3.17 below we combine Lemma 3.15 and Lemma 3.16 with Corollary 3.14 to obtain a bound for the computational complexity of our scheme (17) in terms of the space dimension and the prescribed approximation accuracy.

The next result, Corollary 3.17, proves under suitable assumptions that if $\varepsilon \in (0, \infty)$ is the prescribed approximation accuracy and if $d \in \mathbb{N}$ is the dimension of the considered PDE, then for every $\alpha \in (0, 1/4]$ and every $\delta \in (0, \infty)$ it holds that the computational effort of the approximation method (number of function evaluations of the coefficient functions of the considered PDE and number of used independent scalar standard normal random variables, cf. Section 3.7) is at most $O(d\,\varepsilon^{-(\frac{1}{\alpha}+\delta)})$.

Corollary 3.17. *Assume the setting in Subsection 3.1, assume that $u^\infty \in C^\infty([0,T]\times\mathbb{R}^d,\mathbb{R})$, let $\alpha \in (0, 1/4]$, $\delta \in (0,\infty)$, let $C \in [0,\infty]$ be the extended real number given by*

$$\begin{aligned} C = 16\exp\Big(2\alpha\delta[e^T(1+2L)]^{\frac{1+\alpha\delta}{2\alpha^2\delta}}\Big)\Bigg\{L\Bigg[\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}|u^\infty(t,x)|\Bigg]+\Bigg[\sup_{x\in\mathbb{R}^d}|g(x)|\Bigg] \\ +T\Bigg[\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}|(F(0))(t,x)|\Bigg]+Te^T\Bigg[\sup_{k\in\mathbb{N}}\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}(k!)^{\alpha-1}\left|\big((\tfrac{\partial}{\partial r}+\tfrac{1}{2}\Delta_y)^k u^\infty\big)(t,x)\right|\Bigg]\Bigg\}^{1/\alpha+\delta}, \end{aligned} \tag{80}$$

let $(\mathrm{RN}_{n,M,Q})_{n,M,Q\in\mathbb{Z}}\subseteq\mathbb{N}_0$ be natural numbers which satisfy for all $n,M,Q\in\mathbb{N}$ that $\mathrm{RN}_{0,M,Q}=0$ and

$$\mathrm{RN}_{n,M,Q}\le dM^n+\sum_{l=0}^{n-1}\big[QM^{n-l}(d+\mathrm{RN}_{l,M,Q}+\mathbb{1}_{\mathbb{N}}(l)\cdot\mathrm{RN}_{l-1,M,Q})\big] \tag{81}$$

(for every $N\in\mathbb{N}$ we think of $\mathrm{RN}_{N,N,N}$ as the number of realizations of a scalar standard normal random variable required to compute one realization of the random variable $U^0_{N,N,N}(0,0)\colon\Omega\to\mathbb{R}$), and let $(\mathrm{FE}_{n,M,Q})_{n,M,Q\in\mathbb{Z}}\subseteq\mathbb{N}_0$ be natural numbers which satisfy for all $n,M,Q\in\mathbb{N}$ that $\mathrm{FE}_{0,M,Q}=$ and

$$\mathrm{FE}_{n,M,Q}\le M^n+\sum_{l=0}^{n-1}\big[QM^{n-l}(1+\mathrm{FE}_{l,M,Q}+\mathbb{1}_{\mathbb{N}}(l)+\mathbb{1}_{\mathbb{N}}(l)\cdot\mathrm{FE}_{l-1,M,Q})\big] \tag{82}$$

(for every $N\in\mathbb{N}$ we think of $\mathrm{FE}_{N,N,N}$ as the number of function evaluations of f and g required to compute one realization of the random variable $U^0_{N,N,N}(0,0)\colon\Omega\to\mathbb{R}$). Then it holds for all $N\in\mathbb{N}$ that

$$\mathrm{RN}_{N,N,N}+\mathrm{FE}_{N,N,N}\le Cd\left[\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}\big\|U^0_{N,N,N}(t,x)-u^\infty(t,x)\big\|_{L^2(\mathbb{P};\mathbb{R})}\right]^{-(1/\alpha+\delta)}. \tag{83}$$

Proof. We assume w.l.o.g. that $C\in[0,\infty)$. Throughout this proof let $\tilde{C}\in[0,\infty)$ be the real number given by $\tilde{C}=\frac{1}{16}\exp\Big(-2\alpha\delta[e^T(1+2L)]^{\frac{1+\alpha\delta}{2\alpha^2\delta}}\Big)\,C$. Corollary 3.14, Lemma 3.15, and Lemma 3.16 prove that for all $N\in\mathbb{N}$ it holds that

$$\begin{aligned} &(\mathrm{RN}_{N,N,N}+\mathrm{FE}_{N,N,N})\left[\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}\big\|U^0_{N,N,N}(t,x)-u^\infty(t,x)\big\|_{L^2(\mathbb{P};\mathbb{R})}\right]^{1/\alpha+\delta} \\ &\quad\le\big(8dN^{2N}+8N^{2N}\big)\,\tilde{C}\left[\frac{(1+2L)e^T}{N^{2\alpha}}\right]^{N(1/\alpha+\delta)}=8(d+1)\tilde{C}[(1+2L)e^T]^{N(1/\alpha+\delta)}N^{-2\alpha\delta N}. \end{aligned} \tag{84}$$

This and the fact that $\forall\,N\in\mathbb{N}\colon N!\le N^N$ show that for all $N\in\mathbb{N}$ it holds that

$$\begin{aligned} &(\mathrm{RN}_{N,N,N}+\mathrm{FE}_{N,N,N})\left[\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}\big\|U^0_{N,N,N}(t,x)-u^\infty(t,x)\big\|_{L^2(\mathbb{P};\mathbb{R})}\right]^{1/\alpha+\delta} \\ &\le 16d\tilde{C}\frac{[(1+2L)e^T]^{N(1/\alpha+\delta)}}{(N!)^{2\alpha\delta}}=16d\tilde{C}\left[\frac{[(1+2L)e^T]^{N\left(\frac{1+\alpha\delta}{2\alpha^2\delta}\right)}}{N!}\right]^{2\alpha\delta}\le 16d\tilde{C}\left[\sum_{n=0}^{\infty}\frac{[(1+2L)e^T]^{n\left(\frac{1+\alpha\delta}{2\alpha^2\delta}\right)}}{n!}\right]^{2\alpha\delta} \\ &=16d\tilde{C}\left[\exp\Big([(1+2L)e^T]^{\frac{1+\alpha\delta}{2\alpha^2\delta}}\Big)\right]^{2\alpha\delta}=16d\tilde{C}\left[\exp\Big(2\alpha\delta[(1+2L)e^T]^{\frac{1+\alpha\delta}{2\alpha^2\delta}}\Big)\right]=Cd. \end{aligned} \tag{85}$$

This completes the proof of Corollary 3.17. ☐

The next result, Corollary 3.18, specializes Corollary 3.17 to the case $\alpha = 1/4$.

Corollary 3.18. *Assume the setting in Subsection 3.1, assume that $u^\infty\in C^\infty([0,T]\times\mathbb{R}^d,\mathbb{R})$, let $\delta\in(0,\infty)$, let $C\in[0,\infty]$ be the extended real number given by*

$$\begin{aligned} C = 16\exp\Big(\delta[e^T(1+2L)]^{2+(8/\delta)}\Big)\Bigg\{L\Bigg[\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}|u^\infty(t,x)|\Bigg]+\Bigg[\sup_{x\in\mathbb{R}^d}|g(x)|\Bigg] \\ +T\Bigg[\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}|(F(0))(t,x)|\Bigg]+Te^T\Bigg[\sup_{k\in\mathbb{N}}\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}\frac{\left|\big((\frac{\partial}{\partial r}+\frac{1}{2}\Delta_y)^k u^\infty\big)(t,x)\right|}{(k!)^{3/4}}\Bigg]\Bigg\}^{4+\delta}, \end{aligned} \tag{86}$$

let $(\mathrm{RN}_{n,M,Q})_{n,M,Q\in\mathbb{Z}}\subseteq\mathbb{N}_0$ be natural numbers which satisfy for all $n,M,Q\in\mathbb{N}$ that $\mathrm{RN}_{0,M,Q}=0$ and

$$\mathrm{RN}_{n,M,Q}\le dM^n+\sum_{l=0}^{n-1}\big[QM^{n-l}(d+\mathrm{RN}_{l,M,Q}+\mathbb{1}_{\mathbb{N}}(l)\cdot\mathrm{RN}_{l-1,M,Q})\big] \tag{87}$$

(for every $N \in \mathbb{N}$ we think of $\mathrm{RN}_{N,N,N}$ as the number of realizations of a scalar standard normal random variable required to compute one realization of the random variable $U^0_{N,N,N}(0,0)\colon \Omega \to \mathbb{R}$), and let $(\mathrm{FE}_{n,M,Q})_{n,M,Q\in\mathbb{Z}} \subseteq \mathbb{N}_0$ be natural numbers which satisfy for all $n, M, Q \in \mathbb{N}$ that $\mathrm{FE}_{0,M,Q} =$ and

$$\mathrm{FE}_{n,M,Q} \leq M^n + \sum_{l=0}^{n-1} \left[QM^{n-l}(1 + \mathrm{FE}_{l,M,Q} + \mathbb{1}_{\mathbb{N}}(l) + \mathbb{1}_{\mathbb{N}}(l) \cdot \mathrm{FE}_{l-1,M,Q}) \right] \tag{88}$$

(for every $N \in \mathbb{N}$ we think of $\mathrm{FE}_{N,N,N}$ as the number of function evaluations of f and g required to compute one realization of the random variable $U^0_{N,N,N}(0,0)\colon \Omega \to \mathbb{R}$). Then it holds for all $N \in \mathbb{N}$ that

$$\mathrm{RN}_{N,N,N} + \mathrm{FE}_{N,N,N} \leq Cd \left[\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \left\| U^0_{N,N,N}(t,x) - u^\infty(t,x) \right\|_{L^2(\mathbb{P};\mathbb{R})} \right]^{-(4+\delta)}. \tag{89}$$

Acknowledgement

This project has been partially supported through the research grants ONR N00014-13-1-0338, DOE DE-SC0009248, and by the Deutsche Forschungsgesellschaft (DFG) via research grant HU 1889/6-1.

References

[1] Agarwal, R. *Difference Equations and Inequalities: Theory, Methods, and Applications.* Chapman & Hall/CRC Pure and Applied Mathematics. CRC Press, 2000.

[2] Albeverio, S., and Röckner, M. Stochastic differential equations in infinite dimensions: solutions via dirichlet forms. *Probability theory and related fields 89*, 3 (1991), 347–386.

[3] Bally, V., and Pagès, G. A quantization algorithm for solving multi-dimensional discrete-time optimal stopping problems. *Bernoulli 9*, 6 (2003), 1003–1049.

[4] Bellman, R. *Dynamic programming.* Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 2010. Reprint of the 1957 edition, With a new introduction by Stuart Dreyfus.

[5] Bender, C., and Denk, R. A forward scheme for backward SDEs. *Stochastic Processes and their Applications 117*, 12 (2007), 1793–1812.

[6] Bergman, Y. Z. Option pricing with differential interest rates. *Review of Financial Studies 8*, 2 (1995), 475–500.

[7] Bouchard, B., and Touzi, N. Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations. *Stochastic Processes and their applications 111*, 2 (2004), 175–206.

[8] Briand, P., and Labart, C. Simulation of BSDEs by Wiener chaos expansion. *Ann. Appl. Probab. 24*, 3 (06 2014), 1129–1171.

[9] Burgard, C., and Kjaer, M. Partial differential equation representations of derivatives with bilateral counterparty risk and funding costs. *C. Burgard and M. Kjaer. Partial differential equation representations of derivatives with counterparty risk and funding costs. The Journal of Credit Risk 7*, 3 (2011), 1–19.

[10] Chassagneux, J.-F. Linear multistep schemes for BSDEs. *SIAM Journal on Numerical Analysis 52*, 6 (2014), 2815–2836.

[11] Creutzig, J., Dereich, S., Müller-Gronbach, T., and Ritter, K. Infinite-dimensional quadrature and approximation of distributions. *Found. Comput. Math. 9*, 4 (2009), 391–429.

[12] Crisan, D., and Manolarakis, K. Solving backward stochastic differential equations using the cubature method: Application to nonlinear pricing. *SIAM Journal on Financial Mathematics 3*, 1 (2012), 534–571.

[13] Da Prato, G., and Zabczyk, J. Differentiability of the Feynman-Kac semigroup and a control application. *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl. 8*, 3 (1997), 183–188.

[14] Davis, P. J., and Rabinowitz, P. *Methods of numerical integration.* Courier Corporation, 2007.

[15] Duffie, D., Schroder, M., and Skiadas, C. Recursive valuation of defaultable securities and the timing of resolution of uncertainty. *Ann. Appl. Probab. 6*, 4 (1996), 1075–1090.

[16] E, W., Hutzenthaler, M., Jentzen, A., and Kruse, T. On multilevel Picard numerical approximations for high-dimensional nonlinear parabolic partial differential equations and high-dimensional nonlinear backward stochastic differential equations. *arXiv:1708.03223* (2017).

[17] Elworthy, K., and Li, X.-M. Formulae for the derivatives of heat semigroups. *Journal of Functional Analysis 125*, 1 (1994), 252–286.

[18] Giles, M. B. Improved multilevel Monte Carlo convergence using the Milstein scheme. In *Monte Carlo and quasi-Monte Carlo methods 2006.* Springer, Berlin, 2008, pp. 343–358.

[19] Giles, M. B. Multilevel Monte Carlo path simulation. *Oper. Res. 56*, 3 (2008), 607–617.

[20] Gobet, E., and Labart, C. Solving BSDE with adaptive control variate. *SIAM Journal on Numerical Analysis 48*, 1 (2010), 257–277.

[21] Gobet, E., Lemor, J.-P., and Warin, X. A regression-based Monte Carlo method to solve backward stochastic differential equations. *Ann. Appl. Probab. 15*, 3 (2005), 2172–2202.

[22] Graham, C., and Talay, D. *Stochastic simulation and Monte Carlo methods*, vol. 68 of *Stochastic Modelling and Applied Probability*. Springer, Heidelberg, 2013. Mathematical foundations of stochastic simulation.

[23] Gyöngy, I., and Krylov, N. Existence of strong solutions for Itô's stochastic equations via approximations. *Probab. Theory Related Fields 105*, 2 (1996), 143–158.

[24] Heinrich, S. Monte Carlo complexity of global solution of integral equations. *J. Complexity 14*, 2 (1998), 151–175.

[25] Heinrich, S. Multilevel Monte Carlo Methods. In *Large-Scale Scientific Computing*, vol. 2179 of *Lecture Notes in Computer Science*. Springer, 2001, pp. 58–67.

[26] Henry-Labordère, P. Counterparty risk valuation: a marked branching diffusion approach. *arXiv:1203.2369* (2012), 17 pages.

[27] Henry-Labordere, P., Oudjane, N., Tan, X., Touzi, N., and Warin, X. Branching diffusion representation of semilinear PDEs and Monte Carlo approximation. *arXiv preprint arXiv:1603.01727* (2016).

[28] Henry-Labordère, P., Tan, X., and Touzi, N. A numerical algorithm for a class of BSDEs via the branching process. *Stochastic Process. Appl. 124*, 2 (2014), 1112–1140.

[29] Hutzenthaler, M., and Jentzen, A. On a perturbation theory and on strong convergence rates for stochastic ordinary and partial differential equations with non-globally monotone coefficients. *arXiv:1401.0295* (2014), 1–41.

[30] Hutzenthaler, M., Jentzen, A., and Kloeden, P. E. Strong convergence of an explicit numerical method for SDEs with nonglobally Lipschitz continuous coefficients. *Ann. Appl. Probab. 22*, 4 (2012), 1611–1641.

[31] Karatzas, I., and Shreve, S. E. *Brownian motion and stochastic calculus*, second ed., vol. 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1991.

[32] Kebaier, A. Statistical Romberg extrapolation: a new variance reduction method and applications to option pricing. *Ann. Appl. Probab. 15*, 4 (2005), 2681–2705.

[33] Kloeden, P. E., and Platen, E. *Numerical solution of stochastic differential equations*, vol. 23 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1992. 632 pages.

[34] Maruyama, G. Continuous Markov processes and stochastic equations. *Rend. Circ. Mat. Palermo (2) 4* (1955), 48–90.

[35] Pardoux, É., and Peng, S. Backward stochastic differential equations and quasilinear parabolic partial differential equations. In *Stochastic partial differential equations and their applications (Charlotte, NC, 1991)*, vol. 176 of *Lecture Notes in Control and Inform. Sci.* Springer, Berlin, 1992, pp. 200–217.

[36] Pardoux, É., and Peng, S. G. Adapted solution of a backward stochastic differential equation. *Systems Control Lett. 14*, 1 (1990), 55–61.

[37] Peng, S. G. Probabilistic interpretation for systems of quasilinear parabolic partial differential equations. *Stochastics Stochastics Rep. 37*, 1-2 (1991), 61–74.

[38] Prévôt, C., and Röckner, M. *A concise course on stochastic partial differential equations*, vol. 1905 of *Lecture Notes in Mathematics*. Springer, Berlin, 2007. 144 pages.

[39] Robbins, H. A remark on stirling's formula. *The American Mathematical Monthly 62*, 1 (1955), 26–29.

[40] Skorohod, A. V. Branching diffusion processes. *Teor. Verojatnost. i Primenen. 9* (1964), 492–497.

[41] Yong, J., and Zhou, X. Y. *Stochastic controls*, vol. 43 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 1999. Hamiltonian systems and HJB equations.

[42] Zhang, J. A numerical scheme for BSDEs. *The Annals of Applied Probability 14*, 1 (2004), 459–488.